

Undergraduate Topics in Computer Science

Joseph Migga Kizza

Ethical and Secure Computing

A Concise Module

Third Edition



Undergraduate Topics in Computer Science

Series Editor

Ian Mackie, University of Sussex, Brighton, UK

Advisory Editors

Samson Abramsky , Department of Computer Science, University of Oxford, Oxford, UK

Chris Hankin , Department of Computing, Imperial College London, London, UK

Mike Hinchey , Lero – The Irish Software Research Centre, University of Limerick, Limerick, Ireland

Dexter C. Kozen, Department of Computer Science, Cornell University, Ithaca, NY, USA

Andrew Pitts , Department of Computer Science and Technology, University of Cambridge, Cambridge, UK

Hanne Riis Nielson , Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby, Denmark

Steven S. Skiena, Department of Computer Science, Stony Brook University, Stony Brook, NY, USA

Iain Stewart , Department of Computer Science, Durham University, Durham, UK

Joseph Migga Kizza, College of Engineering and Computer Science, University of Tennessee-Chattanooga, Chattanooga, TN, USA

‘Undergraduate Topics in Computer Science’ (UTiCS) delivers high-quality instructional content for undergraduates studying in all areas of computing and information science. From core foundational and theoretical material to final-year topics and applications, UTiCS books take a fresh, concise, and modern approach and are ideal for self-study or for a one- or two-semester course. The texts are all authored by established experts in their fields, reviewed by an international advisory board, and contain numerous examples and problems, many of which include fully worked solutions.

The UTiCS concept relies on high-quality, concise books in softback format, and generally a maximum of 275–300 pages. For undergraduate textbooks that are likely to be longer, more expository, Springer continues to offer the highly regarded *Texts in Computer Science* series, to which we refer potential authors.

Joseph Migga Kizza

Ethical and Secure Computing

A Concise Module

Third Edition



Joseph Migga Kizza
College of Engineering and Computer
Science
University of Tennessee-Chattanooga
Chattanooga, TN, USA

ISSN 1863-7310 ISSN 2197-1781 (electronic)
Undergraduate Topics in Computer Science
ISBN 978-3-031-31905-1 ISBN 978-3-031-31906-8 (eBook)
<https://doi.org/10.1007/978-3-031-31906-8>

1st edition: © Springer International Publishing Switzerland 2016
2nd & 3rd editions: © Springer Nature Switzerland AG 2019, 2023

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The COVID-19 pandemic brought tremendous changes in every aspect of our lives. From technological changes to workplace policies and etiquettes, life is no longer “normal”, as it used to be. We came out of the environment to face a changed environment and new realities to learn and cope with. The overwhelming growth of technology and the changed and still changing environment seem to have given us unlimited powers making us able to do unthinkable things just a few years before the pandemic. With or without the pandemic, tremendous technological advances have been registered across the board from telecommunication to computing with jaw dropping developments. From pandemic fears and resignation to chatbot excitement followed by fears, we are being jolted into the new and constantly changing realities of our time. Along the way, these developments are creating an unprecedented convergence of communications and computing platform technologies that are reaching into all remote corners of the physical space and beyond into the developing and yet to be metaverse virtual space. These new technological developments have created new communities and ecosystems that are themselves evolving, in flux and difficult to secure and with questionable, if not evolving ethical systems that will take us time to learn, if it remains constant at all. Because of these rapid and unpredictable changes, I found the second edition of *Ethics in Computing: A Concise Module*, needing a review and an update. Without losing my focus and flavor of the previous editions, I have selectively updated the content of the chapters, adding new ones, including the developing and eagerly awaited metaverse, and the exciting but risky chatbots, (including ChatGPT, Dell-E, Bing, and the likes), based on large language models (LLM), all clarifying the message that a time is coming, if not already here, when we, as individuals and as nations, will become totally immersed and dependent on the evolving computing technologies. Evidence of this is seen in the rapid convergence of telecommunication, broadcasting, computing technologies and mobile devices, the miniaturization of these devices, the ever-growing ubiquitousness of computing, the speed of computation, the emergency of virtualization, and immersive virtual reality. These technology characteristics have been a big pulling force sucking in millions of new users every day, sometimes even those unwilling. Other appealing features of technology are ever-growing pervasiveness and applications

both good and bad. Whether small or big, devices based on the growing ability of the changing technology have become a centerpiece of an individual's social and economic activities, the main access point for all information, and the empowerment of the device owners. Individuals aside, computing technology has also become the engine that drives the nations' strategic and security infrastructures that control power grids, gas and oil storage facilities, transportation, and all forms of national communication, including emergency services. These developments have elevated the cyberspace ecosystem and now the evolving virtual space—the metaverse, as the most crucial economic and security environments of nations requiring an *ethical and secure computing environment*.

As we look for ethical and secure computing strategies, the technological race is picking up speed with new technologies that make our efforts and existing protocols on which these strategies based obsolete in shorter and shorter periods. All these illustrate the speed at which the computing environment is changing and demonstrate a need for continuous review of our defensive strategies and more importantly a need for a strong *ethical and secure framework* in our computer, information, and engineering science education. This has been and will continue to be the focus of all my writings on this topic, and it is and remains so in this third edition.

Chapter Overview

This third edition is divided into fifteen chapters as follows:

Chapter 1—“Morality and the Law” defines and examines personal and public morality, identifying assumptions and values of the law, looking at both conventional and natural law, and the intertwining of morality and the law. It, together with Chap. 3, gives the reader the philosophical framework needed for the remainder of the book.

Chapter 2—“Ethics and Ethical Analysis” builds upon Chap. 2 in setting up the philosophical framework and analysis tools for the book discussing moral theories and problems in ethical relativism. Based on these and in light of the rapid advances in technology, the chapter discusses the moral and ethical premises and their corresponding values in the changing technology arena.

Chapter 3—“Ethics and the Professions” examines the changing nature of the professions and how they cope with the impact of technology on their fields. An ethical framework for decision-making is developed. Professional and ethical responsibilities based on community values and the law are also discussed. And social issues including harassment and discrimination are thoroughly covered.

Chapter 4—“Anonymity, Security, Privacy, and Civil Liberties” surveys the traditional ethical issues of privacy, security, and anonymity and analyzes how these issues are affected by computer technology. Information gathering, database, and civil liberties are also discussed.

Chapter 5—“**Intellectual Property Rights and Computer Technology**” discusses the foundations of intellectual property rights and how computer technology has influenced and continues to influence and change the traditional issues of property rights, in particular intellectual property rights.

Chapter 6—“**Social Context of Computing**” considers the three main social issues in computing, namely the digital divide, workplace issues like employee monitoring, and health risks, and how these issues are changing with the changing computer technology.

Chapter 7—“**Software Issues: Risks and Liabilities**” revisits property rights, responsibilities, and accountabilities with a focus on computer software. The risks and liabilities associated with software and risk assessment are also discussed.

Chapter 8—“**Computer Crimes**” surveys the history and examples of computer crimes, their types, costs on society, and strategies of detection and prevention.

Chapter 9—“**Cyberbullying, Cyberstalking and Cyber Harassment**” (rewritten) discusses the growing threat and effects repeated deliberate harm or harassment on other people by using electronic technology that may include devices and equipment such as cell phones, computers, and tablets as well as communication tools including social media sites, text messages, chat, and websites.

Chapter 10—“**Evolving Realities: Ethical and Secure Computing in the New Technological Spaces**” discusses the new frontiers of ethical and secure computing in the new technological spaces that include intelligent entities in cyberspace and their effects on the traditional ethical and social fabric of society.

Chapter 11—“**Ethical, Privacy, and Security Issues in the Online Social Network Ecosystems**” discusses the new realities of global computer social network ecosystems, global linguistic, cultural, moral and ethical dynamisms, and their impact on our traditional and cherished moral and ethical systems.

Chapter 12—“**Virtualization, Virtual Reality and Ethics**” (New) discusses virtualization and virtual reality technologies and how they inform our traditional moral and ethical values through immersive interaction via electronic media.

Chapter 13—“**Artificial Intelligence: Ethical and Social Problems of Large Language Models and the Future of Technology**”. Since its inception in 1956, artificial intelligence has been developed as a the theory upon which the processes of development of computer systems are able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and natural language processes. One type of platform, based on natural language processing, is the Large Language Models (LLM) that currently include ChatGPT, Dell-E, Bing, WebGPT, and others in development. LLMs give rise to chatbots—our humanoids. Our fears of these new humanoids are similar to those of Underhill, if not more. The chapter explores the excitement and fears of the AI chatbots and our possible options (New).

Chapter 14—“**Evolving Cyberspace: The Marriage of 5G and the Internet of Things (IoT) Technologies**” discusses the new frontiers of ethical and secure computing in the new and developing Internet-user interface whose protocols, policies, and standards are yet to be defined, discussed, and accepted by the scientific and user communities. We will explore how this new interface has created an ethical and security quagmire and how this is affecting our traditional ethical and social systems.

Chapter 15—“**Metaverse, the Evolving Realities and Ethics**” (New) discusses the expanding new frontiers of technology, an immersive 3D model of the Internet, supported by virtual and augmented reality that is drawing in large numbers of young technologists, gaming enthusiasts, online shoppers, small start-ups and giant tech companies and investors in NFT (non-fungible token), and the digital assets that will link the physical to virtual item ownership, such as works of art, real estate, music, or videos. With all this interest and rush, we explore the ethical and social responsibilities in this Metaverse.

Audience

The book satisfies the new computing sciences curricula (<https://www.acm.org/education/curricula-recommendations>). In particular the Computing Curricula (CC2020) consist of:

- Computer Engineering
- Computer Science
- Cybersecurity
- Information Systems
- Information Technology
- Software Engineering.

Associate-Degree Computing Curricula

- Associate-Degree Computing Curricula
- Information Technology Competency Model
- Computer Science Transfer
- Computer Engineering Transfer
- Software Engineering Transfer
- Kindergarten through 12th Grade.

CSTA K-12 CS Standards, 2017 Edition

In summary, all these curricula emphasize the student's understanding of the basic cultural, social, legal, and ethical issues inherent in the discipline of computing. To achieve this, the student must

- understand where the discipline has been, where it is, and where it is heading.
- understand the individual roles in this process, as well as appreciate the philosophical questions, technical problems, and aesthetic values that play an important part in the development of the discipline.
- develop the ability to ask serious questions about the social impact of computing and to evaluate proposed answers to those questions.
- be aware of the basic legal rights of software and hardware vendors and users, and they also need to appreciate the ethical values that are the basis for those rights.

Students in related disciplines like computer information and information management systems and library sciences will also find this book informative.

The book is also good for Computer Science practitioners who must practice the principles embedded in the curricula based on understanding:

- that the responsibility that they bear and the possible consequences of failure.
- their own limitations as well as the limitations of their tools.

The book is also good for anyone interested in knowing how ethical and social issues like privacy, civil liberties, security, anonymity, and workplace issues like harassment and discrimination are affecting the new computerized environment.

In addition, anybody interested in reading about computer networking, mobile computing, social networking, information security, and privacy will also find the book very helpful.

Chattanooga, TN, USA
2023

Joseph Migga Kizza

Acknowledgments I appreciate all the help I received from colleagues who offered ideas, criticism, sometimes harsh, and suggestions from anonymous reviewers over the years. Special thanks to my dear wife, Dr. Immaculate Kizza, who offered a considerable amount of help in proofreading, constructive ideas, and wonderful support.

Contents

1	Morality and the Law	1
1.1	Introduction	2
1.2	Morality	3
1.2.1	Moral Theories	4
1.2.2	Moral Decision Making	4
1.2.3	Moral Codes	5
1.2.4	Moral Standards	8
1.2.5	Guilt and Conscience	8
1.2.6	Morality and Religion	9
1.3	Law	10
1.3.1	The Natural Law	10
1.3.2	Conventional Law	11
1.3.3	The Purpose of Law	12
1.3.4	The Penal Code	12
1.4	Morality and the Law	13
1.5	Morality, Etiquettes, and Manners	15
	References	16
2	Ethics and Ethical Analysis	19
2.1	Traditional Definition	20
2.2	Ethical Theories	21
2.2.1	Consequentialism	22
2.2.2	Deontology	22
2.2.3	Human Nature	23
2.2.4	Relativism	23
2.2.5	Hedonism	23
2.2.6	Emotivism	23
2.3	Functional Definition of Ethics	24
2.4	Ethical Reasoning and Decision Making	26
2.4.1	A Framework for Ethical Decision Making	27
2.4.2	Making and Evaluating Ethical Arguments	27
2.5	Codes of Ethics	29
2.5.1	Objectives of Codes of Ethics	30

2.6	Reflections on Computer Ethics	30
2.6.1	New Wine in an Old Bottle	30
2.7	Technology and Values	33
	References	35
3	Ethics and the Professions	37
3.1	Introduction	38
3.2	Evolution of Professions	39
3.2.1	Origins of Professions	39
3.2.2	Requirements of a Professional	40
3.2.3	Pillars of Professionalism	42
3.3	The Making of an Ethical Professional: Education and Licensing	46
3.3.1	Formal Education	46
3.3.2	Licensing Authorities	47
3.3.3	Professional Codes of Conduct	48
3.4	Professional Decision Making and Ethics	51
3.4.1	Professional Dilemmas in Decision Making	51
3.4.2	Guilt and Making Ethical Decisions	53
3.5	Professionalism and Ethical Responsibilities	54
3.5.1	Whistle-Blowing	54
3.5.2	Harassment and Discrimination	57
3.5.3	Ethical and Moral Implications	58
	References	59
4	Anonymity, Security, Privacy, and Civil Liberties	61
4.1	Introduction	65
4.2	Anonymity	66
4.2.1	Anonymity and the Internet	66
4.2.2	Advantages and Disadvantages of Anonymity	67
4.2.3	Legal View of Anonymity	67
4.3	Security	68
4.3.1	Physical Security	68
4.3.2	Physical Access Controls	68
4.3.3	Information Security Controls	70
4.3.4	Operational Security	73
4.4	Privacy	73
4.4.1	Definition	73
4.4.2	Types of Privacy	74
4.4.3	Value of Privacy	75
4.4.4	Privacy Implications of the Database System	76
4.4.5	Privacy Violations and Legal Implications	78
4.4.6	Privacy Protection and Civil Liberties	80

4.5	Ethical and Legal Framework for Information	83
4.5.1	Ethics and Privacy	83
4.5.2	Ethical and Legal Basis for Privacy Protection	84
	References	85
5	Intellectual Property Rights and Computer Technology	87
5.1	Definitions	88
5.2	Computer Products and Services	88
5.3	Foundations of Intellectual Property	91
5.3.1	Copyrights	92
5.3.2	Patents	94
5.3.3	Trade Secrets	96
5.3.4	Trademarks	97
5.3.5	Personal Identity	99
5.4	Ownership	101
5.4.1	The Politics of Ownership	101
5.4.2	The Psychology of Ownership	102
5.5	Intellectual Property Crimes	102
5.5.1	Infringement	102
5.5.2	The First Sale Doctrine	104
5.5.3	The Fair Use Doctrine	104
5.6	Protection of Ownership Rights	105
5.6.1	Domain of Protection	105
5.6.2	Source and Types of Protection	105
5.6.3	Duration of Protection	106
5.6.4	Strategies of Protection	106
5.7	Protecting Computer Software Under the IP	107
5.7.1	Software Piracy	107
5.7.2	Protection of Software Under Copyright Laws	108
5.7.3	Protection of Software Under Patent Laws	109
5.7.4	Protection of Software Under Trademarks	110
5.7.5	Protection of Software Under Trade Secrets	110
5.8	Transnational Issues and Intellectual Property	111
	References	112
6	Social Context of Computing	115
6.1	Introduction	116
6.2	The Digital Divide	117
6.2.1	Access	118
6.2.2	Technology	126
6.2.3	Humanware (Human Capacity)	128
6.2.4	Infrastructure	129
6.2.5	Enabling Environments	130
6.3	Obstacles to Overcoming the Digital Divide	131

6.4	ICT in the Workplace	131
6.4.1	The Electronic Office	132
6.4.2	Office on Wheels and Wings	132
6.4.3	The Virtual Workplace	133
6.4.4	The Quiet Revolution: The Growth of Telecommuting	134
6.4.5	Employee Social and Ethical Issues	138
6.5	Employee Monitoring	139
6.5.1	Workplace Privacy and Surveillance	140
6.5.2	Electronic Monitoring	142
6.6	Employee Health and Productivity in the Workplace	145
6.6.1	Ergonomics	146
	References	149
7	Software Issues: Risks and Liabilities	151
7.1	Definitions	152
7.1.1	Standards	153
7.1.2	Reliability	154
7.1.3	Security	154
7.1.4	Safety	155
7.1.5	Quality	156
7.1.6	Quality of Service	156
7.2	Causes of Software Failures	157
7.2.1	Human Factors	157
7.2.2	Nature of Software: Complexity	158
7.3	Risk	158
7.3.1	Risk Assessment and Management	159
7.3.2	Risks and Hazards in Workplace Systems	160
7.3.3	Historic Examples of Software Risks	161
7.4	Consumer Protection	168
7.4.1	Buyers' Rights	168
7.4.2	Classification of Computer Software	170
7.4.3	The Contract Option	172
7.4.4	The Tort Option	174
7.5	Improving Software Quality	176
7.5.1	Techniques for Improving Software Quality	176
7.6	Producer Protection	177
	References	178
8	Computer Crimes	179
8.1	Introduction	180
8.2	History of Computer Crimes	182
8.3	Types of Computer Systems Attacks	184
8.3.1	Penetration	184
8.3.2	Denial of Service	185
8.4	Motives of Computer Crimes	186

8.5	Costs and Social Consequences	188
8.5.1	Lack of Cost Estimate Model for Cyberspace Attacks	190
8.5.2	Social and Ethical Consequences	192
8.6	Computer Crime Prevention Strategies	193
8.6.1	Protecting Your Computer	193
8.6.2	The Computer Criminal	194
8.6.3	The Innocent Victim	195
	References	196
9	Cyberbullying, Cyberstalking and Cyber Harassment	199
9.1	Definitions	200
9.1.1	Cyberbullying	201
9.1.2	Cyberstalking	201
9.1.3	Cyber Harassment	201
9.2	Types of Cyberbullying	202
9.2.1	Harassment	202
9.2.2	Flaming	202
9.2.3	Exclusion	202
9.2.4	Outing	203
9.2.5	Masquerading	203
9.3	Areas of Society Most Affected by Cyberbullying	203
9.3.1	Schools	203
9.3.2	Cyberbullying in the Workplace	204
9.4	Legislation Against Cyberbullying	204
9.4.1	Federal Laws	205
9.4.2	State Laws	205
9.4.3	International Laws	205
9.5	Effects of Cyberbullying	206
9.6	Dealing with Cyberbullying	207
9.6.1	Awareness	207
9.6.2	Legislations	207
9.6.3	Community Support	208
9.7	Resources	208
	References	210
10	Evolving Realities: Ethical and Secure Computing in the New Technological Spaces	211
10.1	Introduction	213
10.2	Artificial Intelligence	214
10.2.1	Advances in Artificial Intelligence	215
10.3	Cyberspace and the Concept of Telepresence	215
10.4	Securing Cyberspace	216
10.4.1	Detecting Attacks in Cyberspace	216
10.4.2	Vulnerability Scanning in Cyberspace	216
10.4.3	Privacy in Cyberspace	217

10.5	Social Issues in Cyberspace	218
10.6	Artificial Intelligence and Ethics	220
	References	223
11	Ethical, Privacy, and Security Issues in the Online Social Network Ecosystems	225
11.1	Introduction	226
11.2	Introduction to Computer Networks	226
11.2.1	Computer Network Models	227
11.2.2	Computer Network Types	228
11.3	Social Networks (SNs)	231
11.4	Online Social Networks (OSNs)	231
11.4.1	Types of Online Social Networks	232
11.4.2	Online Social Networking Services	233
11.4.3	The Growth of Online Social Networks	234
11.5	Ethical and Privacy Issues in Online Social Networks	236
11.5.1	Privacy Issues in OSNs	236
11.5.2	Strengthening Privacy in OSNs	240
11.5.3	Ethical Issues in Online Social Networks	240
11.6	Security and Crimes in Online Social Networks	244
11.6.1	Beware of Ways to Perpetrate Crimes in Online Social Networks	244
11.6.2	Defense Against Crimes in Online Social Networks	247
11.7	Proven Security Protocols and Best Practices in Online Social Networks	250
11.7.1	Authentication	251
11.7.2	Access Control	251
11.7.3	Legislation	251
11.7.4	Self-regulation	252
11.7.5	Detection	252
11.7.6	Recovery	252
	References	253
12	Virtualization, Virtual Reality and Ethics	255
12.1	Virtualization	256
12.2	Different Aspects of Virtualization	256
12.3	Virtualization of Computing Resources	256
12.3.1	History of Computing Virtualization	257
12.3.2	Computing Virtualization Terminologies	258
12.3.3	Types of Computing System Virtualization	259
12.3.4	The Benefits of Computing Virtualization	263
12.4	Virtual Reality/(Virtual Presence)	265
12.4.1	Different Types of Virtual Reality	267
12.5	Virtualization and Ethics	268
12.6	Social and Ethical Implication of Virtualization	270

12.7	Virtualization Security as an Ethical Imperative	270
12.7.1	Hypervisor Security	271
12.7.2	Securing Communications Between Desktop and Virtual Environment	272
12.7.3	Security of Communication Between Virtual Environments	272
12.7.4	Threats and Vulnerabilities Originating from a Virtual Environment	272
	References	274
13	Artificial Intelligence: Ethical and Social Problems of Large Language Models and the Future of Technology	275
13.1	Introduction	277
13.2	Definition	278
13.2.1	Language Models	278
13.2.2	Large Language Models (LLMs)	278
13.2.3	Problems with Large Language Models (LLMs)	279
13.3	Ethical and Social Problems with Large Language Models	279
13.3.1	Discrimination and Biases	280
13.3.2	Information Hazards	280
13.3.3	Misinformation	281
13.3.4	Privacy and Security Concerns	281
13.4	The Role of Big Technology Companies	282
13.5	Overcoming Challenges	282
13.5.1	Legislative Oversight	282
13.5.2	Sound AI Ethical Framework	283
13.6	Next Steps and the Future of AI	284
	References	284
14	Evolving Cyberspace: The Marriage of 5G and the Internet of Things (IoT) Technologies	287
14.1	Introduction	288
14.2	Fifth Generation (5G) Technology (G5)	288
14.2.1	Overview of 5G Wireless Communications	289
14.2.2	5G Network Architecture and Protocol Stack Perspectives	289
14.2.3	Technical Challenges of 5G Technology	291
14.3	The Internet of Things (IoT)	292
14.3.1	Overview and Growth of Internet of Things	294
14.3.2	Architecture and Networking of IoT	295
14.3.3	Challenges of Using TCP/IP Architecture Over the IoT	298
14.3.4	IoT Governance, Privacy and Security Challenges	301
14.3.5	Governance and Privacy Concerns	301
14.3.6	Security Challenges	302

14.3.7	Autonomy	303
14.3.8	Computational Constraints	303
14.3.9	Discovery	304
14.3.10	Trust Relationships	304
14.4	Ethical, Social and Legal Impacts of 5G and IoT	305
14.4.1	Environment	305
14.4.2	E-waste	305
14.4.3	Conflict Minerals	306
14.4.4	Healthy Issues Emanating from 5G and IoT Technologies	306
14.4.5	Ethics	306
	References	308
15	Metaverse, the Evolving Realities and Ethics	311
15.1	Introduction	312
15.2	Definition	313
15.3	The Evolution of Metaverse	313
15.4	The Architecture of Metaverse	313
15.4.1	Metaverse Avatars	315
15.4.2	Metaverse Platforms	315
15.5	Actualization of Metaverse	320
15.5.1	The Concept of Telepresence	320
15.5.2	Localization	321
15.6	Benefits of Metaverse	321
15.6.1	E-Commerce in Metaverse	321
15.6.2	Metaverse Billing and Payment Systems	321
15.7	Social and Ethical Concerns of the Metaverse	324
15.7.1	Metaverse Immersion and Autonomy	324
15.7.2	Metaverse and Transparency	324
15.7.3	Metaverse and Trustworthiness	324
15.7.4	Metaverse and Privacy	325
15.7.5	Metaverse and Access Inequality	325
15.7.6	Interoperability in Metaverse	325
15.7.7	Metaverse and Freedom of Expression	326
	References	327
Index	329



Morality and the Law

1

Abstract

Morality and the Law defines and examines personal and public morality, identifying assumptions and values and the law, looking at both conventional and natural law, and the intertwining of morality and the law. We define morality as a system that, in addition to setting standards of virtuous conduct for people, also consists of mechanisms to self-regulate through enforcement of the moral code and self-judge through guilt, which is an internal discomfort resulting from disappointment in self-mediated conscience. Based on this definition, we discuss moral theories, moral codes, moral standards and norms and how they are used to judge human actions to determine their goodness or badness. With the discussion of moral standards, we venture into concepts of guilt and conscience. We show how moral guilt is a result of self-judgment and punishment by an individual for not living up to the moral standards set for oneself or for the group. We end the chapter discussion with law noting that conventional laws of a society are anchored by the moral beliefs of that society. We look into the heated argument about this statement and observe that both morality and the legal system serve the purpose of keeping society stable and secure.

Learning Objectives

After reading this chapter, the reader should be able to

1. Learn to make sound moral reasoning.
2. Learn about moral values and ideals in a person's life.
3. Learn about the relationship between morality and religion.
4. Distinguish between morality and etiquette, law, and the professional code of conduct.

5. Learn what it means to have moral principles, the nature of conscience, and the relationship between morality and self-interest.

Scenario 1: With Stem Cell Research We Can Grow Just About Anything Human!

The parliament of the Republic of Kazini passed legislation, and the president signed into law, authorizing its citizens and scientists working on Kazini territory to carry out stem cell research to the best extent possible only limited by the physical resources. Scientists in Kazini have spearheaded such research and have made major breakthroughs in recent years.

Stem cells abound in bodies, but as human bodies age, the number of these cells and their potential and functions start to diminish as well. Embryonic stem cells that are found in the early stages of the body's development have the ability to divide indefinitely in culture and can therefore, at least in the laboratory, develop into virtually any cell type in the body.

The scientists in Kazini and their counterparts from around the world believe in the great benefits of stem cell research, especially embryonic stem cells. Many newspapers and scientific journals, not only in Kazini but also from other countries, have written stories of limitless benefits, the most immediate being the replacement of insulin-producing cells in the pancreas, damaged muscle cells, and dead nerve cells due to strokes, spinal injury, and degenerative diseases that include Alzheimer's and Parkinson's. It may also lead to the development and replacement of liver cells destroyed by a hepatitis and other liver diseases.

Dr. Don Rogan, a brilliant young scientist, is the director of Kazini Clinical Research Laboratory, the leading research nerve center in Kazini. Rogan is convinced that the legislature's action is morally wrong. However, his Laboratory has been chosen for funding and his dedicated scientists and staff are excited by the legislature's actions. They had lobbied hard for the passage of the bill. Now they see a ray of hope for millions of people not only on Kazini but also around the world. Rogan is facing a personal dilemma.

Discussion Questions

1. What options does Rogan have?
2. If you were Dr. Rogan, what would you do?
3. Is Dr. Rogan bound by the legislation?

1.1 Introduction

Whether you believe in a supreme being or you are an atheist, you acknowledge the existence of human life because you are alive. You are alive because someone nurtured you and protected you from all adversities. Whoever did so followed a

set of rules of conduct that kept both of you alive. Such shared rules, written or not, play a vital role in all human existence.

Human beings do not live randomly. We follow a script—a life script. In that script are hundreds of subscripts we follow both for survival (e.g., eating and sleeping) and for specific tasks. For example, when you meet a stranger, you follow a subscript different from the one you follow when you meet a long-lost friend. If you are hungry, the subscript you follow is different from the one you use to overcome anger. Within each subscript are variations we introduce to suit the situation. For example, when meeting an old friend, some people cry and others jump up and down, but both responses remain within the same subscript of meeting an old friend. The most important purpose of all these subscripts is human life, our own as well as others.

Believing in human life implies that we also believe life has a purpose. And because no one wants to live a life of pain, every human being believes in happiness as a purpose for life. To be happy, we need those conditions that create happiness, namely life, liberty, and property. Each condition is embodied in each of the three basic human survival subscripts: morality, ethics, and law. In this chapter, we discuss morality and law, and in Chap. 2 we discuss ethics.

1.2 Morality

Morality is a set of rules for right conduct, a system used to modify and regulate our behavior. It is a quality system in human acts by which we judge them right or wrong, good or bad. This system creates moral persons who possess virtues such as love for others, compassion, and a desire for justice; thus it builds character traits in people. In particular, morality is a survival script we follow in our day-to-day living. According to Wikipedia [1], morality has three different definitions:

A descriptive definition according to which morality means a set of rules (code) of conduct that governs human behavior in matters of right and wrong. An example of the descriptive usage could be “common conceptions of morality have changed significantly over time.”

A normative and universal definition that is more prescriptive and refers to an ideal code of conduct that would be observed by all rational people under specified conditions. An example is a moral value judgment such as “murder is immoral.”

A definition of morality that is synonymous with ethics. Ethics is the systematic philosophical study of the moral domain. We define and discuss ethics in the following chapter.

In each one of these definitions, morality concerns itself with a set of shared rules, principles, and duties, independent from religion, applicable to all in a group or society, and having no reference to the will or power of any one individual whatever his or her status in that group or society. Although moral values are

generally shared values in a society, the degree of sharing these values varies greatly. We may agree more on values such as truth, justice, and loyalty than on others. To paraphrase Shakespeare, life is but a stage on which there is continuous acting from the subscript of morality. Every time we interact in a society or group, we act the moral subscript that was developed by that society or group for its members over time.

Because morality is territorial and culturally based, so long as we live in a society we are bound to live within that society's guidelines. The actions of individuals in a society only have moral values if taken within the context of this very society and the culture of the individual. A number of factors influence the context of morality, including time and place.

1.2.1 Moral Theories

If morality is a set of shared values among people in a specific society, why do we have to worry about justifying those values to people who are not members of that society? In other words, why do we need moral theories? What do moral theories have to do with the moral subscripts? If you write a script for a play, you want both the audience and the cast to understand the message of the play. If you can find a way to help them get that message and believe it, then you have put credibility in the script. This is where moral theories come in. According to MacDonald, moral theories "seek to introduce a degree of rationality and rigor into our moral deliberations" [1]. They give our deliberations plausibility and help us to better understand those values and the contradictions therein. Because many philosophers and others use the words *moral* and *ethical* synonymously, we delay the discussion of moral theories until we discuss ethics.

1.2.2 Moral Decision Making

Every human action results from a decision process. Because every human action follows a subscript, the decision-making process follows a subscript as well. A decision is morally good if the result from it is good. A good moral decision embodies nearly all moral theories and usually takes into consideration the following points:

1. All the facts surrounding the situation, taking into account the interests of all parties involved, and
2. The moral principles involved and how they will affect all others involved.

Combining points 1 and 2 implies there must be reasoning and impartiality in any moral decision. Moral and ethical theorists have outlined four ways of ensuring reason and impartiality in moral decision making, as follows:

1. The use of the rational intuition of moral principles, which helps us perceive moral principles such as the notion of justice and deciding what is good.
2. The use of reason to determine the best way to achieve the highest moral good.
3. The ability to distinguish between primary and secondary moral principles. Primary moral principles are more general; secondary principles are more specific and are generally deduced from the primary ones.
4. The rational calculation of the consequences of our actions. The calculation should tell us whether the action is good or bad depending on the consequences [2].

Nearly all moral theories embody one or more of these themes.

1.2.3 Moral Codes

The *Internet Encyclopedia of Philosophy* defines moral codes as rules or norms within a group for what is proper behavior for the members of that group [2]. The norm itself is a rule, standard, or measure for us to compare something else whose qualities we doubt. Moral codes are often complex definitions of right and wrong that are based upon well-defined group's value systems.

In a way, moral codes are shared behavioral patterns of a group. These patterns have been with us since the beginnings of human civilization and have evolved mainly for the survival of the group or society. Societies and cultures survive and thrive because of the moral code they are observing. History has shown failures of societies and cultures such as the once mighty civilizations and great empires of the Babylonians, the Romans, and the Byzantines, probably because their code failed to cope with the changing times.

Although different cultures have different codes, and we have established that morality is relative to time, there have been some timeless and culture-free (moral) codes that have been nearly universally observed. Such codes include this partial list created by the astronomer Sagan [3]:

1. *The Golden Rule*: “Do unto others as you would have them do unto you.”

Versions of the Golden Rule in Different Religions¹

BUDDHIST: Hurt not others in ways that you would find hurtful.

CHRISTIAN: All things whatsoever ye would that men should do to you, do ye even so to them.

¹ <https://ellemay.wordpress.com/2009/02/09/the-golden-rule-versions-from-many-religions-philosophies/f/>.

CONFUCIAN: Do not unto others what you would not have them do unto you.

HINDU: This is the sum of duty; do naught unto others which if done to thee would cause thee pain.

ISLAMIC: No one of you is a believer until he desires for his brother that which he desires for himself.

JAIN: In happiness and suffering, in joy and grief, we should regard all creatures as we regard our own self.

JEWISH: Whatever thou hatest thyself, that do not to another.

SIKH: As thou deemest thyself, so deem others.

TAOIST: Regard your neighbor's gain as your own gain, and your neighbor's loss as your own loss.

ZOROASTRIAN: That nature alone is good which refrains from doing unto another whatsoever is not good for itself.

2. *The Silver Rule:* "Do not do unto others what you would not have them do unto you." Great men like Mahatma Gandhi followed this rule almost to the letter.
3. *The Bronze Rule:* "Repay kindness with kindness." This rule is widely observed because of its many varying interpretations. Some people call it the "carrot-and-stick" rule. However you interpret it, it seems to support the vendetta syndrome.
4. *The Iron Rule:* "Do unto others as you like, before they do it unto you." This rule, if followed by a leader, can create dictatorships. It seems to say, "He who is on the floor cannot make rules" or "Do it if you can get away with it."
5. *The Tin Rule:* "Pay homage to those above you and intimidate those below you." This is what many call the bully rule.
6. *The Nepotism Rule:* "Give precedence in all things to close relatives, and do as you like to others." This rule legitimizes corruption.

Because most of these rules seem vindictive, corruptible, dictatorial, and abusive, Sagan proposes the following as what seems to be a good culture-free and timeless universal set of moral codes:

1. Be friendly at first meeting.
2. Do not envy.
3. Be generous; forgive your enemy if he or she forgives you.
4. Be neither a tyrant nor a patsy.
5. Retaliate proportionately to an intentional injury (within the constraints of the rule of the law).
6. Make your behavior fair (although not perfectly) clear and consistent.

Other timeless, culture-free, but less widely practiced and less universally accepted codes are those observed by small groups of people with similar interests (e.g., religious and professional groups). Examples of such moral codes include the Native American Ten Commandments, the Jewish and Christian Ten Commandments, and the Unix Users Group Ten Commandments, as outlined here.

1.2.3.1 Native American Ten Commandments [4]

1. Treat the Earth and all that dwell thereon with respect.
2. Remain close to the Great Spirit.
3. Show great respect for your fellow beings.
4. Work together for the benefit of all Mankind.
5. Give assistance and kindness wherever needed.
6. Do what you know to be right.
7. Look after the well-being of mind and body.
8. Dedicate a share of your efforts to the greater good.
9. Be truthful and honest at all times.
10. Take full responsibility for your actions.

1.2.3.2 The Christian Ten Commandments [5]

1. I, the Lord, am your God. You shall not have any other gods besides Me.
2. You shall not take the name of the Lord, your God, in vain.
3. Remember to keep holy the Sabbath day.
4. Honor your father and your mother.
5. You shall not kill.
6. You shall not commit adultery.
7. You shall not steal.
8. You shall not bear false witness against your neighbor.
9. You shall not covet your neighbor's wife.
10. You shall not covet anything that belongs to your neighbor.

1.2.3.3 Unix Users Group Ten Commandments (The Manual, Ex. 20, Verses 1–21) [6]

And lo did Unix² speak these words upon the reboot:

1. Thou shalt use no other operating system than Unix.
2. Thou shalt not make unto thee a false operating system. Thou shalt not program them for I am the Unix and a jealous O/S.
3. Thou shalt not take the mark of trade of Unix in vain, or thou shalt be sued.
4. Remember thy password, and keep it secret.
5. Honour thy parent shell, for if it should die, thou shalt not live long (unless thou hast dissociated thyself).
6. Thou shalt not kill (l)-9 any process, for surely they shalt becometh zombies or defunct.
7. Thou shalt not commit hacking, else thou shalt eat quiche.

² Let Unix be a trademark of AT&T.

8. Thou shalt not use other users' data, for thou shalt be referred to the Data Protection Act, 1984, and sued (again).
9. Thou shalt not create Trojan horses, worms, viruses, or other foul beasts of false programming.
10. Thou shalt not rm-rf thy neighbor's home, nor covet his disk space allocation, nor his workstation account.

The purpose of moral codes in a society is to exert control over actions of members of the group resulting from emotions. Observance of moral codes in most societies is almost involuntary because members grow up with these codes, so they tend to follow them without questioning. In some societies, observance is enforced through superstition and in others it is enforced through folklore and customs. In Chap. 3, we show that professions need to have codes to which their members adhere for them to be ethical and moral in their day-to-day professional activities.

1.2.4 Moral Standards

A moral standard is a moral norm, a standard to which we compare human actions to determine their goodness or badness. This standard guides and enforces policy. Morality is a system that, in addition to setting standards of virtuous conduct for people, also consists of mechanisms to self-regulate through enforcement of the moral code and self-judge through guilt, which is an internal discomfort resulting from disappointment in self-mediated conscience.

1.2.5 Guilt and Conscience

Moral guilt is a result of self-judging and punishing oneself for not living up to the moral standards set for oneself or for the group. If individuals judge that they have not done “good” according to moral standards, they can activate the guilt response, which usually makes them feel bad, hide their actions from both self and others, and find a fitting punishment for themselves, sometimes a very severe punishment. This internal judgment system is brought about because human beings have no sure way of telling whether an action is good or bad based independently on their own “standards.” Individual standards are usually judged based on group standards. So individuals judge themselves based on group standards, and self-judgment comes into play whenever one’s actions fall short of the group’s standards.

The problem with guilt is that it can be cumulative. If individuals commit acts repetitively that they judge to be below moral standards, they tend to become more and more withdrawn. This isolation often leads individuals to become more comfortable with the guilt. As they become comfortable living with the guilt, their previous actions, which were previously judged below standards, begin to look not so bad after all. Individuals become more and more complacent about the guilt and begin to look at the whole moral system as amoral.

Guilt can be eased by encouraging people to focus on the intentions behind the actions. Sometimes the intentions may be good but the resulting action is bad. In such a case the individual should not feel so guilty about the action. Besides looking for intentions of actions, one should also have the will and ability to forgive oneself. Self-forgiveness limits the cumulative nature of guilt and hence helps an individual to keep within the group.

Our moral code, and many times the law, lay out the general principles that we *ought* not do this or that because it is wrong to do it. The law also tells us not to do this or that because it is illegal to do so. However, neither system specifically tells us whether a particular human action just committed is an immoral or illegal act. The link must be done by the individual—a self-realization. It is this individual inner judgment to tell us that the act is right or wrong, lawful or unlawful, that we call our *conscience*. Additionally, conscience is the capacity and ability to judge our actions ourselves based on what we set as our moral standards. The word *conscience* comes from the Latin word *conscientia*, which means *knowing with*. It is an “inner voice” telling us what to do or not do. This kind of self-judgment is based on the responsibility and control we have over our actions. Conscience is motivated by good feelings within us such as pride, compassion, empathy, love, and personal identification. Conscience evolves as individuals grow. The childhood conscience is far different from the adult conscience because our perception of evil evolves with age. The benefits of conscience are that the actions done with good conscience, even if they end up being bad, do not make one guilty of the actions.

Fr. Fagothey [7] writes that conscience applies to three things:

1. The intellect as a faculty for forming judgments about right and wrong individual acts
2. The process of reasoning that the intellect goes through to reach such judgment
3. The judgment itself, which is the conclusion of this reasoning process.

We have seen in this section that morality does not belong to any individual, nor does it belong to any society or group of people. Thus, it cannot be localized. However, those parts of the moral code that can be localized become law.

1.2.6 Morality and Religion

Religion, in contrast to morality, draws much from the divine. Most religious belief systems include or are built around the idea of divine will and divine judgment. However, many of these systems usually correspond to a moral code of conduct, and because of this, many religions claim that religion and morality are intimately connected.

Issues for Discussion

In Roman Catholicism, morality derives from God because God created man and nature and that the ultimate sanction for immorality is the loss of a relationship with God. How does your religion relate to the morality of your society?

How do both Atheism and Pantheism relate to morality?

What values are essential for a person that would allow him/her to starve rather than to steal?

1.3 Law

According to *Merriam-Webster's Dictionary*, a law is a binding custom or practice of a community; a rule of conduct or action prescribed or formally recognized as binding or enforced by a controlling authority; the whole body of such customs, practices, or rules [8]. Black states that law is an art we can create and model, and contemporary critics define law as an instrument of exercising power [9].

Bryan Bourn combines both these definitions of law and describes it as both an art and an instrument for exercising power [9]. He bases his definition on the fact that law on many occasions strives forcefully to create something desirable without following a precise and exact process or formula that can be reproduced (thus the art component). Fr. Fagothey defines laws as a rule and measure of actions directing them to proper ends. It obliges us to make our conduct conform to the norm of morality. He goes on to divide law into two types:

1. Physical law, which directs non-free irrational beings to uniform action toward their ends by inner necessity of their nature, that is, imposing physical necessity
2. Moral law or natural law, which directs free rational beings toward their ends by imposing obligations on the free will—thus imposing moral necessity.

However one defines law, whether as a rule, an injunction, an art, or an exercise of power, there is always a component of force that must be obeyed with the purpose of creating something desirable for the community that the law is intended to serve. This goal is achieved through the reign of equal justice for all in the community. We tend to obey two types of laws: the natural and the conventional.

1.3.1 The Natural Law

Natural law is an unwritten but universal law. It is a theory that an eternal, absolute moral law can be discovered by reason and is derivable from reason. It is distinct from the law of nature, applies to all rational creatures, exists independently of human preferences and inclinations, and is applied cross-culturally. According to

Donald [10], natural law “follows from the nature of man and the world, and consists of rights like the right to self-defense and the right to individual property. So naturally it is ‘higher’ than any other conventional law enacted by a human authority like a government because no conventional law has jurisdiction over natural law.” Natural law has been known since the time of Plato and Aristotle (ca. 500 BC) but has its clear formulation and definition in the writings of Thomas Aquinas, a thirteenth-century philosopher and theologian [1].

Natural law is the anchor of our rights of self-preservation, liberty, and property. Before organized human societies, humans existed because of natural law. It secured the environment in those human settlements for those activities that sustain life, beginning with hunting and progressing through business and commerce. Even today, there are human societies that exist without conventional law. Present-day examples include those states with collapsed governments because of political strife. People in these states, even in the absence of a central governing authority and a functioning legal system, are still living their lives, many of them happily. Although they may not enjoy all the pleasures of life, they have a way of protecting life, liberty, and personal property. Ironically, there are even states that supposedly live with organized authorities resembling government yet have no rule of conventional law; they are surviving on natural law.

The existence of natural law has been debated for centuries. In fact, there are many who do not believe in natural law and are always advocating the supremacy of conventional law. Thomas Hobbes, the famous English philosopher, argued that the nature of man is not such that one could deduce natural law from it, that the natural law so deduced does not place any significant limits on the powers of civil law, and that social order is a creation of state power [1].

1.3.2 Conventional Law

Conventional law is a system created by and for human beings, usually in public deliberations such as a council of elders or representatives in national legislatures. It derives from that part of the moral code that is enforceable and varies from society to society and from culture to culture. Although history and experience have shown that natural law has been used as the basis for some conventional laws, and there are examples such as the English Magna Carta and the U.S. Constitution and Bill of Rights; judgment is not based on natural law [10, 11]. In day-to-day judgment, decisions are based on facts and the matching of facts to words, not on natural law.

Conventional law takes two forms: (1) declarative, which simply restates what the natural law declares, such as forbidding murder, theft, etc., and (2) determinative, which fixes ways of acting in accordance with natural law, such as in contracts, taxes, traffic, and other types of laws. Conventional law has a

long history of evolution from natural law. Some of the outstanding examples follow [7]:

1. *Law of nature.* Originating from the Roman *jus gentium*. The Romans developed *jus gentium* from a mosaic of nations that formed the Roman Empire. *Jus gentium* was a common factor of all laws of all nations in the empire. When the empire collapsed, the resulting states developed this *law of nations* into the modern European legal system.
2. *English common law.* A result of centuries of unwritten precedents and decisions of common courts, statutes, and acts of the English Parliament.

The English common law gave birth to the modern English and American law.

1.3.3 The Purpose of Law

Both conventional and natural laws exist to protect the life, liberty, and property of the group covered by these laws. According to Fr. Fagothey [7], laws are needed for the following reasons:

1. The ignorant need instruction and control by the wise
2. Earthly penalties are required for the safety of society
3. Conceted action demands teamwork and leadership
4. Society must meet changed conditions harmoniously.

1.3.4 The Penal Code

Laws are always useless unless there is a right to punish and an enforcement mechanism is in place. The penal code is a system of set rules prescribing punishment for unlawful acts. In a way, the penal code is that enforcement mechanism. The punishment system consists of three functions [7]:

1. *Retributive*—by paying back the victim for the crime committed, reestablishing the equal balance of justice, and reasserting the authority.
2. *Corrective*—by trying to improve the offender; in other words, rehabilitating the offender back into society.
3. *Deterrent*—by trying to prevent similar actions in the future by the offender, and indeed the offender community, that is, forewarning the offender community by the state, which is the law maker.

The enforcement is different in criminal and civil cases. In criminal cases, the punishment may lead to denial of certain individual rights for a period of time. The

period of incarceration depends on the nature and types of violations. In civil cases, punishments are usually damage awards to those whose rights were infringed upon.

1.4 Morality and the Law

Conventional laws of a society are determined by the moral beliefs of that society. Many people disagree with this statement. In fact, there are two views. The proponents of natural law believe that conventional laws are valid if they meet certain standards of morality, whereas opponents of natural law, usually referred to as legal positivists, do not believe in the validity of conventional laws based on morality [7]. Whatever your camp, both morality and the legal system serve the purpose of keeping society stable and secure. They are both used in making judgments about people's actions, and such judgments are justifiable by reason. Although morality and the law seem to have a common purpose and the means to achieve the stated purpose, the implementation of these means to achieve the purpose is different. The following are some of the major differences:

1. *The process of making codes and laws:* Laws are enacted by authorities such as councils of elders and assemblies of the people's representatives. Moral codes, however, are developed not by one central authority but by all members of a society, over a period of time, from experiences and reason.
2. *Enforcement:* Laws are enforced by the authority that enacted them or representatives of that authority, such as judges and courts, and security forces such as the police. However, morality is self-enforced, not enforceable by courts, nor is it enforceable by any authorized security force. There is no moral or ethical court to judge moral wrongdoers. For example, no one can impose penalties for not obeying the Ten Commandments.
3. *Nature of punishments:* Unlawful acts are punishable by penalties that depend on the type, nature, and civility of the action. If it is criminal, it may result in incarceration, and if it is civil, it may result in payment of damages. However, if the act is judged to be immoral, the judgment is usually based on the individual's perception of that society's morality, and the penalties imposed are also individually based.
4. *Conflict resolution:* Laws are used to resolve interpersonal conflicts in a society. However, morality is mostly used to harmonize intrapersonal conflicts.
5. *Types of judgment:* Morality passes judgment on a person's intentions and character based on what is in your heart. Although courts do not always ignore a person's intention or state of mind, the law cannot normally govern what is in the person's heart.

Because of these differences, it is correct to say that in any society not all laws are based on the morality of that society. Because morality is a higher and superior system, there is only a small area where the two overlap, and there are many times when the two conflict. Let us look at examples. In February 1997 came the startling

news of the results of a bold genetic engineering experiment. The Roslin Institute in Edinburgh, Scotland, reported that a team of researchers led by embryologist Dr. Ian Wilmut had successfully cloned two identical sheep. Wilmut's team beat the odds predicted by researchers around the world by taking a mammary cell from an adult sheep, preparing its DNA to be accepted by the egg from another sheep, moving the egg's own DNA, and replacing it with the DNA from the adult sheep by fusing the egg with the adult cell. The fused egg began to grow normally to form an embryo, which scientists then implanted into another sheep, and that sheep later gave birth to a cloned lamb they named Dolly.

Although the experiment was done purely for animal reproduction, many scientists saw the potential for drug manufacturing and replacing human parts. Animals could be used to produce pharmacologically useful proteins for manufacturing drugs, literally making animals serve as drug factories. Animal clones could also be used to "manufacture" animal parts with human characteristics that could later be used in human transplants.

The cloning experiment created substantial legal, ethical, and moral problems. In many countries, it is not illegal to clone human beings, but because of the potential for abuse, such countries are already scrambling to enact laws that will make such an act illegal. Moral and ethical issues also need to be addressed. For example, what will prevent an unethical scientist from cloning a person he or she loves, or a person on whom to experiment, and what will stop governments strapped by lack of labor from cloning thousands of their best living human beings who have exhibited extraordinary intelligence or skills?

In the rush to create ourselves, we may end up creating monsters that could destroy us, because although the physical characteristics of clones will be similar, behavior characteristics will be as unpredictable as ours! Wilmut acknowledges the potential for misuse of this scientific breakthrough [11]. It is a daunting moral dilemma for which the society must find solutions.

Imagine seeing someone drowning and calling desperately for help while you simply look on and enjoy the show. Your action is not only repugnant, but immoral, and depending on whether the laws of deliberate indifference apply to you, your action may even be illegal. In another example, authorities in some societies fight teen violence by imposing a night curfew on the teens. In such societies, it is illegal for teens to venture out after curfew hours, although it is not immoral. Another good illustrative example is free speech. Consider a situation that occurred on a college campus in which a list of male students, posted by a group of female students led by a faculty member, warned that those male students were potential rapists. Such an act is repugnant, yet it is legal to post such a list. Consider also the trade in pornographic images both in print and on the Internet. These images not only degrade the men, women, and children depicted, they also contribute to other related crimes such as rape. In most cases, however, trading in such images is legal.

These examples illustrate that even though both morality and conventional law are integral parts of human life, they do not cover the same domains. There

are hundreds of similar cases where the legal system, although protecting civil liberties, unintentionally obscures morality.

Issues for Discussion

Name a few of what you consider to be unjust laws and sometimes unjust legal systems that imprison innocent people.

1.5 Morality, Etiquettes, and Manners

Etiquette refers to a code of behavior, a set of norms of correct conduct expected by a society, group, or social class. It is a generally expected social behavior. These rules of the code or the set of norms are usually unwritten, but aspects of these may reflect an underlying moral code.

Manners are unenforced standards of conduct or cultural norms that show that an individual is “refined” and “cultured” with a society or group. These norms codify or set a standard for human behavior. However, in contrast to laws that also codify human behavior, manners, just like morality, have no formal system for punishing transgressions other than social disapproval.

Issues for Discussion

Lapses in etiquettes, the consequences of which may vary depending on the audience, occur when least expected. Discuss these consequences and how etiquettes are related to the moral code of the group.

Discuss your own situations that involved such lapses. What does society expect from the offending individual?

Exercises

1. How do morality and law relate to each other?
2. What is moral relativism?
3. What is the connection between law and art?
4. Why is reasoning so important in morality?
5. Is morality evolutionary or revolutionary? Discuss.
6. Happiness is human. Discuss.
7. What is the role of education in moral behavior?
8. Show how and why the following rules are culture free:
 - (i) The Golden Rule
 - (ii) The Bronze Rule
 - (iii) The Iron Rule

9. If you were charged with creating a “new” human society, what moral code would you design and why?
10. We tend to live a moral script every day. Reflect on what is in your script.
11. Morality is time sensitive. Discuss.
12. Study the Native American Ten Commandments and the Christian Ten Commandments. Without comparing them, discuss the common thread between them.
13. How does guilt help to shape our moral journey?
14. Discuss the interplay between guilt and conscience.
15. What roles does the conscience fill in decision making?
16. Natural law is universal. Discuss.
17. What is the law of nature? Discuss why it is different from natural law.
18. What role does each one of the following have in our lives?
 - (i) Conventional law
 - (ii) Natural law
 - (iii) Law of nature
19. Can there be a common morality? Why or why not?
20. Is common morality possible in cyberspace?
21. Discuss the possibility of common morality in the age of globalization.
22. What is the effect of globalization on morality?

References

1. C. MacDonald, *Moral Decision Making: An Analysis*. <http://www.ethics.ubc.ca/~chrismac/moral.decision.html>
2. *Moral Relativism*, Internet Encyclopedia of Philosophy. <http://www.utm.edu/research/iep/ni/m-ration.html>
3. C. Sagan, A new way to think about rules to live by. Parade Magazine, 28 Nov 1993, p. 12
4. The Native American Ten Commandments, <http://www.indians.org/welker/tencomm.htm>
5. The Christian Ten Commandments, <http://biblescripture.net/Commandments.html>
6. The Unix Ten Commandments, <http://www.pipex.net/people/jasonh/command.html>
7. Fr. A. Fagothey, *Right and Reason*, 2nd edn. (Tan Books and Publishers, Rockford, IL, 1959)
8. Merriam-Webster’s Dictionary, <http://www.merriam-webster.com/dictionary/law>
9. B. Bourn, *Law as Art (with Apologies to Charles Black)*. <http://www.usinternet.com/bdbourn/black.html>
10. J. Donald, *Natural Law and Natural Rights*. <https://jim.com/rights.html>
11. G. Kalota, Scientists report first cloning ever for adult mammal. New York Times, 23 Feb 1997, sec. 1, p. 1

Further Reading

12. Conclusion: words, not laws, should be the weapons. The Ethical Spectacle, Nov 1995. <http://www.spectacle.org/1995/concl.html>
13. D.G. Johnson, *Computer Ethics*, 2nd edn. (Prentice Hall, Englewood Cliffs, 1994)

14. J.M. Kizza (ed.), *Social and Ethical Effects of the Computer Revolution* (McFarland, Jefferson, 1996)
15. D.R.J. Macer, *Bioethics for the People by the People* (Eubios Ethics Institute, Christchurch, 1994), pp. 74–91. <http://bio.tsukuba.ac.jp/~macer/BFPSE.html>
16. Objective Morality, <http://www.percep.demon.co.uk/morality.html18>



Ethics and Ethical Analysis

2

Abstract

This chapter builds upon Chap. 1 in setting up the philosophical framework and analysis tools for discussing moral theories and problems in ethical relativism. We discuss the moral and ethical premises and their corresponding values in the changing technology arena. In particular, we give two fitting definitions of ethics; the traditional definition of ethics and the functional definition of ethics as involving a value mapping. We discuss ethical decision making as a process of making a decision which may result in one or more moral conflicts. We end the chapter with a list of codes of ethics in use by different professional organizations.

Learning Objectives

After reading this chapter, the reader should be able to

1. Analyze an argument to identify premises and draw conclusions.
2. Illustrate the use of ethical argument.
3. Detect basic logical fallacies in an argument.
4. Identify stakeholders in an issue and our obligations to them.
5. Articulate the ethical trade-offs in a technical decision.
6. Evaluate professional codes of ethics for the ACM (Association for Computing Machinery) and other organizations.

Scenario 2: Should We Clone Humans?

Professor John Wesley is a brilliant scientist with an enviable track record of medical successes. In the last 5 years, he has carried out a dozen high-risk

medical operations successfully and has become a must-have on talk shows. He is a sought-after speaker on medical matters, and he is gifted on all reasonable subjects. He has led pioneering research in cloning and has been contemplating cloning some human replacement parts, if he can only get a human body to give him a convincing push.

Mrs. Joan Kaggwa is a well-known and successful entrepreneur, a wonderful wife, and a philanthropist. She is a president of several local and national charity organizations. She sits on the boards of several national and international corporations. For the last 21 years of her marriage, she has worked hard for her family and community. Two years ago, however, her only son, a young man nearing his 18th birthday, was killed in an automobile accident. He was the apple of his parents' eyes. The family was devastated by the death.

For a while now, Mrs. Kaggwa has been following the cloning stories that have appeared on television and in the newspapers, but without seriously giving them much thought until the day of her son's death. Then, with her instance, and to the annoyance of her husband, the family agreed to keep their son's body with Infinite Life Corporation, a company that keeps human frozen bodies in liquid nitrogen for years. Mrs. Kaggwa hoped that someday science would bring her son back. Her prayers were answered, at least according to her, one Sunday morning when she was going through the Sunday paper just before church. A small article caught her eye. The article was about a planned cloning experiment by a young scientist. During the following 2 weeks, Joan made calls that led her and her husband to the waiting room of Professor Wesley to discuss the cloning of their beloved, but dead, son.

Discussion Questions

1. Are there justifiable reasons that lead people to clone their loved ones?
2. Is Mrs. Kaggwa justified in wanting to clone her son?
3. Do you think the Kaggwas' son, if successfully cloned, will be the same as the dead son? Why or why not?
4. What compelling reasons can Professor Wesley give to justify cloning the Kaggwas' son?
5. Do you subscribe to such reasoning?
6. What are the pros and cons of human cloning?
7. Animal cloning is now routine. Why has there been no organized opposition to it?

2.1 Traditional Definition

Fr. Austin Fagothey, in *Right and Reason* [1], traces the origins of ethics from the Greeks. He observes that the Greeks' desire and curiosity to learn about themselves, the human life, and society led to the examination of all human conducts, a

part of philosophy called ethics. Ethics is, therefore, a study of right and wrong in human conduct. Ethics can also be defined as a theoretical examination of morality and as an equivalent of the “theory of morals.” Other philosophers have defined ethics in a variety of ways.

Robert C. Solomon, in *Morality and the Good Life* [2], gives a traditional philosophical definition of ethics as a set of “theories of value, virtue, or of right (valuable) action.” Johnson elaborates on Solomon’s definition by defining ethics as a set of theories “that provide general rules or principles to be used in making moral decisions and, unlike ordinary intuitions, provides a justification for those rules” [3]. The word “ethics” comes from an ancient Greek word *eché* [2], which means character. Every human society, whether civilized or primitive, practices ethics because every society attaches a value on an individual’s actions, on a continuum of good to bad, right to wrong, according to where that individual’s actions fall within the domain of that society’s rules and canons.

Ethics helps us not only in distinguishing between right and wrong but also in knowing why and on what grounds our judgment of human actions is justified. Ethics, therefore, is a field of inquiry whose subject is human actions, collectively called human conduct, which are performed consciously, willfully, and for which one can be held responsible. According to Fr. Fagothey [1], such acts must have knowledge that signifies the presence of a motive, voluntariness to signify that it is willed, and freedom to signify the presence of free choice to act or not to act.

The purpose of ethics is to interpret human conduct, acknowledging and distinguishing between right and wrong. The interpretation is based on a system. This system, according to Fr. Fagothey, uses a process of argumentation consisting of a mixture of inductions and deductions. In most cases, these arguments are based on historical schools of thought called ethical theories. There are different kinds of ethical theories, and within each theory there may be different versions of that theory.

2.2 Ethical Theories

For centuries, in different societies, human actions have been judged good or bad, right or wrong, based on theories or systems of justice developed, tested, revised, and debated by philosophers and/or elders in that society. Such theories are commonly known as ethical theories. Codes of ethics have then been drawn up based on these ethical theories. The processes of reasoning, explanation, and justification used in ethics are based on these theories. There are many ethical theories, but we consider only a few that are most widely discussed and used, namely, consequentialism, deontology, human nature, relativism, hedonism, and emotivism.

2.2.1 Consequentialism

In consequentialism ethical theory, human actions are judged good or bad, right or wrong, depending on the results of such actions—a desirable result denotes a good action and vice versa. There are three commonly discussed types of consequentialism theory:

1. *Egoism*: This theory puts an individual's interests and happiness above everything else. With egoism, any action is good so long as it maximizes an individual's overall happiness. There are two kinds of egoism: ethical egoism, which states how people ought to behave as they pursue their own interests, and psychological egoism, which describes how people actually behave. For example, if a family wanted to be happier, an ethical egoism theorist would prescribe to each family member how he or she ought to behave to achieve individual happiness first before considering the happiness of the family. A psychological egoism theorist, however, would describe how each individual family member should actually behave to achieve his or her happiness and hence the happiness of the family as a whole.
2. *Utilitarianism*: In contrast to egoism, this theory puts a group's interest and happiness above those of an individual, for the good of many. Thus, an action is good if it benefits the maximum number of people. Among the forms of utilitarianism are the following:
 - Act Utilitarianism: Tells one to consider seriously the consequences of all actions before choosing the one with the best overall advantage, happiness in this case, for the maximum number of people [4].
 - Rule Utilitarianism: Tells one to obey those rules that bring the maximum happiness to the greatest number of people. Rule utilitarianism maintains that a behavioral code or rule is good if the consequences of adopting that rule are favorable to the greatest number of people [4].
3. *Altruism*: In altruism, an action is right if the consequences of that action are favorable to all except the actor.

2.2.2 Deontology

The theory of deontological reasoning does not concern itself with the consequences of the action but rather with the will of the action. An action is good or bad depending on the will inherent in it. According to deontological theory, an act is considered good if the individual committing it had a good reason to do so. This theory has a duty attached to it. In fact, the word “deontology” comes from two Greek words, *deon*, meaning duty, and *logos*, meaning science [3]. For example, we know that killing is bad, but if an armed intruder enters your house and you kill him or her, your action is good, according to deontologists. You did it because you had a duty to protect your family and property.

2.2.3 Human Nature

This theory considers human beings as endowed with all faculties and capabilities to live in happiness. We are supposed to discover and then develop those capabilities. In turn, those capabilities become a benchmark for our actions, and our actions are then gauged and judged on how much they measure up to those capabilities. According to the famous Greek philosopher Aristotle, an individual committing an evil action is lacking in some capabilities.

2.2.4 Relativism

This theory is negatively formulated, denying the existence of universal moral norms. It takes right and wrong to be relative to society, culture, or the individual. Relativism also states that moral norms are not fixed in time.

2.2.5 Hedonism

Hedonism is one of the oldest ethical theories. It claims that pleasure is the only good thing in human life, the end of life as the highest good. A hedonist acts only for maximum pleasure, and whatever he or she does, it is done to maximize pleasure or minimize pain. There are two types of hedonism: psychological hedonism, which claims that in fact what people seek in their everyday actions is pleasure, and ethical hedonism, which claims that people ought to seek pleasure, and that pleasure is the moral good. Modern hedonists use the word pleasure to mean happiness.

2.2.6 Emotivism

This theory maintains that ethical statements are neither true nor false and cannot be proven; they are really only statements about how someone feels [4]. Philosophers use these theories as engines to help them to understand and justify human actions. Although over the years and in different places changing values have been attached to human actions, these ethical theories have remained relatively unchanged; this means that although ethics as a discipline is evolving, ethical reasoning has relatively remained the same. In other words, Aristotle and Plato's reasoning to explain and justify human actions is still valid, although the premises surrounding human actions are changing with time and with every new technology.

The process of ethical reasoning takes several steps, which we refer to as layers of reasoning, before one can justify to someone else the goodness or badness, rightness or wrongness, of one's action. For example, if someone wants to convince you to own a concealed gun, he or she needs to explain to you and justify why it is good to have a concealed gun. In such an exercise, the person may start by

explaining to you that we are living in difficult times and that no one is safe. You may then ask why no one is safe, to which the person might reply that there are many bad people out there in possession of high-powered guns waiting to fire them for various and very often unbelievable reasons. So owning a gun will level the playing field. Then you may ask why owning a gun levels the playing field, to which the answer would be that because if the bad guys suspect that you own one just like theirs, they will think twice before they attack you. You may further ask why this is so; the answer may be that if they attack you, they themselves can get killed in the action. Therefore, because of this fear you are not likely to be attacked. Hence, owning a gun may save your life and enable you to continue pursuing the ultimate concept of the good life: happiness.

On the other hand, to convince somebody not to own a concealed gun again needs a plausible explanation and several layers of reasoning to demonstrate why owning a gun is bad. Why is it a bad thing, you would ask, and the answer would be because bad guys will always get guns. And if they do, the possibility of everyone having a concealed gun may make those bad guys trigger-happy to get you before you get them. It also evokes the image of the Wild West filled with gun-toting people daring everyone to get a kick out of what may be a boring life. You would then ask why is this situation dangerous if no one fires? The reply might be because it creates a potential situation in which innocent people may get hurt and therefore an unhappy situation is created, denying people happiness and the good life. The explanation and reasoning process can go on and on for several more layers before one is convinced that owning a gun is good or bad. The act of owning a gun is a human act that can be judged as either good or bad, right or wrong, depending on the moral and ethical principles used.

The spectrum of human actions on which ethical judgments can be based is wide ranging, from simple traditional and easy-to-understand actions, such as killing and stealing, to complex and abstract ones, such as hacking, cellular telephone scanning, and subliminal human brain alterations. On one side of this spectrum, the inputs have straight output value judgments of right and wrong or good and evil. The other end of the spectrum, however, has inputs that cannot be easily mapped into the same output value judgments of good and bad or right and evil. It is at this side of the input spectrum that most new human actions created as a result of computer technology are found. It is at this end, therefore, that we need an updated definition of ethics—a functional definition.

2.3 Functional Definition of Ethics

Let $A = \{a_1, a_2, a_3, \dots, a_n\}$ be a collection of identifiable objects, $a_i, i = 1, 2, \dots, n$. We call this collection a set. A function f defined on set A is a rule that takes elements of A and assigns them values into another set R , called the range of the function. The set A is the domain of f . We represent the function f as $f: A \rightarrow R$. A function defined on two sets A and B takes pairs (a, b) of elements $a \in A$ and $b \in B$ and assigns to each pair a value r in the range set R . For example, let $A =$

$\{a_1, a_2, a_3\}$ and $B = \{b_1, b_2\}$. Then $f(A, B) \rightarrow C$ is a mapping $f(ai, bj) = rk$ for all $ai \in A$, $bj \in B$, and $rk \in C$ where $rk = aj * bj$ for some operation $*$ defined on elements of A and B .

An example of a function such as f would be the mixing of two colors. Suppose A is a can of blue paint and B is a can of yellow paint. Let f be the process of mixing these two colors from both cans. After mixing the contents or some of the contents of can A with those of can B , the resulting mixture is the green color put in can C .

Let us use this model to construct a functional definition of ethics. Let the set A be the set of all possible human actions on which it is possible to pass a value judgment. For example, if you think of an artwork, the human actions on it could be an array of things such as lifting it, hiding it, and stealing it. So define $A = \{a_1, a_2, a_3, \dots\}$. Let the second set B consist of many ethical or moral theories such as the ones we have discussed in the previous sections. So B could contain theories like egoism, act utilitarianism, and others. Define $B = \{b_1, b_2, b_3, \dots\}$. Finally, let R , the third set, be the set of all possible value judgments on the human actions in A based on the ethical theories in B . The function f maps each pair (a, b) of elements, with $a \in A$ and $b \in B$ to a binary value in R . The first set is the set of input parameters. The inputs are human actions on which it is possible to pass a judgment. The second set consists of the ethical theories discussed earlier, like consequentialism, deontology, and human nature. The third set $R = \{\text{RIGHT or WRONG, GOOD or BAD}\}$, the range of the function f on the two sets A and B , is the value set. Now define a function f on a pair of elements (a, b) with $a \in A$ and $b \in B$ to produce an element $r \in R$ as $f: (a, b) \rightarrow r$. We call this function the ethics decision function. Recalling our earlier discussion of ethics, function f represents a sequence of explanations and reasoning on the elements of sets A and B . The elements of R have two values: 1 for GOOD or RIGHT and 0 for WRONG or BAD.

Because the power of reasoning associates to each pair of elements (a, b) , with a in A and b in B , a binary value equivalent to good, bad, right, or wrong using the set B of ethical theories, we represent this function as follows:

$$f(a, b) \rightarrow \begin{cases} 1\{\text{"right," or "good"}\} \\ 0\{\text{"bad," or "wrong"}\} \end{cases}$$

for all $a \in A$ and $b \in B$

What is the relationship between this function, f , and the human mind? If you reflect on it you will see that the human mind seems to employ a similar function in making ethical and moral judgments. Notice that the human mind associates an integer value 0 or 1 to all human actions it perceives through sight, smell, touch, and hearing. Let us use an example to explain this. Suppose you see somebody breaking into a church. Chances are you are going to like or dislike that action. If you like it, you associate the “like” value to an integer 1 and if you dislike it,

you again associate this “dislike” value to an integer value 0. We tend to associate these two integer values to everything our mind perceives.

In making your decision whether you liked the action of the person who broke into the church, you probably based your “judgment” of this action on how that action registers in one of the moral and ethical theories. If it does, and it will always fall in at least one of the theories, then you will associate a weight depending on the hierarchy of reasoning you go through to justify right or wrong. Let us use this new functional model of ethics to get a glimpse into the prospects of ethics in the future. Advances in computer technology can greatly influence this model. An explanation is in order. The presence of computer technology creates multitudes of possibilities for every human action and greatly enhances and expands the input set A. The expansion of set A is likely to bring fuzziness to our traditional definition of ethics. Thus, there is a need to expand our definitions of ethics to meet the constantly changing technology-driven landscape of human actions.

2.4 Ethical Reasoning and Decision Making

Both reasoning and logic are important elements in daily human interactions. Reasoning is a human cognitive process of looking for ways to generate or affirm a proposition. Cognitive processes are mental functions or activities that are grouped based on *experience, interpretation, foreseeing, ordering, analyzing, valuing, and making connections*. Logic on the other hand, based on the Greek meaning, is the tool for distinguishing between truth and falseness. Human beings, on a daily basis, engage in reasoning and logic to achieve the desire results from a problem or an issue. Both reasoning and logic are important in decision making.

Each day we make hundreds of decisions, from what we will wear to what side of the bed to sleep on. When making these everyday decisions, many people tend to rely on a variety of biases and heuristics as they do their reasoning. This kind of reasoning based on intuition unfortunately leads to wrong and ethical decisions. Ethical reasoning is integrating ethical principles in the reasoning process. Each day we are faced with a variety of ethical or moral decisions, ranging from simple ones such as lying about a spouse’s choice of clothing to hard ones such as contributing to an abortion campaign.

Ethical decision making is the process of making a decision that results in a least number of conflicts. Such a process requires the decision maker to do the following [5]:

- Recognize the inherent ethical conflicts through comprehension, appreciation, and evaluation of all ethical dimensions of the problem
- Know the parties involved

- Be aware of alternatives
- Demonstrate knowledge of ethical practices
- Understand how the decision will be implemented and who will be affected
- Understand and comprehend the impact of the decision of the parties involved.

2.4.1 A Framework for Ethical Decision Making

Different elements make a good framework for an ethical decision.

The most common elements that must be in a good framework are these:

- Recognizing inherent ethical conflicts through comprehension, appreciation, and evaluation of all ethical dimensions of problem
- Understanding the problem and the facts of the problem
- Knowing the parties involved
- Being aware of alternatives
- Demonstrating knowledge of ethical practices
- Understanding how the decision will be implemented and who will be affected
- Understanding the impact the decision will have on the parties affected
- Understanding and comprehending the impact of the decision of the parties involved.

Taking these elements of the framework into consideration when making a decision lessens the number of conflicts and the severity of the impact resulting from the decision.

2.4.2 Making and Evaluating Ethical Arguments

In real life, especially in professional life, or in whatever we do, we are going to be faced with an ethical problem for which we need to seek solutions. Many real-life problems have systematic structures on which the search for a solution is based. For example, mathematical problems have rules called algorithms to follow. Many other real-life problems can be modeled in such a way that an algorithm can always be found, or in such cases where no mathematical formula can be used, empirical models can be used. Ethical problems are not like problems in a structured environment, where there are rules to follow. The main question is how to find solutions to ethical problems. We find solutions to ethical problems through a process, or series of steps, which often leads to an ethically justified resolution of the problem. Ethical reasoning either brings a resolution to an ethical problem or, at worst, helps to deepen our understanding of the ethical problem, which may eventually lead to the resolution of the problem at a future date. As we pointed out earlier, the process of ethical reasoning involves a set of layers of reasoning.

The process of ethical reasoning and ethical problem resolution can be likened to the process of software engineering. As in software engineering, the process goes through a number of stages with specific facts and responsibilities before a genuine solution to the problem is found. Before a resolution is embarked on, there must be a clear understanding of the problem. A clear picture of the relevant facts or specifications must be developed. A good description of these facts is then written down and guided by these facts; a set of layers of reasoning is entered into. Although the initial description of the problem is crucial, it should not be the last. As the reasoning process develops, the initial description could be revised and expanded, which may bring more understanding of the problem and may lead to the revision of our reasoning layers as further steps in the reasoning process are added or removed as additional information appears.

The process of ethical reasoning must avail the decision maker with a safe or valid alternative from a multitude of alternatives presented by the ethical problems. This safe alternative is the way out of the ethical muddles presented by the ethical problem. As the process of reasoning progresses, the following information will start to emerge [6]:

- (i) Information to confirm whether the problem is really an ethical problem
- (ii) Information on whether further description of the facts can add anything to the resolution process of the problem
- (iii) Information to identify the key ethical theories, principles, and values that fit the safe alternatives being pursued
- (iv) Information on the strength and validity of the ethical theory chosen and whether there are possible conflicts in the ethical theories, principles, and values with the reasoning processes and facts.

When a final decision has been made, an evaluation of that decision is needed.

The goal of evaluating an ethical argument is to make sure that each of the alternatives being considered is “weighted” against all others using the facts at hand developed earlier, and, in some cases, based on anticipated outcomes to our decisions. In so doing, we determine which alternative is best based on sound reasoning. Two outcomes are possible: one, we pick the best alternative, in which case our reasoning showed more validity of the facts of the problem than all other alternatives, or two, we may find that we are unable to determine a winning alternative. In this case, it means that there is no convincing reasoning in any one of the two or more deadlocking alternatives. This quandary may require any one of the following: the addition of more layers of reasoning, addition of new facts, or replacement of ethical theories and principles in the argument. In either of the two cases, however, justification of the choice of alternatives is based on examining all the reasons given for all the alternatives. A thorough examination of our reasoning is based on the criticism of the ethical reasoning used for each alternative. There

are several critical strategies used to achieve a good examination of the reasoning process, including whether the reasoning used was [6]:

- (i) Based on factual assumptions that are actually false or unsupported by good evidence. If assumptions are false or unsupported by any evidence, the reasons that make use of them are suspect and carry little weight, if any, or
- (ii) Valid. A reasoning is valid if its premises are true. Then the conclusion is also very probably true.

2.5 Codes of Ethics

The main domains in which ethics is defined are governed by a particular and definitive regiment of rules called “codes of ethics.” These rules, guidelines, canons, advisories, or whatever you want to call them, are usually followed by members of the respective domains. Depending on the domain, ethical codes can take any of the following forms:

- (i) Principles, which may act as guidelines, references, or bases for some document
- (ii) Public policies, which may include aspects of acceptable behavior, norms, and practices of a society or group
- (iii) Codes of conduct, which may include ethical principles
- (iv) Legal instruments, which enforce good conduct through courts.

Although the use of codes of ethics is still limited to professions and high-visibility institutions and businesses, there is a growing movement toward widespread use. The wording, content, and target of many codes differ greatly. Some codes are written purposely for the public; others are targeting employees, and yet others are for professionals only.

2018 ACM Code of Ethics and Professional Conduct: Draft 1

Draft 1 was developed by The Code 2018 Task Force. (It is based on the 1992 ACM Code of Ethics and Professional Conduct).

Preamble

Commitment to ethical professional conduct is expected of every member (voting members, associate members, and student members) of the Association for Computing Machinery (ACM). This code, consisting of 24 imperatives formulated as statements of personal responsibility, identifies the elements of such a commitment. It contains many, but not all, issues that professionals are likely to face. Section 2.1 outlines fundamental ethical considerations; Sect. 2.2 addresses additional, more specific considerations of professional conduct. Statements in Sect. 2.2 pertain more specifically to individuals who have a leadership role, whether in the workplace or in a volunteer capacity, for example, with organizations such as ACM. Principles involving compliance with this Code are given in Sect. 2.4.

Each imperative is supplemented by guidelines, which provide explanations to assist members in understanding and applying the imperative.

The Code is intended to serve as a basis for ethical decision making in the conduct of professional work. Secondarily, it may serve as a basis for judging the merit of a formal complaint pertaining to a violation of professional ethical standards.

The Code as a whole is concerned with how fundamental ethical imperatives apply to one's conduct as a computing professional. The imperatives are expressed in a general form to emphasize that ethical principles which apply to computing professionals are derived from broadly accepted ethical principles.

The Code is not an algorithm for solving ethical dilemmas. Words and phrases in a code of ethics are subject to varying interpretations, and a particular imperative may conflict with other imperatives in specific situations. Questions related to these kinds of conflicts can best be answered by thoughtful consideration of the imperatives and fundamental ethical principles, understanding that the public good is a primary consideration.

The rest of the code can be found at: <https://ethics.acm.org/2018-code-draft-1/>.

2.5.1 Objectives of Codes of Ethics

Different domains and groups of people formulate different codes of ethics, but they all have among them the following objectives:

1. Disciplinary: By instilling discipline, the group or profession ensures professionalism and integrity of its members.
2. Advisory: The codes are usually a good source of tips to members and offer advice and guidance in areas where there are fuzzy moral issues.
3. Educational: Ethical codes are good educational tools for members of the domain, especially the new ones who have to learn the do's and don'ts of the new profession. These codes are also a good source of renewal for the older members needing to refresh and polish their possibly waning morals.
4. Inspirational: Besides being disciplinary, advisory, and educational, the codes should also carry subliminal messages to those using them to inspire them to be "good."
5. Publicity: One way for professions to create a good clientele is to show that they have a strong code of ethics and, therefore, their members are committed to basic values and are responsible.

2.6 Reflections on Computer Ethics

2.6.1 New Wine in an Old Bottle

We have so far defined computer ethics as a subset of set A in the functional definition of ethics. We next elaborate on this by pointing out some likely differences between set A in the traditional definition and set A in the functional definition,

which now includes computer ethics. Although the overall picture remains the same, there are differences in the overall implementation of the models because of the changes in set A of the functional definition. These differences are manifested in several places, as discussed in the following sections.

2.6.1.1 Changing Premises

Although it is true that the outcome of the ethics value function remains the same, the domain set itself has changed and will keep changing. The number of input possibilities for every human action keeps on growing with new advances in computer technology. For example, take the act of forgery, which traditionally involves taking somebody's document, making changes to it, and getting a benefit as a result. Suppose the document is a check. Let us also assume, all other acts notwithstanding, that you have the document in your hand, in this case the check. Traditionally, your inputs were limited to making changes to the signature and probably changing the date, and cashing it meant literally walking to the financial institution either to deposit it or asking the teller to cash it after producing identification. Although these acts are still possible and readily accepted, new and cleverer ways have emerged as computer technology has advanced. First, now the inputs to set A of an act like this are numerous, no longer limited to the original two. They range from scanning the check to electronically reproducing almost an original check, to cashing it or depositing it without ever stepping in any financial institution, even in the late hours of the night. All these offerings were of course unheard of just a few years back, but they are giving thieves more ways to do their job and making it very difficult for financial institutions and law enforcement agents to do theirs.

2.6.1.2 Different Temptations

In traditional ethics, there were few temptations prompting unethical actions. But according to Rubin [7], computer technology has generated many more temptations for each input action. He outlines seven of these new temptations:

1. *Speed*: The speed of gathering information has greatly increased, causing unethical actions to be carried out in shorter times, thus decreasing the chances of detection. When the chances of being caught are slim, many perpetrators think that they can get away with it.
2. *Privacy and anonymity*: The great availability of computers and computer related technology in less visible places such as people's homes; high, cheap, and fast communication equipment; and software that can guarantee anonymity are creating a highly tempting environment for unethical acts.
3. *Nature of medium*: The ability to copy digital data without erasing or altering the original in any way causes little or no suspicion and hence encourages unethical activities.
4. *Aesthetic attraction*: Technology, especially when it is new, seems to offer challenges to those who try to use it. Thus, there is a sigh of relief and a sign of great achievement if one overcomes a technological obstacle. In the same way,

if an intruder tries to break into a computer system, the sign of success and the euphoria thereafter overshadows the incivility of the act itself.

5. *Increased availability of potential victims:* With the widespread use of computers and the ever-widening reach of computer networks, an individual can now reach an unprecedented audience. This scope in itself creates an urge to attempt things that one would otherwise not have done.
6. *International scope:* The boundary-less nature of many computer networks, including the Internet, has created a temptation of its own. Now the entire world is well within reach by a touch of a button. This accessibility can tempt many intruders, many trying to circumvent their country's laws, and others thinking that an illegal act done in another country cannot be prosecuted in their own country. There are many temptations here.
7. *The power to destroy:* Computers seem to give this enormous invisible power to those who have them. This seemingly omniscient power may be a temptation to some. Although some of these temptations can still be found in the set of the old temptations, most of them are new.

2.6.1.3 Different Means of Delivery

What used to be the traditional means of carrying out an act such as stealing has changed. With the expanded set of outcome possibilities come expanded delivery systems for the crime. For example, let us go back to the check. The traditional way of cashing a check was to go to the bank. With computers facilitating new ways of banking, you can get your check cashed without ever visiting the bank, even in the middle of the night.

2.6.1.4 Complacent Society

A majority of computer-related actions are either deliberately ignored by society for fear of publicity or they are hailed as novel science: either members of society are still caught in the spell of the new wonder machine or that they have gotten so comfortable with the new wonder machine that they let their moral and ethical standards slide. Whatever it is, society is too complacent about computers, and until this attitude changes, computer ethics is likely to remain different from traditional ethics.

2.6.1.5 Ethical Muddles

With the possibility of numerous inputs from events, new difficulties of choice and justification cause ethical dilemmas, creating conflicting arguments and counterarguments on an input possibility of an event. This situation occurs because computers produce new situations that sometimes fall within our existing laws, rules, and moral principles, and sometimes fall outside these guidelines.

2.7 Technology and Values

Every now and then, a new technology is introduced in our midst, intended to make our lives easier. Some of these technologies do not last for more than a month; others take hold and become revolutionary in magnitude. Those which become successful most often influence society by creating new possibilities that may raise new moral and ethical concerns and consequently create vacuums and new dilemmas in that society's basic sets of moral values. Computer technology has been one of these successful technologies. In its very short duration, it has had such a strong impact and influence on society, and if it continues the present trend unchecked, it is likely to become one of the greatest revolutions in the history of humankind, far greater than the agricultural and industrial revolutions. Society as a whole seems to be engulfed in this revolution and no cultural and/or society norm will, in the end if there is an end, be left unaffected.

Successful technological revolutions tend to create tempting situations that often result in a loosening of individual moral values, and the computer revolution tops that list. Worldwide cultural, political, and social underpinnings and values are undergoing a silent, but tremendous, change as new computer products come on the market and the revolution gathers momentum. It is moving so fast that it is stripping us of our ability to cope. Although we are constantly in need of new moral principles and new ethical values to fit the changing landscape, we cannot formulate, debate, and put in place such principles and values fast enough before they are outdated. More important still, even if we were able to come up with new values and moral principles, we would still lack the conceptual models within which such values and principles can be applied.

Many new situations resulting from the computer revolution are outdated our basic sets of values. Take, for example, the processes of handling forgeries in monetary currencies. There are laws on the books in almost every country against forgeries of any kind, let alone forgeries of currencies. These laws are further reinforced with individual moral values. One can, for example, reproduce and print millions of almost identical notes of a country's currency. Suppose even further that one produces a software program that reproduces the bank notes and enriches oneself. One's conscience of course tells the person that what one is doing is wrong, but the new technological advances are so tempting and making it so easy and so available that one can start rationalizing one's acts: I created or bought the program with my own money, I did all the work by myself, and after all it is highly unlikely that I can be caught because people cannot even tell the difference. All one is doing is creating a vacuum in one's basic set of values, and society needs to find a way to fill that moral vacuum so as to prevent individuals from taking moral vacations! As computer and telecommunication revolutions pick up speed, creating new avenues of use and access such as the Internet and the World Wide

Web, thus giving users room and reasons to take moral vacations, there is an urgent need to do the following:

1. Formulate new laws to strengthen our basic sets of values, which are being rendered irrelevant by computer technology.
2. Construct a conceptual model in which the new laws can be applied successfully.
3. Launch a massive education campaign to make society aware of the changing environment and the impact such an environment is having on our basic values.

The first two objectives are beyond the scope of this book, which mainly focuses on the third objective, educating the public concerning ethical issues raised by the computer revolution.

Issues for Discussion

“Thou shalt not kill.” What does this mean? When can you kill and it is OK? What can you kill and it OK?

Exercises

1. How would you define ethics to the following audiences?
Seventh-graders
College students
Members of the clergy
2. Why are acts such as abortion legal in some societies and not in others?
3. Does technology bring relevant changes in ethics?
4. Use the traditional model of ethics to explain the effects of technology on ethics to seventh-graders.
5. What are the merits of computer ethics education?
6. Why should we study computer ethics?
7. There are two views on teaching computer ethics. State these views. What view do you agree with and why?
8. Why do we need ethical theories?
9. According to the human nature theory, you are supposed to develop your capabilities, and your actions are based on those capabilities. If individuals have few developed capabilities (because of circumstances beyond their control, for example), should they be responsible for their actions?
10. Discuss the existence of universal moral norms.
11. Discuss the effects of time on moral norms.
12. Using graphics, demonstrate the working of the functional definition of ethics.
13. Professional organizations usually use professional codes of ethics to enforce discipline in their members. Do codes always work?

14. Suggest an alternative to the professional codes of ethics and demonstrate that your alternative can work.
15. How does technology affect ethics? morality?

References

1. Fr. A. Fagothey, *Right and Reason*, 2nd edn. (Tan Books and Publishers, Rockford, 1959)
2. R. Solomon, *Morality and the Good Life: An Introduction to Ethics Through Classical Sources*, 4th edn. (McGraw-Hill, New York, 2004)
3. D.J. Johnson, *Computer Ethics*, 4th edn. (Pearson Education, Inc., Upper Saddle River, NJ, 2009)
4. Encyclopedia of Philosophy, https://en.wikipedia.org/wiki/Encyclopedia_of_Philosophy
5. M. Velasquez, D. Moberg, M.J. Meyer, T. Shanks, M.R. McLean, D. DeCosse, C. André, K.O. Hanson, *A Framework for Thinking Ethically* (Markkula Center for Applied Ethics at Santa Clara University). <http://www.scu.edu/ethics/practicing/decision/framework.html>
6. T. Tomlinson, *Nursing Ethics* (Western Schools, 1993)
7. R. Rubin, Moral distancing and the use of information technology: the seven temptations, in *Social and Ethical Effects of the Computer Revolution*, ed. by J.M. Kizza (McFarland, Jefferson, 1996)

Further Reading

8. A. Edel, E. Flower, F.W. O'Connor, *Morality, Philosophy, and Practice: Historical and Contemporary Readings and Studies*, 3rd edn. (Random House, New York, 1989)



Ethics and the Professions

3

Abstract

Ethics and the Professions examines the changing nature of the professions and how they cope with the impact of technology on their fields. An ethical framework for decision making is developed. Professional and ethical responsibilities based on community values and the law are also discussed. Social issues including harassment and discrimination are thoroughly covered. Discussed in depth are the four pillars of professionalism that include commitment, integrity, responsibility, and accountability. We focus our discussion on professional dilemmas and guilt associated with decision making showing that these dilemmas, which are quite common in the everyday activities of a professional, are caused by questioning the values attached to one's premises as inputs to the decision being made; that one's input values may be clouded by conflicting codes of conduct, advances in technology, and/or incomplete or misleading information. We end the chapter with an in-depth discussion of professionalism and ethical responsibilities focusing on whistle-blowing, harassment, and discrimination.

Learning Objectives

After reading this chapter, the reader should be able to

1. Identify ethical issues that arise in professional decision making and determine how to address them.
2. Analyze global computing issues that influence professional decision making.
3. Describe the mechanisms that typically exist for day-to-day ethical decision making.
4. Identify progressive stages in a whistle-blowing incident.

5. Specify the strengths and weaknesses of relevant professional codes as expressions of professionalism and guides to decision making.

Real-Life Experiences: The Kansas City Pharmacist

In August 2001, Robert Courtney, a Kansas City pharmacist was indicted on 20 felony counts of product tampering, drug adulteration, and drug misbranding. Courtney illegally diluted Gemzar and other expensive chemotherapy drugs to make money.

What was more alarming was the fact that he had hundreds of cancer patients most of them relying on chemotherapy treatments for survival. According to the FBI, at least one patient who received the diluted drugs died.

Courtney was caught when a representative of Eli Lilly and Co., the pharmaceutical company that manufactures Gemzar, became suspicious from records that indicated that a Kansas City doctor was receiving much more Gemzar from Courtney's pharmacy than the actual amount of Gemzar the pharmacy was purchasing from the manufacturer.

After the doctor was notified and the drug was tested, U.S. federal agents were then informed. It was found that Courtney was selling up to three times the amount of drugs he was purchasing from the drug manufacturer [1].

Discussion Questions

1. What crime did Robert Courtney commit?
2. Was it proper to arrest Robert Courtney? Why or why not?
3. Do you think Robert Courtney was responsible for the assumed death?

3.1 Introduction

What is a profession? It is a trade, a business, or an occupation of which one professes to have extensive knowledge acquired through long years of experience and formal education and the autonomy of and responsibility to make independent decisions in carrying out the duties of the profession. To profess is to make a public declaration, a claim of something. In the case of a professional, that something is knowledge in the knowledge domain of that which makes up that occupation or trade. Webster's dictionary similarly defines *profession* as "a: a calling requiring specialized knowledge and often long and intensive academic preparation; b: a principal calling, vocation, or employment; c: the whole body of persons engaged in a calling" [2]. Well-known professions are law, medicine, and engineering.

In our study of professions and the people who profess the deep knowledge of the profession, we focus on four themes: (1) evolution of professions, (2) the making of an ethical professional, (3) the professional decision-making process, and (4) professionalism and ethical responsibilities. These four themes cover all

the activities of a professional life. First, we look at the beginnings of professions, describe the characteristics of professionals, and discuss how these characteristics are supported by commitment, integrity, responsibility, and accountability. We then describe the ways professionals are made: through both formal education and informal unstructured in-service. When professionals enter the workforce, their duties involve decision making. We therefore look at the process of decision making, the problems involved, and the guilt felt about what are perceived as wrong decisions and how to avoid them. Professionals in their working environment encounter problems every day that require them to check in with their moral code. We focus on professionalism and ethical responsibilities as one of those areas that requires continual consultation with individual morality and discuss how these affect professions.

3.2 Evolution of Professions

3.2.1 Origins of Professions

The concept of a profession is actually not new; however, the word *profession* today carries a far different connotation than it did during the Middle Ages. The word *profession* referred to a commitment formally *professed* by a person to become a member of a religious order, and a *professional* was the person who has *professed* the commitment. Senior writes, because early universities drew most of their faculty from religious orders, these teachers eventually were called *professors*. Sizer [3] states that professions started in medieval times with the craftsmen's guilds and in inns. These guilds were responsible for apprenticeship standards, competence, and performance of their members. Little distinction was made between manual labor and intellectual groups. But as small intellectual groups developed, such as those of clerics, the first requirements of achievements and maintenance of professional criteria started to emerge. The emphasis on intellectual capabilities for membership in a group became increasingly important as time passed. Sizer states that professions in eighteenth-century England were regarded as "occupations for the 'gentlemen,' offering a safe social niche but not large material rewards." The Industrial Revolution is credited with establishing professions in engineering, accounting, and banking [3]. Over the years, however, material rewards have increased and a set of requirements has evolved.

Over the years, the term profession and its requirements for membership evolved into two categories: the *learned* professions, which required individuals with a deep knowledge of the profession acquired through years of formal education; and *common* professions, which required the individuals to be noblemen who in theory did not really need to work for a living: they were *liberated* from the need to work, but should learn the profession anyway. The first liberal profession was the military career [3]. When the life of the nobility became less influential, especially after the French Revolution, the *common* distinction of professions

came to be known as *trades*, probably as we know them today. However, *trades*, as today, still required one to hold a higher ethical standard.

3.2.2 Requirements of a Professional

There are three basic professional requirements, and over the years as the professions evolved, these three elements have taken different forms.

1. *A set of highly developed skills and deep knowledge of the domain.* Although professional skills are developed through long years of experience, such skills must be backed up by a very well developed knowledge base acquired through long years of formal schooling. Acquiring a sophisticated level of knowledge is crucial because skills based on shallow knowledge of the domain could be damaging to the profession in cases involving decisions that require understanding, analysis, and adoption of concepts to suit the environment or the problem. This requirement alone is enough to differentiate between professionals and skilled laborers who acquire considerable skills from long years of working in the same domain, such as auto mechanics and landscape designers.
2. *Autonomy.* Because professionals provide either products or services, there is always a relationship between the provider of the service and the receiver of the service or the provider of the product and the receiver of the product. In this relationship we are concerned with the power balance. In the case of a professional, the power is in favor of the professional. Take the relationship between a lawyer and a client or a physician and a patient, for example. In either case, there is a power play in favor of the provider of the service. If we consider the example of an auto mechanic, however, there is also a power play in the equation, but this time the power is in favor of the customer, not the provider of the service. There are also marked differences in the way the service is provided by professionals and nonprofessionals. In the case of a professional, there is more room to vary the way a service or a product is provided without consulting the receiver of the service or the product, meaning that professionals have autonomy to vary the way the service is provided without asking the receiver for confirmation or consent. However, in the case of nonprofessionals, the provider of the service cannot vary the way the service is to be delivered without checking with the customer. For example, when you take a car for repair, the mechanic cannot vary from what you agreed on without formally asking you.
3. *Observance of a code of conduct.* A working professional usually observes these four types of codes [4]:
 - *The professional code:* a set of guidelines provided to the professional by the profession spelling out what a professional ought to do and not do. A professional code protects both the image of the profession and that of the individual members. Thus, it is a requirement for the profession that members adhere to the code.

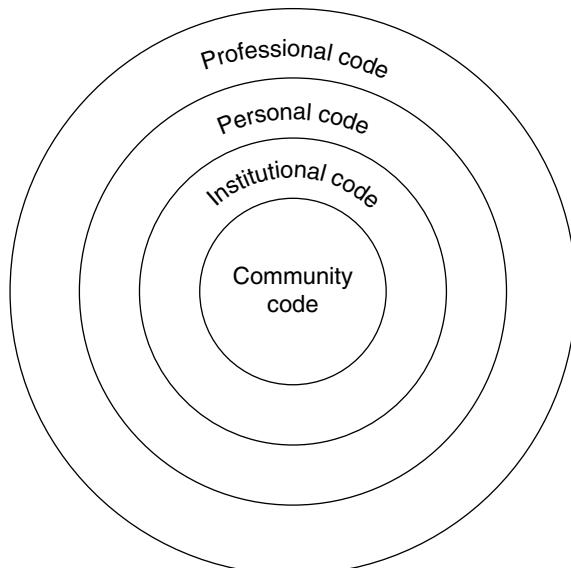
- *A personal code*: a set of individual moral guidelines on which professionals operate. In many ways these guidelines are acquired by professionals from the cultural environment in which they grow up or live in and the religious beliefs they may practice. Whatever the case, a personal code supplements the professional code significantly.
- *The institutional code*: a code imposed by the institution for which the professional is working. This code is meant to build and maintain the public's confidence in the institution and its employees.
- *The community code*: a community standard code developed over a period of time based on either the religion or culture of the indigenous people in the area. It may be imposed by civil law or the culture of the community in which the professional works.

The interaction among the four codes can be explained as follows: consider each code as a circle inside another circle with the community code at the center of these concentric circles. Outside the community code is the institutional code enclosed by the personal code, which is enclosed by the professional code (see Fig. 3.1).

Any action performed by a professional working at a local institution is contained in the outermost circle. Therefore, for such action to be ethical, moral, and legal, it must be in conformity with all the codes and intersect all codes.

Let us consider an example. Suppose a physician is working in a community hospital where the hospital and the community do not support euthanasia. If the doctor is asked by his or her patients to assist them in taking their own life, the doctor must examine all four codes before coming to a decision. First, the professional code may not support euthanasia whether the doctor's individual moral

Fig. 3.1 Codes governing human actions



code does or does not. So, because the institutional, community, and the professional codes do not support euthanasia, the doctor may not find it in his or her best interest to grant the patients their wishes even if he or she agrees with the patient. As we discuss later, the requirement that any action taken by a professional must fall within the intersection of the four sets of codes may present moral dilemmas for the professional in the decision-making process and consequently tarnish the professionalism of the individual.

3.2.3 Pillars of Professionalism

Professionalism is supported by four pillars: commitment, integrity, responsibility, and accountability.

3.2.3.1 Commitment

Commitment, according to Humphreys, has these six characteristics [5]:

1. *The person making the commitment must do so willingly without duress.* The person executing the commitment must like what he or she is doing. If commitments are in the form of assignments with little autonomy, it is more likely the commitment may not be present.
2. *The person responsible must try to meet the commitment, even if help is needed.* Because commitments are not assignments, the person who has made the commitment is assumed to have the know-how, the autonomy to vary steps, and the skills to do the job. Professionals possess these characteristics, plus they have the ability to seek the necessary skills from others to circumvent obstacles that may arise, so more commitment is expected of them.
3. *There must be agreement on what is to be done, by whom, and when.* Professionals entering into a commitment must have advance knowledge of what is to be done and who is likely to do what part. Entering into a commitment without adequate advance knowledge is highly unprofessional. When the work is divided among other professionals, they themselves must make the same commitment for their respective parts and, in this case, commitment for those smaller parts is as important as the commitment for the whole job. If the smaller parts are assigned to nonprofessionals, they are considered assignments, and the commitment must lie with the professional assigning the parts. Such commitment is carried out through supervision of the nonprofessional members of the team.
4. *The commitment must be openly and publicly stated.* Open commitments are transparent and easily correctable if there are problems. Professional commitments must fall within the allocated resources of time, material, and money. If a commitment is public, there are more chances that most of the sourcing, acquisition, distribution, and use of resources will be transparent, and thus the job is likely to be done more smoothly.
5. *The commitment must not be made easily.* Before entering into a commitment, professionals should do research to make sure that what they are entering into

is not a Trojan horse (something or someone intended to defeat or subvert from within).

6. *Before the committed date, if it is clear it cannot be met, advance notice must be given and a new commitment negotiated.* It is a sign of responsibility and commitment to have the courage to tell others of shortfalls in parts of the agreement, so if there is anything to be done to meet the deadlines, it is done without acrimony.

3.2.3.2 Integrity

Integrity means a state of undivided loyalty to self-belief. It is honesty, uncompromising self-value, and incorruptible. The word “integrity” comes from the Latin word *integratas*, which means entire, undivided, or whole. To stay undivided in one’s beliefs professionally requires three maxims of integrity: namely, vision, love of what one is doing, and commitment to what one has to do.

Vision. Having vision is the capacity to anticipate and make a plan of action that will circumvent obstacles and maximize benefits. Vision is a sign of good leadership, and professionals who have the initiative, the autonomy, and the authority in the provider-client relationship exemplify leadership.

Love. Numerous studies have shown that people who love what they do do it better than those who do it because they have to. In school, children who have a love for a subject perform far better than those who do it because it is a requirement. When people choose professions, they should do so because they have a love for the work. The amount of love put in helps maintain morality in one’s actions because what is being done is no longer a chore but a creation, and as we know, people love their own creations.

Commitment. The vision and love applied to the work bonds the individual to whatever he or she is doing until it is done: this is commitment as we defined it earlier.

3.2.3.3 Responsibility

Responsibility concerns roles, tasks, and actions and their ensuing consequences. For example, as parents we have an obligation and a duty to bring up our offspring.

That is parental responsibility. But responsibility also depends on a person’s value system, which is based on his or her environment and culture. There are various types of responsibilities, including personal, communal, parental, and professional, and these responsibilities vary depending on the age of the individual and his or her position in society. For example, the responsibilities of a 5-year-old are far different from those of a 40-year-old. Clearly the responsibilities of a country’s chief executive are different from those of a janitor. When individuals choose a lifestyle implied in a career or a vocation, they choose and must accept the package of responsibilities that go with that lifestyle.

Responsibilities of a Professional as a Provider A professional in either a provider-client or a provider–customer relationship represents the provider of either a service or a product. This relationship, as we pointed out earlier, is a contract between the two parties. The relationship consists of three major types of responsibilities: service, product, and consequential.

Service Responsibilities For a professional to provide a service to a client, there must be a contract binding the professional and the client. In this contract, as in any other contract, the professional has specific responsibilities regarding the time of delivery of the service, the quality of the service, and the consequences after the service has been rendered. For example, in the time-constraint responsibility, the service must be rendered within an agreed timeframe; if not, a new time must be negotiated. In the quality of service responsibility, the service must meet its desired goal so far as the client is concerned, and it must have the expected value. The consequence responsibility involves the safety of the client from harm, both physical and financial, after receiving the service. The provider must take all these responsibilities seriously.

Product Responsibilities If the contract between the provider and the client involves a product, the provider has the responsibility to deliver the product agreed upon on time, in good shape and of quality, and to provide documentation for the safe use of the product. The provider of the product is responsible for all liabilities that might arise as a result of use of the product. In liability cases, the provider responsibility depends on the contract and the degree of harm. We say more about liabilities in Chaps. 6 and 8.

Consequential Responsibilities In a television medical drama episode I watched, an operating room scene showed a female doctor dancing to a reggae tune while operating on a patient and unknowingly sewing the patient up with a surgical metal clip still in the patient's chest. In the next scene the patient has died and the autopsy shows the metal clip is still in his chest. Before the results of the autopsy, the doctor remembers her error and naturally becomes remorseful, not knowing whether to tell the family right away or wait until the medical report is issued. She knows full well that whatever the case, the family is going to sue her and the hospital, and probably her job at that hospital and her medical career are over. There is remorse on the part of the doctor and anger on the part of the patient's family, all because one person did not fulfill her responsibilities.

Remorse and anger are aftereffects of an action gone wrong, in this case a professional service. Whether a professional has provided a service or a product, there are always aftereffects of that action. Oftentimes one is praised for a service well done and the best product ever provided, but there are also times when one is remorseful because a service did not produce what it was intended to or a product did not live up to expectations. In the worst case scenario the service or product may cause physical or financial harm to the client. In such cases, one expects liabilities for the service or product, and the professional must accept those consequential responsibilities. In the case of the doctor, the service she provided fell

short of what was expected, and she had to face the consequential responsibilities of her actions, which at times not only include the parties involved but may also involve innocent bystanders.

3.2.3.4 Accountability

One way we can define accountability is the obligation to answer for the execution of one's assigned responsibilities. This process involves the “cycle of setting measurable goals, planning what needs to be done to meet those goals, reporting progress towards goals, evaluating the reports, and using that feedback to make improvements” [6]. Accountability involves these three key elements [7]:

1. *A set of outcome measures that reliably and objectively evaluate performance:*
In every profession there is a minimum set of measures that every individual in that profession must meet. This set must be carefully selected, and those measures must be attainable. However, these measures vary according to the profession and the individual activity to be performed by the professional. For example, in the teaching profession, one of the measures might be the success rate of students when they take standardized examinations.
2. *A set of performance standards defined in terms of these outcome measures:*
Similar to outcome measures, performance standards must be carefully chosen and attainable. These standards are also very dependent on the profession, but each profession must have a set of common performance standards for all its members for every type of service or product provided by that profession. For the teaching profession, the standard of output measures might be the passing of standardized examinations at a certain predetermined level. In the law profession, it might be the ability of a judgment to stand on subsequent appeals. Whatever standard measure is chosen, it must be plausible and measurable.
3. *A set of incentives for meeting the standards and/or penalties for failing to meet them:* The incentives chosen must be good enough so as not to create undesirable motives. For example, if the incentives are too good, they may force professionals to put the interest of their customers and clients below the interest of attaining the measures. If the incentives are monetary, they may force professionals to put the interest of making money ahead of the services they are supposed to offer. Similarly, the penalties prescribed must not be so harsh that they drive away those who intend to enter the profession. Harsh penalties also tend to make people who are in the wrong hide their actions and dig in deeper for fear of being discovered.

3.3 The Making of an Ethical Professional: Education and Licensing

In our discussion of the evolution of the professions, we have noticed the never-ending requirements of an individual seeking membership in the chosen profession or trade to either have a deep knowledge of the profession acquired through formal education or to be intrinsically of a “gentleman’s calling,” willing to hold a higher ethical standard. To continue to uphold these essential requirements in both professions and trades, let us now discuss three items that encourage, maintain, and improve that higher ethical standard: these are formal education, licensing, and professional codes of conduct. Professionals must follow a specific process to meet and maintain those professional requirements.

3.3.1 Formal Education

For formal education to be effective in teaching and enforcing the pillars of professionalism, it must be targeted and incremental. Let us walk through the making of an information technology professional as an example. In elementary school, as students are introduced to information technology, they should be told not to use machines to destroy other people’s property or to hurt others. But these cautions should be explained in an age-appropriate way. For example, children should be taught responsible ways of using computers and the Internet.

They should be told not to visit certain Web pages, to avoid getting involved in relationships online, to not give personal and family information online, and not to arrange for a rendezvous online or offline. In addition, they should be told to respect others’ work and property, whether they are online or off. Cases have already been reported of children as young as 14 years old breaking into computer systems and destroying records. In fact, many of the computer network attacks, and a good number of the headline-making incidents, have been perpetuated by young people, sometimes as young as 10 years of age. For example, in a certain county in Tennessee, several ninth-graders broke into their school computer system and infected it with a virus that wiped out most of the school records. It is believed the students got the virus off the Internet [8]. The content of what is taught must be relevant and sensitive to different age groups and professionals.

As students go through high school, content should become progressively more sophisticated. The message about responsible use of computers should be stressed more. The teen years are years of curiosity and discovery, and many young people find themselves spending long hours on computers and other online devices. Those long hours should be spent responsibly. Although a significant portion of the message should come from parents, schools should also be part of this partnership by offering courses in responsible use of computers. The teaching could focus on ethics: students should be given reasons why they cannot create and distribute viruses, download copyrighted materials off the Internet, and use the Internet to

send bad messages to others. These ethical reasons go beyond the “do it and you will be expelled from school” type of threats.

In college, of course, the message is more direct. There are several approaches that bring the message across to students:

- (i) Students take formal courses in professional ethics in a number of professional programs in their respective colleges.
- (ii) Without taking formal courses as part of their curriculum, students are taught substantial amounts of information ethics sprinkled throughout their courses in both general education and their major.
- (iii) A capstone course is used in the general education requirements, and information ethics content is added to that course. Many colleges now require computer literacy as a graduation requirement. Use that course to add ethics content.
- (iv) At exit, a 1-h information ethics course is required, which can be taken online.

Once they join the workplace environment, these professionals should be required to attend informal refresher sessions, seminars, and in-service workshops periodically.

3.3.2 Licensing Authorities

Licensing grants individuals formal or legal permission to practice their profession, which tips the balance of power in the giver-receiver equation in favor of the giver. Before a license is issued, certain formalities must be accomplished; for example, testing the competence of the aspirant for the specific knowledge and skills required. If such a test is not passed, the licensing authority may deny issuing the license. Beside testing for competence, the licensing authority also provides the licensee with a set of rules required to keep the license. If the rules are violated, the authority may have the prerogative of either sanctioning the licensee or recalling the license. Clearly, a license is a privilege, not a right, and if licensees want to maintain that right, they must follow the prescribed code. Licenses can be (and have been) used as both control and educating instruments to enforce rules, laws, and certain group or society norms.

Many professions license members, and most of these professions require the potential licensee to take and pass examinations that sometimes test both knowledge and skills. Many professions, to keep members updated and compliant with their codes, limit the validity of their licenses to specific time periods so members must renew their licenses. They then tie license renewal to passing of continuing examinations, which helps ensure that members stay knowledgeable in the profession. Professions also use periodic licensing examinations to check on member compliance with codes. If members have in the past violated a code and been reported, such members may not have their licenses renewed even if they pass all examinations.

It is quite evident that in those professions with no licensing requirements, discipline has been and continues to be a problem, whereas in those maintaining vigorous licensing requirements, relatively few disciplinary problems have emerged. Because every profession strives for a good image, many legitimate professions require licensing for membership. Licensing enables professions to enforce their rules by law. For example, physicians can lose their license to practice if they do anything unlawful or unethical. Once such a license is withdrawn, a physician cannot practice medicine. Although there are many professions with licensing requirements but with no enforcement mechanism, an increasing number of professions are opting to require enforceable licensing to keep their image untainted.

3.3.3 Professional Codes of Conduct

The primary purpose of professional codes of conduct is to promote the public image of the profession by specifying and enforcing the ethical behavior expected from its members. Accordingly, and in most cases, professional codes consist of standards, canons, and rules of conduct that address the following areas [9]:

- Moral and legal standards
- Professional-client relationship
- Client advocacy
- Professional-public relationships
- Sanction mechanics
- Confidentiality
- Assessment
- Compliance
- Competence
- Certified professional credentials for those professions that use certification.

For professional codes of conduct to be effective, a profession must institute a system of enforcement, reporting, hearing procedures, sanctions, and appeals. Codes without such systems in place are completely ineffective.

3.3.3.1 Enforcement

Experience and studies have shown that professions with enforceable codes have fewer discipline problems among their members than those with no codes or those with codes but without enforcement mechanisms. Those professions with fewer disciplinary problems naturally have better images. Because the purpose of codes for any profession is to create and maintain a good image, those professions without codes should come up not only with codes, canons, and guidelines, but also with enforcement mechanisms, and those with codes but with no enforcement system should add the enforcement. It is common knowledge that laws, codes, canons, or guidelines are not obeyed until and unless some type of enforcement machinery is in place. There are various techniques of enforcement, most of which have

no civil authority. The most widely used are professional ethics boards, standing committees, or review boards charged with the following actions:

- Drawing up the codes of ethics for the profession if none exist
- Revising codes if and when necessary
- Conducting education campaigns at the membership level
- Distributing copies of the codes to every member
- Developing disciplinary procedures
- Receiving complaints, conducting hearings, counseling members, and sanctioning members found guilty
- Promoting the image of the profession [9].

3.3.3.2 Reporting of Grievances

There are two main reporting procedures. The first is the typical organizational route in which a complaint is reported first to the local chapters if it exists. The complaint then makes its way to the top, usually to the national ethics committee. The second is the short-circuit procedure in which reporting can be done at any level and then from there a complaint is forwarded all the way to the top. Professions may vary these two types of reporting, mainly in the area of who is supposed to report a professional in violation. In some professions, the reporting must be done by a member of the profession in good standing and nobody else. In this case, concerned members of the public must report their complaint to a member of the profession, who then forwards the complaints to the committee. In other professions, any member of the public can initiate a complaint with the local professional board. Whichever way a complaint is reported, there should be a way to inform members of the profession and the public on the procedures of reporting and who can and cannot file a complaint, and there must be established channels of communication.

3.3.3.3 Hearing Procedures

Generalization of hearing proceedings is because of the many factors involved: for example, the nature, the financial standing, and the structure of the profession; the kind of enforcement procedures being used; and the penalty to be imposed. If there is no enforcement procedure or if the penalty is not significant, the accused member may not even appear for the scheduled hearing. Professions should consider all these factors when formulating the hearing procedures. For example, hearings should be held at the profession's nearest field office to save members from traveling long distances. If there is no field office, arrangements should be made to find a location convenient to both the accused and the hearing committee members, and the hearing process itself should be as short as possible.

3.3.3.4 Sanctions

If a hearing committee decides that a member is guilty of the offenses charged, then the committee must propose sanctions to fit the violations committed by the

member. The committee may decide to recommend any one or a combination of the following: probation, revocation of certification, request for resignation, and suspension from the profession at the member's expense. If a probation option is taken, the committee must clearly specify the nature, duration, and conditions of the probation. Also, there must be a person to whom the professional is to report for all requirements of the probation, including supervision. After the sanctioned member fulfills the requirements of the penalty, a recommendation can be made to reinstate the member in good standing.

3.3.3.5 Appeals

A profession must have an appeals process on the books for the sanctioned professional who is not satisfied with either the ruling of the committee or the penalty imposed. Such guidelines should state clearly the procedure for appeals, how the appeal instrument is to be handled, who attends to the appeals, and the maximum amount of time an individual has between the time he or she receives a judgment and the filing of the appeal. The time allotted for a judgment on the appeal should also be stipulated. The profession must also state whether an appealing member should continue executing his or her duties or be prohibited from doing so until the appeal is complete. In certain professions, appealing members are either put on administrative leave, suspended, or allowed to carry on with their duties pending the decision of the appeal.

Here is an example of a professional code of conduct for

The Institute of Electrical and Electronics Engineers, Inc.

CODE OF ETHICS 2

We, the members of the IEEE, in recognition of the importance of our technologies affecting the quality of life throughout the world, and in accepting a personal obligation to our profession, its members and the communities we serve, do hereby commit ourselves to the highest ethical and professional conduct and agree:

1. To accept responsibility in making engineering decisions consistent with the safety, health, and welfare of the public and to disclose promptly factors that might endanger the public or the environment
2. To avoid real or perceived conflicts of interest whenever possible and to disclose them to affected parties when they do exist
3. To be honest and realistic in stating claims or estimates based on available data
4. To reject bribery in all its forms
5. To improve the understanding of technology, its appropriate application, and potential consequences
6. To maintain and improve our technical competence and to undertake technological tasks for others only if qualified by training or experience, or after full disclosure of pertinent limitations
7. To seek, accept, and offer honest criticism of technical work, to acknowledge and correct errors, and to credit properly the contributions of others
8. To treat fairly all persons regardless of such factors as race, religion, gender, disability, age, or national origin
9. To avoid injuring others, their property, reputation, or employment by false or malicious action

10. To assist colleagues and coworkers in their professional development and to support them in following this code of ethics.

3.4 Professional Decision Making and Ethics

Now we come to our third theme on professionalism and ethics: the process of professional decision making. Here we focus on professional dilemmas and guilt associated with decision making.

3.4.1 Professional Dilemmas in Decision Making

Dilemmas in decision making are quite common in the everyday activities of a professional. The process of decision making resembles mapping with input parameters and an output decision. The input parameters in the decision-making process are premises. To each premise a value is attached. The mapping uses these values along with the premises to create an output, which is the decision. For example, if I have to make the decision whether to walk to church or take the car, the set of premises might include, among others, time, parking, safety, and gas. The values attached to these premises are that if I go by car, I save time, I need a parking space, walking is good exercise, I need money to buy gas. If I decide to walk, my decision might be based on a set of premises including health and money to which I may attach the following values: walking to church 1 day a month is good exercise and it saves me money for gas. The mapping function takes these premises together with the values and outputs a “logical” decision. This mapping function is similar to the one we used in the ethics definition in this chapter. Dilemmas in decision making are caused by questioning the values attached to one’s premises as inputs to the decision being made. One’s input values may be clouded by conflicting codes of conduct, advances in technology, or incomplete or misleading information.

3.4.1.1 Conflicting Codes of Conduct

In Fig. 3.1 of Sect. 3.2.2, we showed that every decision made by a professional must take into account the interrelationships of professional, personal, institutional, and local codes. The decision must be made in such a way that all four codes agree.

Decisions outside the core intersection must be weighted carefully because they always result in controversy. Take the case of the famous Michigan pathologist Dr. Kevorkian, the so-called Doctor Death. Dr. Kevorkian became a hero to some who believed in assisted suicide and Doctor Death to others who did not. He managed to force a debate over assisted suicide on the entire nation by repeatedly helping people to kill themselves using his “death machine,” for a total of at least 47 people. In the 7 years in which he accelerated his killing and before he was

eventually charged and put in prison, Dr. Kevorkian actually scoffed at the law, scorned elected and religious leaders, and won over juries. Dr. Jack Kevorkian became more known for his stand on assisted suicide than on his long years of professional service. The controversy was generated by the conflict in the codes, namely, the medical professional code of conduct, which includes the Hippocratic oath, and the local code, that is, the code of the town, county, and the state of Michigan (the institutional code does not apply because he was retired).

3.4.1.2 Advances in Technology

Dilemmas in decision making may also be caused by advances in technology. Computer technology in particular has created more muddles in the decision-making process than any other technology. Advances in computer technology create a multitude of possibilities that never existed before. Such possibilities present professionals with myriad temptations (see Sect. 3.5.1.2).

3.4.1.3 Incomplete or Misleading Information

Not having all the information one needs before making a decision can be problematic. Consider the famous prisoners' dilemma. Two people are caught committing a crime, and they are taken to different interrogation rooms before they have a chance to coordinate their stories. During the interrogation each prisoner is told that the other prisoner has agreed to plead guilty on all charges. Authorities inform the prisoner that agreeing to plead guilty on the charges as the other prisoner has done will bring him or her a reduced sentence. But rejecting the plea will of course mean that the accused is not cooperating with the investigation, which may result in him or her receiving the maximum punishment allowable. Each prisoner has four recourses:

- Plead guilty without the friend pleading guilty, which would mean deserting a friend
- Refuse to plead guilty and the friend pleads guilty, which would mean betrayal and probably a maximum sentence
- Plead guilty and the friend pleads guilty, which means light sentences for both of them
- Both refusing to plead guilty and probably both receiving a light sentence or a maximum sentence.

Whatever option the prisoners take is risky because they do not have enough information to enable them to make a wise decision. There are similar situations in professional life when a decision has to be made without enough information available and within time constraints. In such a situation the professional must take extra care to weigh all possibilities in the input set of premises and their corresponding values.

Taking all these into account and using the ethical framework we developed in the previous chapter can help the professional in making decisions that are just, fair, and plainly ethical.

3.4.2 Guilt and Making Ethical Decisions

In an ethical decision-making process, decisions are made based on, and reflect, consequences, individual liberties, and justice. To achieve this, individuals can use any other ethical theories to frame or make ethical choices that reflect the selected criteria. However, whatever theory used, the outcome falls into one of the following three criteria:

- *Utilitarian criterion*: in which decisions are made solely on the basis of their intended outcomes or consequences
- *Rights criterion*: in which decisions are made based on the set of liberties the society enforces, such as the Magna Carta and the Bill of Rights
- *Justice criterion*: which reflects justice; decisions are made so that they are fair, impartial, and equitable to all.

As we saw in Chap. 2 (Sect. 2.2.5), guilt is our natural internal judgment system, punishing ourselves based on our moral standards or the group's standards. Guilt therefore plays a crucial part in ethical decision making. In the decision-making process, guilt normally sets in right after the decision or a choice is made. And because guilt stays with the individual over a period of time, sometimes becoming cumulative, as we pointed out earlier, it may affect that individual's future decisions. Its effects on future decision-making processes center round new values being attached to the premises of the input set to the decision function. A guilty person reexamines his or her value set attached to all premises that come into play in the decision-making process. Sometimes guilt produces doubts about the present values attached to the premises without producing new and better values. Guilt causes decision makers to agonize over decisions. As we noted in Chap. 2, an excess of guilt could cause an individual to withdraw from society, which could be more dangerous because a withdrawn person may start to challenge the values attached to the premises as he or she tries to justify the guilt, resulting in bad decisions being made.

Although decisions are based on the outcome of an individual's deliberations, considering all input parameters and attaching values to these premises calls for a thorough examination of each premise by the individual. This process is aided by the individual reflecting on these basic steps:

- Examining the ethically relevant issues, principles, standards, and practices
- Determining the different parties (and their special interests) who will be affected by the decision
- Deciding on an alternative course of action if and when the outcome of the decision is not what is expected
- Considering the probable consequences (short and long term) of each alternative on each of the different parties involved
- Thinking of consulting with a trusted colleague if the situation is complex, risky, or there is undue personal involvement

Determining how personal values, biases, beliefs, or self-interests influenced the decision (either positively or negatively) and whether the consequences of the decision have been evaluated

Being prepared to (1) assume responsibility for the consequences of the action, including correction of negative consequences, if any; (2) reengage in the decision-making process if the ethical issue is not resolved; and (3) evaluate the system(s) within which the issue arose, to identify and remove the circumstances that might facilitate and reward unethical practices.

3.5 Professionalism and Ethical Responsibilities

This is the last of our four themes in professionalism and ethics. We focus here on professionalism and ethical responsibilities that include whistle-blowing, harassment, and discrimination.

3.5.1 Whistle-Blowing

The term whistle-blowing gives the impression of an act of seeking public attention. This behavior is what we see in a sports event whenever a foul is committed. The referee blows a whistle to call public attention, including that of the athlete, to the unsportsmanlike act committed. In some countries, law enforcement personnel use whistles to draw public attention to what they deem unlawful acts and to seek help.

The purpose of whistle-blowing in the workplace and the goal of a whistle-blower are the same as in the sports arena: calling to public attention, including and especially to that of a higher authority such as a government, what is considered an illegal or mismanaged act. Whistle-blowing can be internal, in which case the attention is sought internally and remains within organizational channels, or it can be public, in which case it alerts everyone.

Every day people, especially employees, witness wrongdoing on the job. What they witness usually can jeopardize not only their health, safety, or lives but the well-being of others. Quite often many witness such illegal acts but choose to remain silent in the face of such misconduct because they think it is not their responsibility or they believe it will not make a difference. Yet others fear to cause problems on the job. A few brave it out to save lives. However, quite often, instead of receiving praise for their brave actions and high integrity, they are often targeted for retaliatory acts such as investigations, ridicule, blacklisting (especially in their trade), harassment, intimidation, demotion, and sometimes outright dismissal.

So, in light of these threats, the most important aspect of whistle-blowing is to remain anonymous. Revealing the identity of a whistle-blower could be dangerous. Besides the obvious risks of potential job loss and poor or inadequate legal protection, there is also a psychological and sometime emotional price to

pay for whistle-blowing. Personal friends and family may turn against you. At work you may be labeled a troublemaker, leading people with whom you work to treat you as an outcast. Thus, care must be taken before whistle-blowing to ensure anonymity. The most difficult decision may involve finding a good medium that will ensure that confidentiality and anonymity. It is difficult and almost impossible to expect total anonymity for a whistle-blower, however, because the need for sufficient information to support allegations may result in revealing one's identifying details.

Different whistle-blowing methods have been used for years, ranging from traditional ones to more modern computer-aided means.

3.5.1.1 Computer-Aided Methods

Most common methods are anonymous, including anonymous remailers who use a software program to strip the header and all other identifying data from an original e-mail before forwarding it to its destination. Because the remailer does not include any return address on the e-mail, it attaches a pseudonymous address in case you need a reply. Before using anonymous remailers, however, exercise caution because the authorities can force the server administrator to reveal the owner of the pseudonymous name and address in cases of emergencies and other coercion.

3.5.1.2 Traditional Methods

A cross section of traditional methods is used in whistle-blowing. Historically, whistle-blowing has used spy-like methods to pass on information to either the public or a higher authority. All methods that ensure anonymity can be used: the most common include face-to-face communication with a public person who will ensure your anonymity; talking with the news media, which can keep your identity a secret; hotlines that alert the caller identity; and writing letters.

Whistle-blowing has been praised by many as courageous actions taken by a few good people with a moral conscience who risk everything to call public attention to illegitimate business practices and illegal and immoral actions. Others have condemned whistle-blowing as acts of vendetta, revenge, and greed that should not be encouraged. In fact, most whistle-blowers are either fired employees or unhappy ones. The following situations can complicate whistle-blowing.

Fear of reprisals: Many illegal and immoral acts go unreported because would-be whistle-blowers fear reprisals such as physical harm, job loss, reassignment to less desirable, sometimes demeaning jobs, suspension from work, and denial of promotions or training. In one survey, Keenan and Charles Kreuger report that this fear is real in potential whistle-blowers. Their survey indicated that not many organizations were willing to protect the whistle-blower. In fact, only one-half of the managers surveyed indicated a willingness to protect a whistle-blower [10, 11].

Suspicion surrounding whistle-blowing: Not every whistle-blower should be taken seriously because not all of them are sincere. Some whistle-blowers

are driven by greed, vendettas, anger, or revenge. In fact, many known cases of whistle-blowing were provoked when management and the employee disagreed. In other cases, whistle-blowing is caused by booty promises, especially by governments, to reward anybody with a certain percentage of the proceeds from whistle-blowing. In the United States, for example, private employees can sue any company on behalf of the government under the 1986 amendment to the False Claims Act (see Appendix B available to download from this book's Springer webpage) commonly known as a qui tam action, if that company has any dealings with the federal government. Under this law, a person who discovers fraud against the government can file a civil suit if the government does not take the case. As an incentive to whistle-blowing, any money recovered is shared between the government and the plaintiff, who can receive as much as 30% of the amount involved in the fraud.

Membership in organizational channels: Sometimes a whistle-blower act may be ignored because the whistle-blower is a member of the company or business organizational channel. Weil [12, 13] cites two whistle-blowers who are not considered as such because they called public attention to a serious ethical and moral problem but remained within the lines of command and, therefore, were not taken seriously. Both Roger Boisjoly and colleague Allan MacDonald of Morton Thiokol in Utah are known to have opposed the launch of the fated "Challenger" but were overwhelmed by management, and they then blew the whistle in the hearings of the Presidential Commission set by President Ronald Reagan.

Because whistle-blowing can save lives and reduce waste, the U.S. government encourages people who witness illegal acts and government waste to whistle blow. Besides enacting laws such as the False Claims Act that seek to expose fraud in federal contracts, the government also suggests that the would-be whistle-blower observe the following steps [14]:

1. Before taking any irreversible steps, talk to family and close friends about your decision to blow the whistle.
2. Be alert and discreetly attempt to learn of any other witnesses who are upset about the wrongdoing.
3. Before formally breaking ranks, consider whether there is any reasonable way to work within the system by going to the first level of authority. If this is not possible, think carefully about whether you want to "go public" with your concerns or remain an anonymous source. Each strategy has implications: the decision depends on the quantity and quality of your evidence, your ability to camouflage your knowledge of key facts, the risks you are willing to assume, and your willingness to endure intense public scrutiny.
4. Develop a plan, such as the strategically timed release of information to government agencies, so that your employer is reacting to you, instead of vice versa.
5. Maintain good relationships with administration and support staff.

6. Before and after you blow the whistle, keep a careful record of events as they unfold. Try to construct a straightforward, factual log of the relevant activities and events on the job, keeping in mind that your employer will have access to your diary if there is a lawsuit.
7. Identify and copy all necessary supporting records before drawing any suspicion.
8. Break the cycle of isolation, research, and identify and seek a support network of potential allies, such as elected officials, journalists, and activists. The solidarity of key constituencies can be more powerful than the bureaucracy you are challenging.
9. Invest the funds to obtain a legal opinion from a competent lawyer.
10. Always be on guard not to embellish your charges.
11. Engage in whistle-blowing initiatives without using employer resources.
12. Do not wear your cynicism on your sleeve when working with the authorities.

3.5.2 Harassment and Discrimination

Harassment is to verbally or physically create an environment that is hostile, intimidating, offensive, severe, pervasive, or abusive based on a number of parameters including one's race, religion, sex, sexual orientation, national origin, age, disability, political affiliation, marital status, citizenship, or physical appearance. Discrimination, on the other hand, is a process of making decisions that negatively affect an individual, such as denial of a service, based wholly, or partly, upon the real or perceived facts of one's race, religion, sex, sexual orientation, national origin, age, disability, political affiliation, marital status, or physical appearance. Harassment and discrimination are serious breaches of human rights. In fact, harassment is a form of discrimination. If not attended to, harassment not only affects just a few individuals but eventually grows to affect everyone in the organization. The following steps are needed in fight against harassment and discrimination:

- (i) *Awareness.* There are no clear signs of harassment, but in most cases harassment is manifested in the following signs: unhappiness, anxiety, discomfort, stress, and lifestyle changes. If some or all of these signs start to appear in the environment of an individual, then there is harassment. Discrimination is even harder to detect than harassment. However, there is discrimination if the decisions made are based upon the discriminatory factors previously listed.
- (ii) *Prevention.* The main tool for the prevention of harassment and discrimination is that an organization has a clear and simply written policy framework setting out the procedures that must be utilized if harassment and discrimination occur. The procedures must include awareness/education, the complaint process, sanctions, and redress.

3.5.3 Ethical and Moral Implications

The act of whistle-blowing is meant to alert and call the public to be witnesses to illegal acts that may be hazardous to their health and well-being or that may waste public resources. Of course, as we pointed out earlier, there are many other reasons for whistle-blowing. Are whistle-blowers living saints who fight evil to bring serious problems to light, thus contributing to the protection of the public's welfare? Does this explain the small numbers of whistle-blowers, although it is known that there are organizations in which a high potential for catastrophe can develop and somehow remain unexposed despite many people being aware of the problems?

Even people with high moral standards can be prevented from doing what is morally right because of the privileges, rights, and freedoms they stand to lose within the organization if they become known. People who feel accused and those allied to them tend to hit back to deflect attention from the accused. Retaliation is very damaging. So a would-be whistle-blower either decides to stay in line and live with a moral dilemma, but survive, or resign and live with a clear conscience. For a professional, a decision such as this presents a complex moral conundrum because if he or she stays within the organization, retaliation is almost predictable. Staying with the organization also presents other problems to both whistle-blower and colleagues. For example, collegial relationships and networks are disrupted. However, whistle-blowing is morally justifiable when the activities involved pose serious danger and harm to human life. The moral concept of whistle-blowing is good; it helps those who dare not speak out and all others who are affected.

Harassment and discrimination are both evil acts that not only challenge the conscience of an individual doing the acts but also create a situation that brings discomfort and inferiority to the targeted individual. It is, however, unfortunate that most individuals perpetuating the acts of discrimination and harassment lack the moral conviction and conscience.

Exercises

1. Define professionalism and list three elements of professionalism.
2. Why are the following concepts important for professionalism? Justify your answers.
 - Commitment
 - Integrity
 - Responsibility
 - Accountability
3. Discuss the merits of licensing software professionals.
4. Give an example of a decision that involves the examination of all four categories of codes.
5. Why is whistle-blowing so controversial? What are its pros and cons?
6. Why are harassment and discrimination so difficult to detect?

7. Is computer technology negatively or positively affecting harassment and discrimination?
 8. Discuss the effects of whistle-blowing on a whistle-blower.
 9. Study and discuss the False Claims Act.
 10. Has the False Claims Act been successful?
 11. Why is ethical decision making like software engineering?
 12. Are whistle-blowers saints or blackmailers?
 13. Why is it so difficult to make an ethical decision in today's technologically driven society?
 14. What role does guilt play in professional decision making? Why is it so important?
 15. Does every valid ethical argument involve a set of layers of arguments?
 16. Suggest a more fitting role for licensing authorities.
-

References

1. Pharmacist Robert Courtney admits he diluted drugs. The Kansas City Star, 10 July 2014
2. Webster's dictionary
3. R. Sizer, A brief history of professionalism and its relevance to IFIP, in *Ethics of Computing: Codes, Spaces for Discussion and Law*, ed. by J. Berleur, K. Brunnstein (Chapman and Hall, London, 1996)
4. J.M. Kizza, Professionalism, ethical responsibility, and accountability, in *Social and Ethical Effects of the Computer Revolution*, ed. by J.M. Kizza (McFarland, Jefferson, 1996)
5. W.S. Humphreys, *Managing for Innovation: Leading Technical People* (Prentice Hall, Englewood Cliffs, 1987)
6. Y.-F. Ng, Developing an accountability framework: political advisors in the Westminster system of governance. <http://www.ippapublicpolicy.org/file/paper/593a29bd6e1f6.pdf>
7. California Community Colleges (6870). https://lao.ca.gov/analysis_1995/chf6870.html
8. M. McMahan, Hearing set for school computer hackers. Chattanooga Free Press, 22 Mar 1997, sec. C1 (1997)
9. J.M. Kizza, The role of professional organizations in promoting computer ethics, in *Social and Ethical Effects of the Computer Revolution*, ed. by J.M. Kizza (McFarland, Jefferson, 1996)
10. T. Tomlinson, *Nursing Ethics* (Western Schools, 1993)
11. J. Keenan, C.A. Kreuger, *Internal Whistle-Blower: How Your Organization Can Profit by Listening* (Wingtips, 1995), pp. 28–30
12. The Canadian Code of Ethics for Psychologists, 4th edn. 2017 (1991)
13. V. Weil, Whistle-blowing: what have we learned since the challenger? Ethics Education Library. <http://ethics.iit.edu/eclibrary/biblio/whistleblowing-what-have-we-learned-challenger>
14. B.A. Senior, Profession: brief history and epistemology of a challenging word for construction (2006)

Further Reading

15. C. Johansson, L. Ohlsson, An attempt to teach professionalism in Engineering education. <http://www.hk-r.se/bib/rapport/2-93.html>
16. T. Devine, *The Whistleblower's Survival Guide: Courage Without Martyrdom*, First Printing edn. (Fund for Constitutional Government, 1997)



Anonymity, Security, Privacy, and Civil Liberties

4

Abstract

This chapter surveys the traditional ethical and privacy issues including security, anonymity and the analysis of how these issues are influenced by computer technology. This dialog also looks at privacy and the protection of civil rights. But in the absence of and agreed upon set of civil liberties by scholars, the discussion focuses on the following four accepted categories: (i) criminal justice that includes police powers, personal liberty, and the right to a fair trial; (ii) basic freedoms of speech, assembly, association, movement, and no discrimination; (iii) freedom of information; and (iv) communications and privacy. With the rapid advances in computer technology, and in particular the advent of the Internet mobile telecommunication technologies, the reader is challenged and brought into the discussion of finding ways, best practices and in some cases protocols and frameworks to protect these civil liberties. The chapter ends with a challenge to the reader to find a fitting ethical framework to protect us and our ethical and social values against the avalanche of these technologies. What should be included in it? Is there a need for a legal framework also? The reader is prompted!

Learning Objectives

After reading this chapter, the reader should be able to

1. Summarize the legal bases for the right to privacy and freedom of expression.
2. Analyze stated security procedures for “weak points” that an attacker could exploit and explain how they could (or will) fail.
3. Propose appropriate security measures for different situations.
4. Describe current computer-based threats to privacy.

5. Explain how the Internet may change the historical balance in protecting freedom of expression.
6. Describe trends in privacy protection as exemplified in technology.

Scenario 1: Did You Say Privacy? What Privacy?

Surveillance technology has progressed to the point that it is possible to identify individuals walking city streets from satellites in orbit. Telephone, fax, and e-mail communications can routinely be monitored. Personal information files are maintained on citizens from cradle to grave. There is nowhere to run ... nowhere to hide. Personal privacy is dead.

24/7 OF THE (UN)KNOWN CITIZEN

W. H. Auden (1907–1973) notes in his poem “The Unknown Citizen” that his unknown citizen was a “modern man” for “he was fully insured, and his Health-card shows he was once in hospital but he left cured. Both Producers Research and High-Grade Living declare he was fully sensible to the advantages of the installment plan, and he had everything necessary to the Modern Man, A phonograph, a radio, a car, and a Frigidaire” [1]. Our citizen is definitely a modern man, more so than the archaic Auden’s, for he owns all modern life’s amenities, plus a company car, a cellular phone, and a reserved parking spot. He is computer savvy, with a computer in his office and at home. The normal day of our citizen starts at 6 a.m. when he is awakened by soft music from a radio. The radio reports that futures are up, indicating higher stock price opening. He jumps out of bed and switches on his computer to put in a “buy” order for a stock he has been meaning to buy.

Snapshot #1: Tempest

Several yards outside his private home, someone with a Tempest (a criminal, government agent, private investigator) is recording all that our citizen is doing. The information our man has used on the computer has been recorded by the Tempest.

HEADLINER #1: “PRIVACY LOST: THE DEMISE OF PRIVACY IN AMERICA” [2]

As our citizen pulls out of the garage, he notices that he is low on gas, so he pulls up at the nearest Conoco gas station. Being a modern man as he is, he decides to pay at the pump.

Snapshot #2: The transaction record is entered into a Conoco database and, of course, Visa database.

Without worry he pulls away, speeding to work.

HEADLINER #2: “SPYING ON EMPLOYEES: BIG BROTHER HAS HIS EYE ON YOU ... SNOOPING SISTER IS SUPERVISING YOU” [3]

Snapshot #3: “Wherenet Tracking Knows Where You Are at Every Moment”

Just before 8 a.m., he arrives at his workplace, pulling into his reserved parking spot. He enters the lot with an electronic key.

Snapshot #4: *Company network and surveillance cameras start their full day of recording his activities. At 8:02 a.m. he settles into his office. He starts the day’s activities with several e-mails to read and a few calls to return.*

HEADLINER #3: “SNOOPING BILL TO BECOME LAW BY NOVEMBER” [4]

Snapshot #5: *Echelon*

At 12:01 p.m. he receives a call from the company representative in Greece to discuss the new company marketing strategy he proposed last month.

At 3:15 p.m. he heads to his doctor’s appointment. He has been complaining of knee pains. His doctor orders for a few tests. They are done quickly, the labs taking as much detail from him as possible and his doctor updating his medical record. His insurance will pay for part of the visit. So the insurance is billed and his other medical record is updated. He pays by check. His other financial record is updated.

As he leaves his doctor’s office at 4:30 p.m., he decides to call it a day and head home. He calls his office to inform his secretary that he is heading home!

Snapshot #6. “*Big Brother in the Flesh: New Technology Could Make Us All a Part of the Collective, Permanently Supervised from Above*” [5]. *On the way home, he remembers that he needs a few groceries. So he heads to Kroger’s. At the grocery store, he picks up a few things and he is picked up.*

Snapshot #7: *Kroger’s Surveillance Cameras. At the checkout counter, he gives in a dollar coupon for the chicken soup and he also hands in a Kroger’s card to save 75 cents off chili on the week’s special for Kroger’s most valuable customers.*

Snapshot #8: *Kroger’s database records the day’s transaction. To receive the card, the citizen provided Kroger with his home address, income, family size, and age.*

Snapshot #9: *Celeria*

At 5 p.m. he leaves Kroger’s and heads home. But on the way home, he receives a call from his girlfriend on his private cellular phone inviting him for dinner at her place.

At 5:30 p.m. he turns into his private driveway only to notice spilt garbage. He wonders whether the city garbage collectors did it. He puts the car in the garage and comes back to clean the driveway. The neighbor informs him that he noticed two guys going through his garbage, and they later drove away.

HEADLINER #4: “FORGET THE FIREWALL: GUARD YOUR GARBAGE AGAINST ‘DUMPSTER-DIVING’ HACKERS” [6]

After cleaning the driveway, he checks his snail mail. He notices that they are all bills!

HEADLINER #5: “WHO IS READING YOUR BILLS” [7]

At 6 p.m. before he leaves for his girlfriend’s house, the citizen decides to check his e-mail and complete some correspondence.

Snapshot #10: Carnivore

At 7 p.m. he leaves for his girlfriend’s house. He might spend the night there! The girlfriend is also modern and lives in a “Digital Home.”

HEADLINER #6: “LATEST SURVEILLANCE LEAVES NOTHING TO CHANCE: EXPLORING THE DARK SIDE OF THE DIGITAL HOME” [8]

Next morning, he will drive the company car to work! What else do we need to know? Is he happy? That is absurd. As Auden would put it: Had he been unhappy, we would have known!

Discussion Questions

1. Where do we go from here?
Legislation
Regulation
Self-help
2. Do they work?
3. Do you believe we still have individual privacy?
4. What do you think is the best way to safeguard privacy?
5. How much interference by government in your life can you tolerate to feel secure?
6. How much privacy are you willing to give up to feel secure?

Scenario References

1. W.H. Auden, The unknown citizen, in *Literature: An Introduction to Fiction, Poetry, and Drama*, 6th edn, ed. by X.J. Kennedy, D. Gioia (Harper Collins, New York, 1995), pp. 611–612

2. Privacy lost: the demise of privacy in America, <http://dorothyseeseonline.tripod.com/newsline/id3.html>
3. Moyers and Company, Big brother's prying eyes, 14 June 2013. <https://billmoyers.com/episode/big-brothers-prying-eyes/>
4. Snooper's Charter passes into law—what it means. <https://www.ft.com/content/40d2ede4-adac-11e6-9cb3-bb8207902122>
5. K. Mieszkowski, Big brother in the flesh: new technology could make us all a part of the collective, permanently supervised from above. Las Vegas Weekly, 21 Sept 2000 edn
6. S. McClure, J. Scambray, Forget the firewall: guard your garbage against 'Dumpster Diving' hackers. LISTSERV@SecurityFocus.com. Friday, 7 July 2000
7. D. Eisenberg, Who is reading your bills? A court ruling on privacy riles the FCC. <http://www.cnn.com/ALLPOLITICS/time/1999/08/30/privacy.html>
8. Possessed: the dangers of the digital home. <https://www.ft.com/content/bd8e187e-7799-11e7-a3e8-60495fe6ca71>

4.1 Introduction

Social, economic, and technological advances have dramatically increased the amount of information any individual possesses. Increasing demand for information and easier access to it have also created challenges. We have come to learn that information is a treasure in itself: the more you have, the better. Having valuable intellectual, economic, and social information creates enormous opportunities and advantages for an individual because information has become a vital resource in this information age.

Even though information is a treasure, it can also be a liability; for example, we are constantly seeking ways to acquire, keep, and dispose of it. We want to make sure that what is seen and heard privately does not become public without our consent.

In our technologically advanced society, a number of factors have contributed to the high demand for information and the subsequent need for anonymity, security, privacy, and the safeguard of our civil liberties. Among the main contributing factors are the following:

- High digitalization of information and increasing bandwidth
- Declining costs of digital communication
- Increased miniaturization of mobile computing devices and other communications equipment
- Greater public awareness by the news media of the potential abuse of digital communication, especially the Internet.

4.2 Anonymity

The Greeks used the word *ανώνυμία* to describe the state of being nameless. Anonymity is being nameless, having no identity. Because it is extremely difficult for anybody to live a meaningful life while being totally anonymous, people usually use some type of anonymity. Consider these several types:

Pseudo-identity: An individual is identified by a certain pseudonym, code, or number (compare with a writer's pen name): this is referred to as pseudonymity. It is used frequently in the "Witness Protection" program. This is the most common variant of anonymity.

Untraceable identity: One is not known by any name including pseudo-names.

Anonymity with a pseudo-address to receive and send correspondence with others: This technique is popular with people using anonymous remailers, user groups, and news groups.

4.2.1 Anonymity and the Internet

The nature of the Internet, with its lack of political, cultural, religious, and judicial boundaries, has created a fertile ground for all faceless people to come out in the open. In particular, the Internet provides two channels through which anonymous acts can be carried out.

1. *Anonymous servers*: With advances in software and hardware, anonymity on the Internet has grown through anonymous servers. There are two types of anonymity servers:
 - (a) Full anonymity servers, where no identifying information is forwarded in packet headers.
 - (b) Pseudonymous servers, which put pseudonym in forwarded packet headers, keeping the real identity behind a pseudonym, but being able to receive and forward all packets sent to the pseudonym to the real server.

Anonymity servers are able to accomplish this through the use of encryption. We are not going to discuss further how this encryption is done.

2. *Anonymous users*: It is increasingly impossible to ensure anonymity online. Some ways to be anonymous is to use pseudonyms. However, even this, anonymity can no longer be assured.

4.2.2 Advantages and Disadvantages of Anonymity

There are several advantages and disadvantages to anonymity. We consider some of these here, starting with advantages.

- Anonymity is good when a whistle-blower uses it to check unhealthy activities within the organization. Although whistle-blowers are controversial, they are good in a number of cases, especially when there is abuse of the office and resources. We discussed whistle-blowing in Chap. 3.
- Anonymity is good in case of national security so that underground spies can gather information that is good for national defense.
- Where there is intimidation and fear of reprisals, anonymity is good because useful information may be revealed.
- Anonymity is good for some relationships and the security of some people.

There are also disadvantages to anonymity.

- Criminals and embezzlers can use it to their advantage, especially in online social networks.
- Many disputes could be solved if information from individuals who are party to these disputes can reveal the necessary information.

4.2.3 Legal View of Anonymity

As we have pointed out in the last section, anonymity has its good and bad sides. More important, society may not be safe if many criminals use anonymity to hide their criminal activities. Anonymity can also bring suffering in social relationships in society. So, in a number of cases, it is necessary for either a local authority or national legislatures to pass laws that regulate when and who can use anonymity legally. In the current environment of the Internet, there are serious debates on the freedoms of individuals on the Internet and how these freedoms can be protected in the onslaught of people under anonymity in cyberspace.

Discussion Issues

1. List and discuss roles in society that require one to be anonymous and if this is beneficial to society.
2. Discuss the major disadvantages of anonymity, especially in cyberspace.

4.3 Security

In general, security can be considered a means to prevent unauthorized access, use, alteration, and theft or physical damage to property. Security involves these three elements:

1. *Confidentiality*: To prevent unauthorized disclosure of information to third parties. This is important in a number of areas including the disclosure of personal information such as medical, financial, academic, and criminal records.
2. *Integrity*: To prevent unauthorized modification of files and maintain the status quo. It includes system, information, and personnel integrity. The alteration of information may be caused by a desire for personal gain or a need for revenge.
3. *Availability*: To prevent unauthorized withholding of information from those who need it when they need it. We discuss two types of security: physical security, which involves the prevention of access to physical facilitates such as computer systems, and information security, which involves prevention of access to information by encryption, authentication, and other means.

4.3.1 Physical Security

A facility is physically secure if it is surrounded by a barrier such as a fence, has secure areas both inside and outside the facility, and can resist penetration by intruders. Physical security can be guaranteed if the following four mechanisms are in place: deterrence, prevention, detection, and response [1].

1. *Deterrence* is used to defend systems against intruders who may try to gain access. It works by creating an atmosphere intended to scare intruders.
2. *Prevention* is used in mechanisms that work by trying to stop intruders from gaining access.
3. *Detection* should be the third line of defense. This mechanism assumes the intruder has succeeded or is in the process of gaining access to the system, so it tries to “see” that intruder who has gained or who is trying to gain access.
4. *Response* is an aftereffect mechanism that tries to respond to the failure of the first three mechanisms. It works by trying to stop or prevent damage or access to a facility.

4.3.2 Physical Access Controls

To ensure physical security, a regimen of access controls must be put in place. In physical access control, we create both physical barriers and electronic protocols that will authenticate the user of the resource whose security we are safeguarding.

4.3.2.1 Physical Security Barriers

The physical barrier can be a fence made of barbed wire, brick walls, natural trees, mounted noise or vibration sensors, security lighting, closed-circuit television (CCTV), buried seismic sensors, or different photoelectric and microwave systems [1]. The area surrounding the facility can be secured using locks and keys, window breakage detectors, infrared and ultrasonic detectors, interior microwave systems, animal such as dogs, and human barriers such as security guards.

4.3.2.2 Electronic Access Controls

With advances in technology, we are moving away, although not totally, from the physical barriers to more invasive electronic controls that include card access control systems, firewalls, and the third, and probably the most important area, the inside, may be secured using electronic barriers such as firewalls and passwords.

Passwords

A password is a string of usually six or more characters to verify a user to an information system facility, usually digital systems. Password security greatly depends on the password owner observing all these four “never” cardinal rules:

1. Never publicize a password.
2. Never write a password down anywhere.
3. Never choose a password that is easy to guess.
4. Never keep the same password for an extended period of time.

Password security is not only important to individuals whose files are stored on a system, it is also vital to the system as a whole, because once an intruder gains access to one password, he or she has gained access to the whole system, making all its files vulnerable. Thus, system security is the responsibility of every individual user of the system.

Firewalls

A firewall is hardware or software used to isolate the sensitive portions of an information system facility from the outside world and limit the potential damage that can be done by a malicious intruder. Although there is no standardization in the structure of firewalls, the choice of firewalls depends on the system manager's anticipated threat to the system. Most firewalls are variations of the following three models:

Packet filters: Packet-level filters contain gates that allow packets to pass through if they satisfy a minimum set of conditions, and choke or prevent those packets that do not meet the entry conditions. The minimum conditions may include packets to have permissible origin or destination addresses, as determined by the network administrator. The filter firewalls can also configure and block packets with specific TCP or UDP packet port numbers, or filter based on IP protocol types. As we see later, packet filters have a weakness in that

they cannot stop or filter a packet with malicious intent if the packet contains the permissible attributes.

Proxy servers: These servers work on the protected portions of the network that usually provide information to outside users requesting access to those portions. That is, the firewall protects client computers from direct access to the Internet. Clients direct their requests for an Internet connection through the proxy server. If individual client requests conform to the preset conditions, then the firewall will act on the request; otherwise, it is dropped. These firewalls require specialized client and server configurations depending on the application.

Stateful inspection: These firewalls combine both the filter and proxy functions. Because of this, it is considered complex and more advanced. The conditions for a stateful inspection are, as the filter, based on a set of rules. In contrast to filters, these rules are not based on TCP or UDP but on applications as are proxy servers. They filter packets by comparing their data with archived friendly packets.

4.3.3 Information Security Controls

Information security includes the integrity, confidentiality, and availability of information at the servers, including information in files and databases and in transition between servers, and between clients and servers. The security of information can be ensured in a number of ways. The most common are cryptography for information transmission and authentication and audit trails at the information source and information destination servers. Cryptography, the science of writing and reading coded messages, forms the basis for all secure transmission, through three functions: symmetrical and asymmetrical encryption, and hash functions.

4.3.3.1 Encryption

Encryption is a method that protects the communications channel from sniffers, programs written for and installed on the communication channels to eavesdrop on network traffic, examining all traffic on selected network segments. Sniffers are easy to write and install and difficult to detect. Cryptography uses an encryption algorithm and key to transform data at the source, called plaintext, turn it into an encrypted form called ciphertext, usually an unintelligible form, and finally recover it at the sink. The encryption algorithm can either be symmetrical or asymmetrical.

Symmetrical encryption, or secret key encryption as it is usually called, uses a common key and the same cryptographic algorithm to scramble and unscramble the message as shown in Fig. 4.1. The security of the transmitted data depends on the fact that eavesdroppers with no knowledge of the key are unable to read the message. One problem with symmetrical encryption is the security of the keys, which must be passed from the sender to the receiver.

Asymmetrical encryption, commonly known as public key encryption, uses two different keys, a public key known by all and a private key known by only the sender and the receiver. The sender and the receiver each has a pair of these keys, one public and one private. To encrypt a message from sender A to receiver B

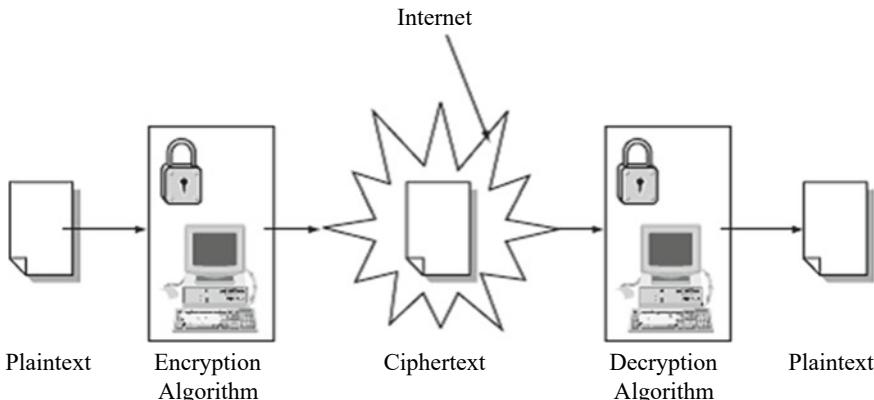


Fig. 4.1 Symmetrical encryption

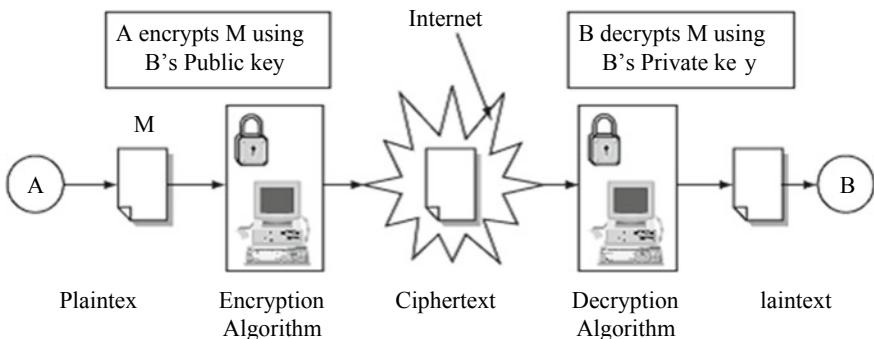


Fig. 4.2 Asymmetric encryption

(Fig. 4.2), both A and B must create their own pairs of keys. Then A and B exchange their public keys: anybody can acquire them. When A is to send a message M to B, A uses B's public key to encrypt M. On receipt of M, B then uses his or her private key to decrypt the message M.

A *hash function* takes an input message M and creates a code from it. The code commonly referred to as a *hash* or a *message digest*, is discussed more in the next section. A one-way hash function is used to create a digital signature of the message, just like a human fingerprint. The hash function is therefore used to provide the message's integrity and authenticity.

4.3.3.2 Authentication

Usually it is difficult for a system to verify the identity of a user, especially a remote user. Thus, authentication is a process whereby the system gathers and builds up information about the user to assure that the user is genuine. In data

communication, authentication is also used to ensure the digital message recipient of the identity of the sender and the integrity of the message. In computer systems, authentication protocols based on cryptography use either secret-key or public-key schemes to create an encrypted message digest that is appended to a document as a digital signature.

The digital signature is similar to a handwritten signature in printed documents. Similar to handwritten signatures, digital signatures ensure that the person whose signature the system is authenticating is indeed the true person, but digital signatures provide a greater degree of security than handwritten signatures. Also, digital signatures once submitted can never be disowned by the signer of a document claiming the signature was forged: this is called non-repudiation. A secure digital signature system consists of two parts: (1) a method of signing a document, and (2) authentication that the signature was actually generated by whomever it represents.

The process of signing the document, that is, creating a digital signature, involves a sender A passing the original message M into a hash function H to produce a message digest. Then, A encrypts M together with the message digest using either symmetrical or asymmetrical encryption, and then sends the combination to B. Upon receipt of the package, B separates the digital signature from the encrypted message. The message M is put into a one-way hash to produce a message digest, and B compares the output of the hash function with the message digest A sent. If they match, then the integrity of the message M together with the signature of the sender are both valid (see Fig. 4.3).

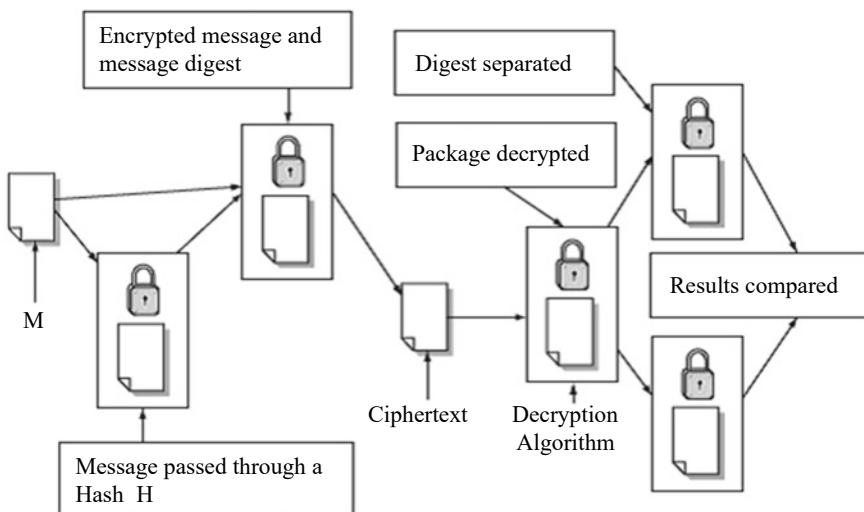


Fig. 4.3 Digital signature and authentication

Physical Authentication Methods. Authentication of users or user surrogates is usually based on checking one or more of the following user item:

- *User name* (sometimes *screen name*).
- *Password*.
- *Retinal images*: The user looks into an electronic device that maps his or her retinal image: the system then compares this map with a similar map stored on the system.
- *Fingerprints*: The user presses on or sometimes inserts a particular finger into a device that makes a copy of the user fingerprint and then compares it with a similar image on the system user file.
- *Physical location*: The physical location of the system initiating an entry request is checked to ensure that a request is actually originating from a known and authorized client machine. To check the authenticity of such a client, the network or Internet Protocol (IP) address of the client machine is compared with the one on the system user file. This method is used mostly in addition to other security measures because it alone cannot guarantee security: if used alone, it provides access to the requested system to anybody who has access to the client machine.
- *Identity cards*: Increasingly, cards are being used as authenticating documents. Whoever is the carrier of the card gains access to the requested system. As is the case with physical location authentication, card authentication is usually used as a second-level authentication tool because whoever has access to the card automatically can gain access to the requested system.

4.3.4 Operational Security

Operation security involves policies and guidelines that organizations, including all employees, must use to safeguard the assets of the organization, including its workers. These policy guidelines are spelt out in a document we call a security policy. It also includes guidelines for security recovery and response in case of a security incident.

4.4 Privacy

4.4.1 Definition

According to Jerry Durlak [2], privacy is a human value consisting of four elements he calls rights. We put these rights into two categories. The first category includes three rights that an individual can use to fence off personal information seekers; the

second category contains those rights an individual can use to control the amount and value of personal information given out.

1. Control of external influences:

- *Solitude*: The right to be alone without disturbances
- *Anonymity*: The right to have no public personal identity
- *Intimacy*: The right not to be monitored.

2. Control of personal information:

- *Reserve*: The right to control one's personal information including the methods of dissemination of that information.

The notion of privacy is difficult to accurately define because the definition of privacy depends on matters such as culture, geographic location, political systems, and religious beliefs.

4.4.2 Types of Privacy

Although there are varied definitions of privacy, the several types of privacy we discuss here are not influenced by the factors we have outlined in the previous section.

4.4.2.1 Personal Privacy

This type of privacy involves the privacy of personal attributes. The right to privacy of all personal attributes would mean the prevention of anyone or anything that would intrude or violate that personal space where those attributes are: this would include all types of intrusions including physical searches, video recording, and surveillance of any type. In a number of countries, there are statutes and acts similar to the U.S. Fourth Amendment, which guarantees the right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures.

Discussion Issue

Discuss a few of these statutes and acts.

4.4.2.2 Informational Privacy

In the previous section, we discussed the privacy of an individual meaning that we want to ensure the privacy of an individual by preventing any intrusions through physical and electronic access to that individual's attributes. Informational privacy, in contrast to personal privacy, concerns the protection of unauthorized access to

information itself. Of course, there are different strands of information that we have to protect, including the following:

- *Personal information:* Most personal information of value includes information on personal lifestyles such as religion, sexual orientation, political affiliations, or personal activities.
- *Financial information:* Financial information is important not only to individuals but also to organizations. Financial information is a very valued asset because it gives the organization the autonomy it needs to compete in the marketplace.
- *Medical information:* Medical information is very personal and very important to all of us. For personal, employment, and insurance purposes, many people want their medical information to be private.
- *Internet:* In this new age, the Internet keeps track of all our activities online. With an increasing number of people spending an increasing number of time online in social networks and the digital convergence becoming a reality with every passing day, not only will our social life be online but soon all our lives will be also. We want those activities and habits private.

4.4.2.3 Institutional Privacy

Institutions and organizations want their data private not only for business advantages but also for the life of the business. The research data, the sales and product data, the marketing strategies, and the activities of the organization all need to be private.

4.4.3 Value of Privacy

Privacy has traditionally been perceived as valuable and has even gained more importance in the information age because it guards an individual's personal identity, preserves individual autonomy, and makes social relationships possible.

However, these days in the information age, the value of privacy has been eroded. We can no longer guarantee our privacy. It has left many wondering whether there is such a thing as privacy any more. As the scenario at the start of this chapter demonstrates, no one has guaranteed privacy any more unless such an individual is no longer part of the society. From the telephone calls you make that identify you through caller ID, to every transaction you pay for either by credit card or by check, and to the multitude of forms you fill out from getting your pet groomed to having a prescription filled, you are identifiable and you have nowhere to hide. The most abused number, the Social Security number, is used as a personal ID by many companies, including health insurance companies, that use it as a customer ID in spite of the repeated warning from the federal government not to do so. In its effort to help stop the erosion of individual privacy, the U.S. Congress passed the Gramm–Leach–Bliley Financial Services Modernization Act of 1999,

but put in reverse conditions, the so-called opt-out discussed in Sect. 4.4.3.1, that make the Act useless.

We consider three attributes of privacy: personal identity, autonomy, and social relationships.

4.4.3.1 Personal Identity

As information becomes more precious, it becomes more important for individuals to safeguard personal identity. Personal identity is valuable because it enshrines personal privacy. Unfortunately, with rapid advances in technology, especially computer technology, it has become increasingly difficult to protect personal identity.

4.4.3.2 Autonomy

Humans need to feel that they are in control of their own destiny. They need their autonomy. The less personal information other people have about an individual, the more autonomous that individual can be, especially in decision making. However, other people will challenge one's autonomy depending on the quantity, quality, and value of information they have about that individual. People usually tend to establish relationships and associations with individuals and groups that will respect their personal autonomy, especially in decision making.

4.4.3.3 Social Relationships

In some societies where marriages are arranged, parents on both sides try to collect as much information about each other as possible before they commit their offspring in marriage. In societies where there are no arranged marriages, the parties involved usually spend a period of time dating. The dating time is spent collecting as much information as possible about each other. The couple then uses this information to make a decision about marrying. However, each party may try to conceal some information because some seemingly valuable information may not be worthwhile and may even lead to the breakup of the relationship.

4.4.4 Privacy Implications of the Database System

4.4.4.1 Information Gathering

Have you paid enough attention to the number of junk mails, telephone calls during dinner, and junk e-mails you have been getting? If so, you may have thought about who has your name on a list and what they are doing with it. In recent years, telemarketers have been having a field day as technological advances have replaced the door-to-door salesman. Many companies you have done business with may have sold or shared your personal information to other companies, and tracing the source may be difficult. In many cases, we do not preserve our privacy as we defined privacy earlier. We have helped information seekers, such as companies, in gathering and databasing information from us. We do this every time we apply for discount cards from grocery stores, gas stations, and merchandise stores, and

every time we fill out information on little cards to enter contests, and every time we give out our Social Security number and telephone numbers to store clerks in department stores. The information they collect from us is put into databases and is later sold to the highest bidder, usually a marketer.

Information gathering is a very serious business that is increasingly involving a growing number of players which traditionally were governments gathering mostly defensive information on weapon systems. However, with globalization and the Internet, the doors to the information-gathering field have been thrown open. Now individuals, companies and organizations, and of course governments, are all competing, sometimes for the same information.

The tools of the trade have also improved tremendously, becoming more stealthy, much smaller, and of course more accurate. With the modern tools of gathering information, no one is safe anymore. Because of our habits online, Internet crawlers are in action visiting our machines stealthily and gathering a wealth of information. There is no longer the need to get your information from the cards you fill out at shopping malls and grocery stores. There are better and faster ways now. There are tremendous legal and privacy issues that we have to confront. First, most of the information collected from us, that we come to know about, which is a fraction of what they take, is collected without our consent.

Although the problem is skyrocketing, there is minimum effort to curtail the practice. This is a result of a number of reasons, the most important of which is that the rate at which technology is developing is continuously outstripping our legal systems and our ability to legislate, let alone enforce, the new laws. Several attempts have been made, including the Gramm–Leach–Bliley Financial Services Modernization Act, aimed at restricting financial institutions such as banks and brokerages from sharing customers' personal information with third parties.

Although the Financial Services Modernization Act has given financial institutions an information bonanza, the Act also tries in some ways to protect the customer through three requirements that the institutions must disclose to us:

- (i) *Privacy Policy*: Through this, the institution is bound to tell us the types of information the institution collects and has about us and how it uses that information
- (ii) *Right to Opt-Out*: Through this, the institution is bound to explain our recourse to prevent the transfer of our data to third-party beneficiaries
- (iii) *Safeguards*: Through such, the institution must put in place policies to prevent fraudulent access to confidential financial information.

However, this same law, as many of its kind, has allowed these same U.S. financial institutions to merge and form what have been called financial supermarkets. This one Act has opened a door for these companies to merge and consolidate customer data from several sources.

4.4.5 Privacy Violations and Legal Implications

Privacy, as we have defined it, is a basic human value that is at the core of human dignity and autonomy. Because of this recognition, many major and historical documents such as the Fourth Amendment to the U.S. Constitution, the UN Universal Declaration of Human Rights, the Council of Europe, and many national and multi-national treaties contain enshrined clauses of the individual's right to privacy. It is believed that privacy forms the foundation of a free and democratic society.

However, this fundamental right is violated every day in many ways. Although individual privacy rights have been violated for years, the advent of the Internet has accelerated the rate and scale of violations. There are numerous contributing factors or causes of violations. Let us look at some here:

- (i) Consumers willingly give up information about themselves when they register at websites, at shopping malls to win prizes, and in mailing solicitations.
- (ii) Consumers lack the knowledge of how what they consider a little bit of information can turn into a big invasion of privacy.
- (iii) Inadequate privacy policies.
- (iv) Failure of companies and institutions to follow their own privacy policies.
- (v) Internet temptation, as discussed in Sect. 3.5.1.2, that enables businesses to reach individuals in a very short time in the 'privacy' of their homes and offices.

Because of the Internet's ability to reach many people with ease, major privacy violators have been online companies such as DoubleClick, an Internet advertising company. DoubleClick used its dominant position on the Internet to reach as many people as possible. With this opportunity, DoubleClick quietly abandoned its original policy of providing only anonymous data collected from its Internet users to marketers and started combining its online user profiles with information from direct mailers and others. The combined information could easily be used to identify the Internet users as DoubleClick's tracking could now reveal names, addresses, and purchasing habits.

After massive user protests, the threat of lawsuits, and notices from the U.S. Federal Trade Commission (FTC), DoubleClick gave up on their intended plan, putting the proposal on hold and claiming that they would seek user permission before launching the matching data plan [3]. Yahoo!, Inc, another Internet company, was sued by Universal Image, Inc., a company that makes educational videos. Universal sued Yahoo! for not living up to a contract to provide data, including customer e-mail information that Universal had signed with Broadcast.com, Inc., before Yahoo acquired it [4].

Because of the anticipated growth of the Internet, there is widespread agreement that privacy rights are under serious attack and that something has to be done. The measures that are needed to protect both the user and consumer need to be varied to include legislation like the U.S. Consumer Protection Act, enforcement, and self-help.

Other privacy violations include intrusion, misuse of information, interception of information, and information matching.

4.4.5.1 Intrusion

Intrusion is an invasion of privacy by wrongful entry, seizing, or acquiring possession of the property of others. For example, hackers are intruders because they wrongfully break into computer systems whether they cause damage or not. With computer network globalization, intrusion is only second to viruses among computer crimes, and it is growing fast.

4.4.5.2 Misuse of Information

Human beings continually give out information in exchange for services. Businesses and governments collect this information from us honestly to provide services effectively. The information collected, as discussed in Sect. 5.4.3, is not just collected only to be stored. This information is digital gold to these companies. They mine the gold from us and sell it to the highest bidder. There is nothing wrong with collecting personal information when it is going to be used for a legitimate reason, for the purpose for which it was intended. However, the problem arises when this information is used for unauthorized purposes; collecting this information then becomes an invasion of privacy.

4.4.5.3 Interception of Information

Interception of information is unauthorized access to private information via eavesdropping, which occurs when a third party gains unauthorized access to a private communication between two or more parties. Information can be gathered by eavesdropping in the following areas:

- At the source and sink of information, where either client or server intrusion software can listen in, collect information, and send it back to the sender
- Between communication channels by tapping into the communication channels and then listening in.

4.4.5.4 Information Matching

The threat of information matching can best be highlighted by an old story recounted by Mason [5], who says it has been retold so many times that its accuracy is probably in doubt; however, its message remains the same.

Here is the story:

A couple of programmers at the City of Chicago's computer center began matching tape files from many of the city's different data processing applications on name and I.D. They discovered, for example, that several high-paid city employers had unpaid parking fines. Bolstered by this revelation they pressed on. Soon they uncovered the names of several employees who were still listed on the register but who had not paid a variety of fees, a few of whom appeared in the files of the alcoholic and drug abuse program. When this finding was leaked to the public, the city employees, of course, were furious. They demanded to

know who had authorized the investigation. The answer was that no one knew. Later, city officials established rules for the computer center to prevent this form of invasion of privacy from happening again [5].

The danger with information matching is that there is no limit to what one can do with the collected information, and no one knows what the profiles built from the matched information will be used for and by whom. Hundreds, maybe thousands, of databases with individual records are gathered from an individual over a lifetime. Can you recall how many forms you have filled in since you were a child? They may be in the thousands. Each one of these forms contains a set of questions asking for specific information about you. Each time an individual gives a certain answer to any one of these questions, the answer is used to establish a link with hundreds of other databases [5]. Hundreds, perhaps thousands, of databases have personal Social Security numbers. Such databases include driver's records, vital statistics, Social Security administration, medical records, schools, work, and public local, county, state, and federal databases. With the Social Security number as the search key, all these databases can very easily be linked together. In addition to these two links, many other keys can be used as links between databases.

The threat to information matching does not originate only from linking individual records in different databases: it also can come from erroneous or outdated (stale) information. Errors can enter information in basically three areas: (i) at the source, where it occurs mainly through incorrect input such as typing the letter "l" of the alphabet instead of the numeral "1" (one); (ii) during transmission because of transmission interference; and (iii) at the sink, mainly as a result of poor reception. Information becomes stale when it becomes outdated. Unfortunately, erroneous and stale information is frequently used. For example, in the U.S. alone, according to Mason, more than 60,000 local and state agencies by 1986 had routinely provided data to the National Crime Information Center, where on a daily basis close to 400,000 requests were made to the center from law enforcement agents across the country. However, studies showed that the data were in error 4–6% of the time [5]. Thus, an equal number of requests from law enforcement agents were filled with false information, probably placing many innocent individuals in awkward situations. Another example, which may not involve crime information, would be erroneous information collected by a credit reporting agency and used in approving such services as loans, mortgages, and credit cards. If stale information is used, there is a danger that an individual could be denied credit unfairly, and it is widely known how difficult it is to remove that stale information from an individual's credit record.

4.4.6 Privacy Protection and Civil Liberties

Perhaps there is no one agreed-upon set of civil liberties. Many rights scholars have different sets of rights that they put under the umbrella of civil liberties. But the most accepted set of civil liberties are grouped into the following four categories:

(i) criminal justice, that includes police powers, personal liberty, and the right to a fair trial; (ii) basic freedoms of speech, assembly, association, movement, and no discrimination; (iii) freedom of information; and (iv) communications and privacy.

Rapid advances in computer technology, and in particular the advent of the Internet, have all created an environment where detailed information on individuals and products can very easily and cheaply be moved, merged, compared, and shared. With the help of sophisticated network scanning and spying software such as STARR, FreeWhacker, Stealth Keyboard Logger, Snapshotspy, Surf Spy, Net Spy, and NPC Activity Monitor, no personal information on any computer on any network is safe.

Although this is good for law enforcement agencies such as the local police and FBI to track down criminals, and for banks to prevent fraud, and businesses to move data and process customer orders quickly and efficiently, the accessing and sharing of personal data by companies, associations, government agencies, and consumers without an individual's knowledge is a serious threat to the security and well-being of the individual. So, there must be ways to take precautions to protect against the misuse of personal information without consent. We have already indicated that personal privacy is a basic civil liberty that must be protected as is any other civil liberty such as the right to free speech. In many countries, there are guidelines and structures that safeguard and protected privacy rights. These structures and guidelines, on the average, fall under the following categories:

1. *Technical*: Through the use of software and other technically based safeguards, and also by education of users and consumers to carry out self-regulation. For example, the Electronic Frontier Foundation has the following guidelines for online safeguards [5]:
 - (a) Do not reveal personal information inadvertently.
 - (b) Turn on cookie notices in your Web browser, and/or use cookie management software or infomediaries.
 - (c) Keep a “clean” e-mail address.
 - (d) Do not reveal personal details to strangers or just-met “friends.”
 - (e) Realize you may be monitored at work. Avoid sending highly personal e-mails to mailing lists, and keep sensitive files on your home computer.
 - (f) Beware of sites that offer some sort of reward or prize in exchange for your contact or other information.
 - (g) Do not reply to spammers, for any reason.
 - (h) Be conscious of Web security.
 - (i) Be conscious of home computer security.
 - (j) Examine privacy policies and seals.
 - (k) Remember that you alone decide what information about yourself to reveal—when, why, and to whom.
 - (l) Use encryption!
2. *Contractual*: Through determination of which information such as electronic publication, and how such information is disseminated, are given contractual

and technological protection against unauthorized reproduction or distribution. Contractual protection of information, mostly special information like publications, is good only if actions are taken to assure contract enforceability.

3. *Legal:* Through the enactment of laws by national legislatures and enforcement of such laws by the law enforcement agencies. For example, in the United States the following acts are such legal protection instruments [2, 6]:
 - (a) Children's Online Privacy Protection Act.
 - (b) Consumer Protection Act.
 - (c) Freedom of Information Act (1968) as amended (5 USC 552).
 - (d) Fair Credit Reporting Act (1970).
 - (e) Privacy Act (1974): regulates federal government agency record keeping and disclosure practices. The act allows most individuals to seek access to federal agency records about themselves and also requires that personal information in agency files be accurate, complete, relevant, and timely.
 - (f) Family Educational Right and Privacy Act (1974): requires schools and colleges to grant students or their parents access to student records and limits disclosure to third parties.
 - (g) Tax Reform Act (1976): restricts disclosure of tax information for nontax purposes.
 - (h) Right to Financial Privacy Act (1978): provides bank customers the privacy of financial records held by banks and other financial institutions.
 - (i) Electronic Funds Transfer Act (1978): requires institutions providing EFT to notify its customers about third-party access to customer accounts.
 - (j) Privacy Protection Act (1980): prevents unannounced searches by authority of press offices and files if no one in the office is suspected of committing a crime.
 - (k) Federal Managers Financial Integrity Act (1982).
 - (l) Cable Communications Policy Act (1984).
 - (m) Electronic Communication Act (1986): broadens the protection of the 1968 Omnibus Crime Control and Safe Streets Act to include all types of electronic communications.
 - (n) Computer Matching and Privacy Protection Act (1986): sets standards for the U.S. government computer matching programs, excluding matches done for statistical, law enforcement, tax, and certain other causes.
 - (o) Computer Security Act (1987).
 - (p) Video Privacy Protection Act (1988): prohibits video rental stores from disclosing which films a customer rents or buys.
 - (q) Driver's Privacy Protection Act (1994): prohibits the release and use of certain personal information from state motor vehicle records.
 - (r) Telecommunication Act (1996): deregulates the cable and telephone companies to enable each company to become involved in the business of the other.
 - (s) Medical Records Privacy Protection Act (1996):
 - (i) Recognizes that individuals possess a right of privacy with respect to personally identifiable health information

- (ii) Provides that this right of privacy may not be waived in the absence of meaningful and informed consent, and
 - (iii) Provides that, in the absence of an express waiver, the right to privacy may not be eliminated or limited except as expressly provided in this act.
- (t) Digital Millennium Copyright Act (2000).
- (u) The Gramm–Leach–Bliley Financial Services Modernization Act (2000).

4.5 Ethical and Legal Framework for Information

4.5.1 Ethics and Privacy

The issues involving ethics and privacy are many and cover wide areas including morality and law. The rapid advances in computer technology and cyberspace have resulted in rapid changes in these issues and the creation of others. For example, before the Internet, the best way to correspond with a colleague was to either write or type a note, mail it, and, of course, trust a postal carrier. Your worry was not that the carrier would snoop and read its contents, but whether the carrier would deliver it in a timely fashion. Many people never worried because they knew that tampering with mail was a federal offense.

Now, however, with the advent of the Internet and electronic messages, confidentiality is a great concern. Computer technology has raised more privacy questions than it can answer. Is there any confidentiality in electronic communication? Is anything that goes in the clear over public communication channels secure anymore? Are current encryption protocols secure enough? What laws need to be in place to secure any one of us online? Who should legislate them? Who will enforce them? We need a first an ethical framework resembling the one we developed in Chap. 2. But in addition to this, we also need a legal framework. Both these frameworks would probably help. The questions are who will develop these frameworks? and who will enforce them?

Discussion Issues

Attempt to draft an ethical framework discussed here. What do you need to include?

What should be in the legal framework? Who should enact the laws in the framework?

4.5.2 Ethical and Legal Basis for Privacy Protection

The explosion of interest in the Internet, with growing numbers of people obtaining access to it, has also increased the potential for Internet-related crime. The arrest of Kevin D. Mitnick, one of the Federal Bureau of Investigation's (FBI) most wanted computer criminals in 1995, ignited anew the debate on the issue of ethics and security. Mitnick was arrested by the FBI after several years on the agency's most wanted computer criminals list. His arrest was a result of months of work by Tsutomu Shimomura, a renowned cybersleuth.

Mitnick's acts, and many after his, highlight how vulnerable the Internet is and how vulnerable we are whenever we use it. Security and ethical issues do not and should not come into play only when a crime is committed. These issues are also raised when individuals and companies act in ways that are considered harmful or have the potential of being harmful to a sector of society. Consider online postings, for example. As the Internet grows, companies and individuals are flocking to the Internet to post and advertise their wares. Until recently, the focus of Internet security and ethics was on pornographic images accessible to children. But of late a multitude of concerns have sprung up as new Internet technologies and services have sprung up, for example, online social networks.

What is the way forward? How can we ethically and legally encounter the new Internet technologies and services without interfering in people's loves and businesses as they use these new services?

Exercises

1. Define security and privacy. Why are both important in the information age?
2. What is anonymity? Discuss two forms of anonymity.
3. Discuss the importance of anonymity on the Internet.
4. Is total anonymity possible? Is it useful?
5. Develop two scenarios: one concerned with ethical issues involving security, and the other considering ethical issues involving privacy.
6. Is personal privacy dead? Discuss.
7. List and discuss the major threats to individual privacy.
8. Identity theft is the fastest growing crime. Why?
9. Why is it so easy to steal a person's identity?
10. Suggest steps necessary to protect personal identity.
11. Governments are partners in the demise of personal privacy. Discuss.
12. Anonymity is a double-edged sword. Discuss.
13. Are the steps given in Sect. 4.4.5 sufficient to prevent identity theft? Can you add more?
14. What role do special relationships play in identity theft?
15. Modern-day information mining is as good as gold! Why or why not?
16. How do consumers unknowingly contribute to their own privacy violations?

-
17. How has the Financial Services Modernization Act helped companies in gathering personal information?

References

1. G. Seven, Lex Luthor and the legion of doom/hackers presents: identifying, attacking, defeating, and bypassing physical security and intrusion detection systems. *Lod/H Tech. J.* **1**(3) (1990)
2. *Second Amended Verified Original Petition and Application for TRO and Temporary Injunction* (Universal Image, Inc. v. Yahoo, Inc.). <http://www.tomwbell.com/netlaw/universal/yahoo.html>
3. R. Will, Sites targeted for privacy violations. *USA Today*, (2000)
4. W. Rodger, Sites targeted for privacy violations. *USA Today*, 13 June 2000
5. R. Mason, Four ethical issues of the information age, in *Ethical Issues in Information Systems*, ed. by R. Dejoie, G. Fowler, D. Paradice (Boyd & Fraser, Boston, MA, 1991)
6. K. Landon, Markets and piracy. *Commun. ACM* **39**(9), 92–95 (1996)

Further Reading

7. Privacy lost: the demise of privacy in America. <http://dorothyseeseonline.tripod.com/newsline/id3.html>
8. Spying on employees: big brother has his eye on you ... snooping sister is supervising you. <http://www.successunlimited.co.uk/related/snoop.htm>
9. Tactics of a high-tech detective. *New York Times*. <http://www.takedown.com/coverage/tactics.html>
10. W.H. Auden, The unknown citizen, in *Literature: An Introduction to Fiction, Poetry, and Drama*, 6th edn., ed. by X.J. Kennedy, D. Gioia (Harper Collins, New York, 1995), pp.611–612
11. D. Eisenberg, Who is reading your bills? A court ruling on privacy riles the FCC. <http://www.cnn.com/ALLPOLITICS/time/1999/08/30/privacy.html>
12. T. Garrison, Latest surveillance leaves nothing to chance: exploring the dark side of the digital home. *Realty Times*, 10 Feb 1999. <http://realtytimes.com/rtnews/rtipages/19990210digitalhomes.htm>
13. I. Lynch, Snooping bill to become law by November, 27 July 2000. <http://www.vnunet.com>
14. S. McClure, J. Scambray, Forget the firewall: guard your garbage against ‘Dumpster Diving’ hackers. *LISTSERV@SecurityFocus.com*. Friday, 7 July 2000
15. K. Mieszkowski, Big brother in the flesh: new technology could make us all a part of the collective, permanently supervised from above, 21 Sept 2000 edn. <http://www.lasvegasweekly.com>
16. J. Rachels, Why privacy is important, in *Ethical Issues in Information Systems*, ed. By R. Dejoie, G. Flower, P.A. Radice (Boyd & Fraser, Boston, 1991)
17. S. Schiesel, On the web, new threats to young are seen. *New York Times*, 7 Mar 1997



Intellectual Property Rights and Computer Technology

5

Abstract

This chapter discusses the foundations of intellectual property rights and how computer technology has influenced and changed the traditional issues of property rights. The reader is immersed into a discussion of controversial issues of ownership in a rapidly amalgamating global cultures, languages, beliefs and values as a result of rapid globalization technologies like telecommunication that is casting a far and wide net that is likely, in the near future, to create one global commons. The controversial issues focused on here include the politics and psychology of ownership and the changing infringement landscape. Another issue of interest in our focus is intellectual property crime (IPC), activities that involve infringement, counterfeiting, piracy of products and services for profit without permission from the creator, misappropriation, misrepresentation, corruption and bribery, and espionage.

Learning Objectives

After reading this chapter, the reader should be able to

1. Distinguish among patent, copyright, and trade secret protection.
2. Discuss the legal background of copyright in national and international law.
3. Explain how patent and copyright laws may vary internationally.
4. Outline the historical development of software patents.
5. Discuss the consequences of software piracy on software developers and the role of relevant enforcement organizations.

Scenario 4: Cybersquatting: Is It Entrepreneurship or Intellectual Theft?

Just before the 2000 New York senatorial campaign, Chris Hayden paid \$70 each for the exclusive 2-year rights to the following Internet addresses: www.hillary2000.com, www.hillaryclinton2000.com, and www.clinton2000.com. A few weeks later, Mrs. Hillary Clinton, the then U.S. first lady, declared her candidacy for the state of New York senatorial race. The Clinton campaign team wanted her presence on the web, but they could not use any of the three names, though they rightly belonged to Mrs. Clinton. Deciding not to challenge Mr. Hayden in the middle of an election campaign, the team opted to buy the rights for www.hillary2000.com from Mr. Hayden. However,

Mr. Hayden decided to engage a broker to demand \$15,000 for the use of the name [1].

Cybersquatting, as the practice of grabbing somebody's name and registering it with an Internet registration company in anticipation of reaping huge rewards, is becoming widespread.

Discussion Questions

1. *Is Mr. Hayden violating Mrs. Clinton's intellectual rights?*
2. *Can Mr. Hayden claim free speech protection for the use of the names?*
3. *Should there be laws to make the practice illegal?*

5.1 Definitions

Intellectual property (IP) broadly describes tangible things such as ideas, inventions, technologies, artworks, music, and literature to which one can claim ownership. Ownership of IP for any of these things may result in economic gain as rewards to personal initial investments before they acquire value. It is a set of legal rights that result from intellectual activity in the industrial, scientific, literary, and artistic fields [2]. Intellectual property rights (IPR) are legal rights bestowed on an individual or a group that created, designed, or invented the activities or processes which led to the intellectual property in domains such as science and technology, business, industry, and the arts. These legal rights, most commonly in the form of patents, trademarks, and copyright, protect the moral and economic rights of the creators, in addition to the creativity and dissemination of their work [2].

5.2 Computer Products and Services

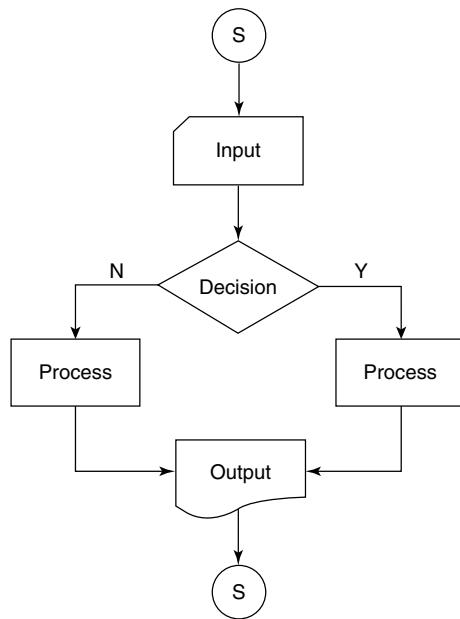
Computer products consist of those parts of the computer you can see and touch (e.g., the keyboard, CPU, printer, and monitor). These parts are considered products because they have tangible form and intrinsic value. A service is an act carried

out on behalf of someone, usually a customer. If the service is to be paid for, the provider must strive to please the customer; it is crucial. If the service is not to be paid for, the act must then be performed to the liking of the provider.

Services have intrinsic value to the customer or recipient but have no tangible form. For example, a patient going to a doctor for treatment receives a service that has an intrinsic value, especially if the patient gets better, but it has no tangible form. A computer service can take the form of repairing a computer product and/or configuring and installing a computer network, neither of which has a tangible form but does offer considerable intrinsic value to the owner. Computer products can be defined easily because they have tangible form and intrinsic value. Services can also be defined easily because they have intrinsic value to the customer or the recipient, and in most cases to the provider, although they have no tangible form. Computer software, however, cannot be so easily classified as either a product or a service. Since it entered the marketplace, therefore, legal protection of computer software has been very problematic.

Computer software is a set of logical instructions to perform a desired task. This logical sequence follows an algorithm. The development of this sequence of instructions goes through the following phases:

1. *Logic map*: The plan of the idea, process, or an algorithm to accomplish the task. A plan is a flowchart of logic control with four major stations, namely, the input, output, processing, and decision box connected by arrows indicating the direction of flow of logic (see the flowchart in Fig. 5.1). Within the flowchart itself, there may be loops that express repetition of certain parts of the logic of that flowchart. Effective implementation of the flowchart achieves the desired effects of the program for the designated task.
2. *The source code*: A result of the implementation of the flowchart, turning the flowchart into a set of instructions using a programming language of choice.
3. *The object code*: The second stage of the implementation of a flowchart in which the source code, with the help of either a compiler or an assembler, is turned into strings of zeros and ones. In this form, the program is referred to as an object code. This step is not yet fully an operational form of the program because at this stage the program lacks variable addresses that have yet to be reconciled. It is also missing library routines that are added in the next stage.
4. *Memory-based executable code*: After the object code is finished, it is passed to the linker, another one of the system programs, whose job is to look for all missing variable addresses and library and personal routines and include them in the object code to produce a load module. The load module is now ready to be loaded if execution is needed. The job of loading the module is done by another system program called a loader. Most programs bought from vendors and software producers are in this executable form. The program at this stage has intrinsic value to the software producer and probably to the buyer, but it has no tangible form if you do not consider the medium it is on.
5. *Microcode*: Another form of executable code, but differing from the type of code we have just described, this code is not loaded in physical memory. It

Fig. 5.1 Flowchart

is coded and loaded on ROM (the read-only memory of the computer, which cannot be written on by the user) or burned into the computer hardware at the time of manufacture. If this code is loaded in ROM it can only be erased electronically or by using ultraviolet light. If, however, it is incorporated into the hardware, it is not easily changed and it cannot be erased. In this form the program is referred to as microcode. For a program in hardware, execution is normally achieved through the logical flow in the assemblies of hardware components.

In support of either hardware-based or memory-based programs, programmers, usually the creators of these programs, write documentation and technical manuals to accompany the program and help the user of the software. Our references to computer programs here, therefore, include memory-based and hardware-based programs together with the technical manuals and all related documentation.

In our definition of computer software, whether hardware based or memory based, including technical writings, note that computer software, if it is considered at the execution stage without the technical documentation, has an intrinsic value both to the developer and to the buyer, but it may not have a tangible form unless you consider the medium it is on (e.g., the disk). For example, during tax filing season when you buy a tax program to help you with taxes, you either download the program or get a couple of CDs with the program on them and a number of manuals and flyers. You can ignore the flyers in your package because they are usually for commercial purposes, but the purpose of the manuals is to help you learn how to use the program. You can touch the manuals, the CDs, and so on,

but you cannot touch the program itself. That is, the manuals and CDs all have a tangible form and probably some intrinsic value, but the program itself does not have a tangible form, although it has the most intrinsic value. In this case, we can classify such software as a service.

Not having a tangible form, however, does not by itself rule out software as a product. According to Johnson [3], the courts do not equate products with what is tangible. Courts have defined things such as energy and leases as products, although none has a tangible form. So there are cases when we can consider software as a product.

In his article “Negligence for Defective Software,” Prince [4] puts software into three categories. The first category includes off-the-shelf software such as Windows and others that one can buy ready to use with no alterations allowed by the producer. This category he calls the “canned” software. The customer gets it as is. The second category is the software specifically ordered by the customer from the software house or producer to fit the customer’s very specific needs, similar to going to your physician for a specific illness. The third category is software the customer buys off the shelf, but with changes allowed, or the customer adds certain parts to the software to meet some specific needs. According to Prince, category 1 software is considered to be a product, whether it has a tangible form or not; category 2 software is considered to be a service; and category 3 software is a new class he calls “mixed case.”

5.3 Foundations of Intellectual Property

Gaining the skills to provide computer technology products, services, and software requires a considerable investment in both time and money. Thus, the individuals who do this work should reap financial rewards for their efforts. Such rewards create an atmosphere of creativity and competitiveness, which in turn creates jobs that drive the economy. This creativity must therefore be protected, for if it falters because of lack of protection, then the economy of the country falters along with it.

Computer technology in particular was born of this individual creativity and the adventurism of young entrepreneurs. To encourage these innovators, society must protect their efforts and resources. To do this, a specific set of rights, collectively known as intellectual property rights, has been recognized, and laws have protecting intellectual rights been enacted and extended to cover software by different countries and groups of countries to protect those rights.

Intellectual property rights form a wide scope of mechanisms that include copyrights, patents, trademarks, protection of trade secrets, and, increasingly, personal identity rights. Each of these instruments of protection is regulated by a body of laws and statutes we discuss throughout this chapter. Unfortunately, some of these laws are not universal; they only apply in one country. And even within the United States, the same laws may not apply in all states. In particular, we look

at intellectual property rights as they apply to computer products, services, and software.

5.3.1 Copyrights

Internationally, copyright is a right, enforceable by law, accorded to an inventor or creator of an expression. Such expressions may include creative works (literary, dramatic, musical, pictorial, graphic, artistic) together with audiovisual and architectural works and sound recordings. In general, every original work that has a tangible form and is fixed in a medium is protectable under the copyright law. The history of copyright laws can be traced back to eighteenth-century England with the so-called statute of Queen Anne around 1710 setting a pattern for formal copyright statutes. England was followed by the United States in 1790 when the first U.S. copyright law was enacted by Congress, and by France in 1793 [5].

Since then, copyright laws have spread worldwide. Examples of international copyright bodies include the Berne Convention in 1886, of which the United States was not a signatory until 1989, the 1952 Universal Copyright Convention (UCC), and the Berne and Paris conventions in 1971. To ensure that conventions stay current and signatory countries observe them, a number of world bodies have been created mainly to administer the conventions. The World Intellectual Property Organization (WIPO) created in 1967 was the first to be charged with such a task. Later, the UN Educational Scientific and Cultural Organization (UNESCO) together with WIPO were assigned to administer the UCC, and finally the World Trade Organization (WTO) is now charged with administrating the Trade-Related Aspects of Intellectual Property Rights (TRIPPS) agreement concluded under the Uruguay round of the General Agreement on Tariffs and Trade (GATT). Besides these large and more comprehensive organizations, there are also numerous small regional organizations such as the North American Free Trade Agreement (NAFTA) [5].

These organizations, together with national legislatures, keep these conventions and national copyright acts current through amendments. For example, in the United States the 1790 copyright law was amended in 1831 and then again in 1870. In 1909, Congress enacted the Copyright Act, which underwent two amendments, and in 1976 Congress enacted the current Copyright Act, which came into effect in 1978. This act has already undergone several amendments, mainly because of advances in computer technology.

Each country has its own requirements for the issuance of a copyright. In the United States, for example, there are three requirements for copyright protection of a work under the 1978 U.S. Copyright Act: originality, fixation, and expression; in Canada, it is originality and fixation. The U.S. copyright laws cover all original works fixed in tangible forms regardless of medium, and such works must be expressions, not ideas [4]. The scope of works or creations meeting these criteria is wide: it includes artistic, pictorial, graphic, and sculptural works; musical and sound recordings; audiovisual works including television and motion pictures; and

literary works and other printed materials such as books, greeting cards, journals, flyers, and leaflets; the list goes on.

However, a number of creative fixable works are excluded from this extensive list because they are considered either trivial or utilitarian. The list for these is also a long one: it includes such items as calendars, schedules of events, scorecards, ideas, facts, names, common language phrases, titles, and blank forms. Although some of these may not be protected by the copyright laws, they may be protected somewhere else, for example, by trademark or patent laws.

5.3.1.1 Works in the Public Domain

When the copyright on a work expires, that work goes into the public domain. Other works in the public domain include those owned by governments, non-copyrightable items we listed earlier such as ideas and facts, works intentionally put in the public domain by the owner of the copyright, and works that lost copyrights for various reasons before the copyrights expired.

Works in the public domain are not protected by the copyright law and can be used by any member of the public without prior permission from the owner of the work. Examples of such works in the United States include works published before 1978 whose copyright has not been renewed and, therefore, have no valid copyright notice. A copyright notice consists of a copyright symbol denoted by ©, the word “copyright,” the year the copyright was granted, and the name of the copyright owner. For example: copyright © 1995 John Mukasa.

5.3.1.2 Application for a Copyright

For authors and creators of works who need this kind of protection, the process begins with an application to the copyright office. Each country’s copyright office has different requirements. The U.S. Copyright Office requires an applicant to include with the application a copy of the work for which a copyright is sought and to file for copyright within 3 months of the first distribution of the work.

Upon receipt of the application by the Copyright Office, it is reviewed to ensure it meets the three criteria of originality, fixation, and expression for the issuing of a copyright.

Fixation, a remnant of the Gutenberg era, and now the most controversial element in the current debate about intellectual property rights in the digital age, refers to the tangible form in which the creation is perceived by others. For example, computer programs are fixed in binary code. For performing arts such as drama, the script is the fixation. In the fine arts, the painting or the sculpture is the fixation. It is important in creative work to have fixation because it clearly demonstrates and defines the tangible form of the creation and the domain and parameters of such a creation.

Similar to fixation, the protection of a creation requires proof of originality. The originality requirements differentiate among facts, ideas, and expressions as creations. Facts are considered common property for all humanity; no one has a right to an unknown or known fact because it is not considered an invention; the same principle applies to theories, mathematical, scientific, and others. Ideas,

as are facts, are also considered common property and therefore are not copyrightable. Thus, originality is only possible through expressions. Such expressions may include ideas, theories, and other inventions. The packaging must be original if protection is sought. Packaging includes remakes of protected works. This type of packaging is what scholars of copyrights call derived works. For an interesting discussion of this, refer to *The Copyright Book* by Strong [6]. The review process is very extensive and thorough and takes some time before it is complete. Upon approval, the recipient must place a notice of copyright ownership in all parts and copies of the work.

5.3.1.3 Duration of a Copyright

In the United States the duration of copyright protection falls into two periods: those copyrights granted before the 1978 Copyright Act and those granted after that date. If a copyright was received for a published work before 1978, that copyright lasts for 75 years after the date of issuance. For unpublished works, the copyrights will expire on December 31, 2002 regardless of when they were issued. If the copyright was received after 1978, the work remains protected by copyright laws for the lifetime of the author plus 50 years. In the case of more than one author of the work, the protection lasts for the lifetime of the longest living author plus 50 years. For all works made for hire, that is, works made as part of contracted employment, the coverage lasts 75 years from the date of the first publication or 100 years from the date of creation [5].

5.3.2 Patents

In contrast to the copyright, which protects expressions, patents protect inventions or discoveries. In the United States, patent rights are protected just like copyright rights.

In many countries, patent protection rights, as are those of copyrights, are provided for by the constitution. The U.S. Constitution, for example, states it this way: “to promote the progress of science and useful arts, by securing for limited times to authors and inventors the exclusive right to their respective writing and discoveries.” In accordance with this, Congress has enacted and continually revised patent laws to suit the times. These laws give inventors or discoverers living within U.S. borders and territories the exclusive rights to make, use, or sell their invention or discovery for a specific period of time, so long as there was full disclosure of the invention or discovery at the time the patent application was filed.

Because of the disclosure requirement that every patent applicant must meet, the patent is more of a contract between the inventor or discoverer and the government. For this contract to be binding, each party to the contract must keep its part: the government’s part to protect the exclusive rights of the inventor or discoverer while such a person recovers his or her investments for a period of time, that of the inventor or discoverer to make a full disclosure to the government of the invention or discovery. With this disclosure the government makes sure the public benefits

from the discovery or the invention while it is still under patent protection and more after.

5.3.2.1 What is Patentable?

In the United States, an invention or discovery is patentable if it meets two basic requirements. The first requirement is that the invention or discovery for which the patent is sought is new and useful, or is a new and useful improvement of any of the following: process, manufacture (covering all products that are not machines), and machine (covering all mechanisms and mechanical products and composition of matter, which includes all factory-manufactured life forms).

The second requirement is that the invention or discovery must satisfy the following four conditions, and all must apply:

1. *Utility*: An invention or discovery serves a basic and minimum useful purpose to the general public or to a large percentage of the public without being a danger to the public, illegal, or immoral.
2. *Novelty*: The invention or discovery for which a patent is sought must be new, not used, known, or published somewhere before.
3. *Nonobviousness*: The invention or discovery for which patent protection is sought must not have been obvious to anyone with ordinary skills to produce or invent in its disclosed form.
4. *Disclosure*: There must be adequate disclosure of the product for which a patent is sought. Such a disclosure is often used by the Patent Office in its review to seek and prove or disprove the claims on the application form and also to enable the public to use the invention or discovery safely and gainfully after the period of protection.

5.3.2.2 Application for a Patent

In many countries, the process of obtaining a patent begins with the filing of an application with the patent office. As we already discussed, the application must give a clear and detailed disclosure of the invention or discovery including its workings, experiments made, data used, results obtained, safety record, and effectiveness if used properly. Its weaknesses, if observed, and all pertinent information that may be required if the Patent Office is to carry out a similar experiment must also be submitted.

5.3.2.3 Duration of a Patent

After the review process is completed—and this may take some time depending on the disclosure provided and the type of invention or discovery—the patent is then issued to the applicant for the invention, and only for that invention, not including its variations and derivatives. The protection must last for a number of years, 17 years in the United States. During this time period, the patent law protects the inventor or discoverer from competition from others in the manufacture, use, and sale of the invention or discovery.

5.3.2.4 Public Domain

The patent law does not protect ideas, but only the process of carrying out an idea. Competitors may take the same idea and use a different process to arrive at their own inventions or discoveries, which can then be patented as novel. When the patent protection expires, the patent together with all disclosures go into the public domain for anyone to use.

5.3.3 Trade Secrets

A trade secret is information that gives a company or business a competitive advantage over others in the field. It may be a formula, a design process, a device, or trade figures. Thus, there is no one acceptable description or definition of trade secrets. The generic definition is that it is a collection of information in a given static format with strategic importance. The format may be a design expressing the information, a formula representing the collection of information, a pattern, a symbol, or an insignia representing the information. Whatever the format the collected information takes, it must have given or offered an advantage to the owner which places that owner a degree above the competition.

In the United States, in contrast to the other intellectual properties we have described so far, trade secrets have no federal protection. All trade secret laws are state laws.

However, trade secret owners are further protected by express or implied contract laws and laws of unfair competition, which are backed by federal statutes.

5.3.3.1 Characteristics of Trade Secrets

Because it is difficult to define a trade secret, it is important that we characterize what makes a trade secret. According to Neitzke [7], a trade secret is characterized by the following:

1. The extent to which the information is known outside the business. If many people outside the company or business know or have access to the collection of information that constitutes the trade secret, then it is no longer a trade secret.
2. The extent of measures taken by individuals possessing the trade secret to guard the secrecy of the information. If the information is to remain known by as few people as possible, there must be a detailed plan to safeguard that information and prevent it from leaking.
3. The value of the information to the owner and to the competitor. If the collection of information forming the trade secret has little or no value to the competitor, then it can no longer be a trade secret because it offers no definite advantage to the owner over the competitor. It does not matter whether the owner values the information; so long as it is not valued in the same way by the competitor, it is not regarded as a trade secret.
4. The amount of effort or money spent by the owner to develop or gather the information. The logic here is usually the more money the developer puts in a

project, the more value is placed on the outcome. Because there are some information or project outcomes that do not require substantial initial investments, the effort is what counts here.

5. The ease or difficulty with which the information could be properly acquired or duplicated by others. If it will take much effort and money to duplicate the product or the information, then its value and therefore advantage to the competitor diminishes.

The conditions that characterize a trade secret are in direct conflict with the requirements of a patent. Remember the main requirement for obtaining a patent is the full disclosure of all the information surrounding the product and its workings; this directly conflicts with the need for secrecy in trade secrets. So, the patent applicant cannot claim a patent and at the same time claim protection using the trade secret laws.

5.3.3.2 Duration of Trade Secrets

Trade secrets have an indefinite life of protection so long as the secrets are not revealed.

5.3.4 Trademarks

A trademark is a label identifying a product or service. It is a mark that attempts to distinguish a service or a product in the minds of the consumers. The label may be any word, name, picture, or symbol. It is very well known that consumers tend to choose between products through association with the product's brand name. For example, the Golden Arch is a trademark for McDonald's restaurants. There are many other fast-food restaurants (e.g., Burger King), but none of them can use the Golden Arch as their trademark. The Golden Arch differentiates McDonald's from all other fast-food restaurants and may give it an advantage over its competitors in the industry. Because trademarks are used by consumers to choose among competing products, they are vigorously protected by their owners.

Differing from patents and copyrights, however, trademarks are not so protected and enshrined in constitutions. For example, in the United States, trademark laws, as are trade secrets, are based on state statutes. At the federal government level the trademark laws can be found in the Lanham Act and the new Trademark Cyberpiracy Prevention Act 1999 [8].

Although the patent gives owners the exclusive right to use, sell, and make use of the invention or discovery, and the copyright law gives owners the exclusive rights to copying their works, the trademark gives its owner the right to prevent others, mostly competitors, from using the same or similar symbol to market their products.

5.3.4.1 Categories of Trademarks

Trademark is a general term that includes a service mark, a certification mark, and a collective mark. A service mark is usually used in the sale or advertising of a service. It is supposed to uniquely identify that service. A circle with checked lines distinguishes AT&T telecommunication services from those of BTT, MCI, and many others. A certification mark is used as a verifier or to authenticate the characteristics of a product, a service, or group of people who offer a certain service. For example, colleges attach seals to diplomas as marks to certify the educational attainment of the holders. A collective mark is mainly used by a group of people to indicate membership in an organization or association. For example, people who take and pass certain specialty examinations can use a mark such as CPA, Ph.D., or M.D. to indicate they belong to those groups.

5.3.4.2 Characteristics of Trademarks

It is said that a picture is worth a thousand words. So, it is assumed by trademark owners that a symbol is worth 1000 words and, therefore, their marks are always saying something to the customers, at least in theory. A variety of marks are used by product and service companies to enhance the commercial value of their products or services in the eyes of the public by association. General wisdom is that the more recognizable is the mark, the more valuable will be the product or service.

In addition to categorizing trademarks by what they cover, as we did in the last section, let us also group them according to what they say. Trademarks as symbols of sales to the consumer are generally supposed to tell the consumer about the services or the products they are intended to boost. The impression the mark gives to a consumer, the likelihood of the existence of such a mark, and the ease of obtaining registry put it in one of the following characteristic groups:

- *Arbitrary marks*: Trademark symbols that say nothing about the product or service are usually used arbitrarily with a product or service, but over time they become associated with that product or service. Most arbitrary marks are one or more words or a collection of letters already in linguistic use but with no associated meaning. Many established trademarks start as arbitrary marks, and consumers eventually come to associate them with the product or service. For example, McDonald's Golden Arch may have had no meaning at the beginning and still may have no meaning to "outsiders" unless they know of the association.
- *Suggestive marks*: Symbols or writings that are usually in the public domain may be twisted by people to say something about their products or services. These marks may suggest to the customer the features, qualities, and other characteristics of the product or service. A good example here would be a mark such as "GolfRight" as a trademark for a company manufacturing a new brand of golf balls. The two words "golf" and "right" were taken out of the public domain and combined to describe the product with the creation of a new word, "GolfRight."

- *Descriptive marks:* These usually contain a description of the intended purpose of the mark but say nothing about the product or service. For example, if you create a program that you think simplifies the tax-preparing process, you may use a trademark called “Easy-Tax.”
- *General marks:* New marks are unrelated and with no suggestive features, qualities, and characteristics of the products or services they are said to represent. In contrast to the arbitrary marks, general marks are not linguistically bound. A general mark could be any symbol. General marks are desirable because they are easy to register, because the likelihood of the existence of a similar mark is minimal. An example is the use of a graphic symbol such as an arrow for a product or service.

5.3.4.3 Registration of a Trademark

An application for a trademark must contain and present all relevant information. It must also describe the product or service for which the trademark is being sought, the class of goods and services, and the date of first issue of the mark. Marks are registered only if they meet certain criteria. The core requirement is that the mark must not cause confusion with similar marks used by others. In the United States a mark is registered as a trademark only if it meets the following criteria:

1. It must be in good “taste” for the public: not immoral, deceptive, or illegal.
2. It must not have suggestive connotations to its origin.
3. It must not be a symbol of any recognized country.
4. It must not use people’s likenesses, either after death or if living, without prior consent.

5.3.4.4 Duration of a Trademark

In the United States, a valid trademark is protected for 10 years. If an extension is needed it can be granted for another 10 years.

5.3.5 Personal Identity

Identity theft is a crime committed when one misrepresents oneself, with or without success, as another person to get the victim’s information so that the perpetrator can receive goods and services in the name of the fraud victim. Identity theft is now one of the fastest growing crimes in the United States and in a number of other countries as well. Although it is still not considered to be a crime in some countries, national legislatures are in full gear enacting laws to criminalize it. When it happens, it takes probably an instant, but it can take a long period of time before it is discovered. By then, the information misuse, financial loss, and psychological damage can be devastating, but this is nothing compared to the agony one goes through trying to control, manage, and recover from the damage caused. Doing this can sometimes take years and be very costly.

Techniques to steal personal identity include the following [9]:

- (i) Advertising in newspapers and mostly on the Internet. The most common technique now, pretext calling, is where people misrepresent themselves as law enforcement agents, social workers, and potential employers to obtain the private data of others from banks and other financial institutions.
- (ii) From readily available how-to books and discussion groups, perpetrators get foolproof methods of wangling financial information from bank employees.
- (iii) Use of telemarketing scams to trick consumers into revealing personal data.
- (iv) Abundant authentic-looking fake IDs, including Social Security cards, birth certificates, and driver's licenses, are on sale online.
- (v) Going through one's trash for personal information.
- (vi) Using the post office to redirect one's mail to a perpetrator's box number.
- (vii) Criminals are increasingly using radio scanners to eavesdrop on personal calls.

5.3.5.1 Prevention

After being the victim of identity theft, it is extremely difficult to straighten out one's record, let alone recover the stolen personal attributes. The best course of action is for individual defense. The following steps are considered minimal but effective:

- (i) Shred all credit card receipts, canceled checks, and other financial documents.
- (ii) Seek employer personal information protection plans.
- (iii) We are leaking vessels of personal information. At every stop we make, we involuntarily give out crucial personal information such as sensitive financial data, telephone numbers, Social Security numbers, and other vital personal data.
- (iv) Where possible, have all your payments deposited electronically in your bank account.
- (v) Periodically check your credit report. It is better still if you review credit reports from all three credit bureaus for erroneous data on your personal credit report. When you get your credit report, look for things such as who is using your information. Check and make sure you know who requests for your information from these companies.
- (vi) Shred all your credit card solicitations and all other mail that bears personal identification.
- (vii) If you become a victim, report the incident to law enforcement personnel.

Although not as effective so far, legislation is also important. The U.S. Congress recently passed a law that makes it a federal crime, punishable by up to 5 years in prison, for anyone to misrepresent himself or herself to obtain someone's private financial data.

5.4 Ownership

As we discuss ownership in this section, and indeed in the whole book, we confine ourselves to intellectual property ownership. Ownership of everything else other than intellectual property is outside our present scope. An idea is novel if it is original, authentic, and new. Inventiveness, creativity, and discoveries are born of individual ideas. Good ideas are the source of substantial benefits to individuals and the public. Before an idea can be useful, however, it must be put into utilizable form, either as a process or as an application. It is this idea in a utilizable form that is the core of intellectual property rights. In many countries the owner of such an application for the idea has a set of legal rights to its expression as a creation, work, invention, discovery, information, or any other form.

Via the copyright law, the patent law, the trademark law, and trade secret statutes, governments have indicated that they can protect owners' rights under certain conditions, and, therefore, legal ownership of creations, discoveries, information, inventions, and the like, is protectable. As we have already seen, the domain of all these rights constitutes intellectual property. Within this domain, these rights are grouped into subsets defining specific areas of interests such as the right to make, use, or sell one's works of discovery or creation. Each such subset is protected by the four well-known instruments we discussed in Sect. 5.3.

5.4.1 The Politics of Ownership

Recently, much has been written about the concept of intellectual property rights, and the issue has been in the news, such as when the U.S. government negotiated with China, the country that many in the West believe has the highest rate of abuse of intellectual property laws. There have been many statements made about the effects of cultural differences between Western countries and other cultures of the world regarding the issue of intellectual property rights.

In fact, this issue alone has become a defining factor between Western and other cultures. Western culture emphasizes individuals, rewards individual achievements, and hence upholds intellectual property issues as a golden egg. Non-Western cultures, in contrast, which emphasize community responsibility, do not understand the West's focus on intellectual property rights. To many non-Westerners, a good manifestation of an idea that benefits a community should be enjoyed by the whole community, not just by one or a few members, because individuals make up the community.

As global economies become more and more intertwined, and as the West continues to keep the lead in technological development, many of the non-Western cultural underpinnings are likely to change as these cultures devour the new imported Western technology. Already a number of countries in Southeast Asia have been forced to abide by the intellectual property laws as dictated by the West.

In addition to the cultural politics of intellectual property issues, there is also a perception controversy. Many people believe the protection of the manifestation of one's ideas by copyrights, patents, trademarks, and trade secrets laws automatically constitutes a monopoly of the benefits that come with the ideas. Because of this misconception, the U.S. Congress and indeed other governments have passed antitrust laws to calm the public. The antitrust laws in themselves prevent or restrict patent, copyright, or trademark holders from collecting large royalties beyond the term of the license by opening up the competition.

5.4.2 The Psychology of Ownership

Whether we grew up in a culture that rewards individual achievements or in those cultures that pride themselves on community achievements, we are aware of the distinct psychology about individual ownership. We tend to classify those items that we individually own, whether they are few or abundant, according to their intrinsic value to us. We may believe our self-worth and status in the community depend on the intrinsic value of the items we own. So the intrinsic value we attach to these items is one of the most important aspects of ownership. Another aspect of ownership is the tangibility of what we own. When what we own has a tangible form with glamour and value to others, whether it has any intrinsic value to us, it tends to raise our status in the community. We therefore gain the respect of others, which in turn affects our self-esteem and our egos.

5.5 Intellectual Property Crimes

An intellectual property crime (IPC) is the act of infringement on the rights of the owners of the intellectual property. IPC refers to all activities that involve infringement, counterfeiting, and piracy of products and services for profit. IPC also includes misappropriation, misrepresentation, cybercrimes, corruption and bribery, and espionage. The cost of intellectual property crimes to industry and nations is huge.

Technological advances have allowed these crimes to grow like wildfire in the past decade because committing these crimes is much easier and the field of crimes has become global, thus decreasing the threat of apprehension. Technology has also increased these crimes because duplicated products are easy to make and the costs are low.

5.5.1 Infringement

In Sect. 5.3 we discussed the legal protection over a domain of rights for an individual's manifested idea. This legal protection is offered in subsets of laws that define the boundaries within which such laws can be enforced. Anybody else

with no rights within this domain is considered an infringer, defined as one moving within the protected domain to claim rights for the use of someone else's manifestation of an idea without permission from the holder of the rights.

This is an abstract concept, and the difficulty in understanding it illustrates the elusiveness of the boundaries of these rights. There are three types of infringements:

1. *Direct infringement*: The infringer knowingly or otherwise makes, uses, sells, or copies a protected item without any alteration.
2. *Inducement infringement*: The infringer intentionally supports infringement activities on a protected item without individually taking part in the infringement activities.
3. *Contributory Infringement*: The infringer takes part in the infringement of a protected item. Let us now look at infringement under each one of the subdomains of the intellectual property domain.

5.5.1.1 Copyright Infringement

Copyright infringement is very difficult to prove. However, U.S. courts have provided guidelines that many courts follow [6]. Here are some of the items that courts look for in an infringement suit:

- Whether the infringer has knowledge or visual contact with the work
- Whether the individual claiming to be the owner has a valid copyright
- Whether the work under dispute is a major revision with substantially new contents of the original or just a variation.

5.5.1.2 Patent Infringement

Similar to copyright infringement, patent infringement is also difficult to detect. Highly sophisticated methods of policing and investigative work need to be laid down and followed. No public law enforcement can be used in these cases. It is purely the effort of the owner of the patents, and he or she must meet all expenses incurred during the investigation and prosecuting of the infringer if caught. Once the infringer is caught and determined guilty by the court, a hefty settlement is collected from the perpetrator. There may also be punitive damages.

Because the policing and investigation can be difficult, lengthy, and expensive, patent owners tend to use a device that uses the public to do the policing for them. They achieve this by using patent markings on their products, for example "Pat." followed by the number of the patent. With this mark on the product, patent owners hope the public will police the marketplace and inform them if they suspect any patent infringement. If the patent owner confirms that an infringement on his or her patent has taken place, the first course of action is usually litigation to collect the damages and most importantly to send a message to the public and mostly to those who had intentions of infringing on the patent to keep off. Another channel

of action open to the patent owner is through an independent arbitrator to obtain some compensation from the infringer.

5.5.1.3 Trademark Infringement

To prove infringement of a trademark, one must prove beyond doubt that the infringer's action was likely to confuse the public. Because of this, it is very difficult to prove trademark infringement. If the owner of the trademark can successfully prove and convince the courts that the infringer's mark has or is likely to cause confusion, then the infringer may be asked to pay any or a combination of the following: monetary awards based on the profits he or she made on the product displaying the mark, losses the owner supposedly incurred as a result of the infringement, and/or punitive damages and legal fees.

5.5.2 The First Sale Doctrine

A copyright owner under the first sale doctrine has the right to distribute copies of copyrighted materials by means of sale, transfer of ownership, rental, release, or by any other means. In the United States, under the first sale doctrine section 109(a) of the Copyright Act, artists, authors, inventors, or discoverers can control subsequent use of their works through a lease or license. Anybody else who uses that work without either a lease or license is an infringer.

5.5.3 The Fair Use Doctrine

The fair use doctrine establishes a bridge between the protection of rights of artists, authors, inventors, or discoverers to benefit from their works and the basic rights of the community to gain from each member's contributions for the betterment of all and the upholding of the principle of economic competition. The use of copyrighted material is considered fair if it does not exploit the commercial value of the work. There are four ways to judge whether the use of an invention, discovery, or work is fair. We list them here and discuss them in depth in Chap. 12:

1. The purpose of use, whether commercial or educational
2. Nature of use
3. Percentage of use
4. The effect of use on the commercial value of the invention, discovery, or works

The fair use doctrine has also given rise to conflicts between the separation of free speech and copyrights. According to Strong [6], a “citizen may be free to speak, but he is not entitled to speak his mind in the same words as his neighbor. He is free to speak the idea if you will, but not the expression.” There are so many exceptions and inclusions under the fair use doctrine that it is difficult to be sure what is fair use unless one talks to a copyright lawyer or other experts

knowledgeable in copyright law. The rule of thumb advocated by many copyright lawyers is that any time you have to copy any part of a copyrighted work outside personal educational use, even in the case of just one copy, talk to somebody who knows the law.

5.6 Protection of Ownership Rights

In Sects. 5.3 and 5.5 we discussed the intellectual property rights instruments of protection and how they can be infringed. In this section we consider how an owner of these property rights can use these instruments. We approach this by discussing the domain, source and types, duration, and the strategies of protection.

5.6.1 Domain of Protection

During our discussion of intellectual property rights, we defined the domain as the set of all different rights enjoyed by the owner of a manifested idea. Within this domain, there are subsets of rights enjoyed by the owner, depending on the manifestation of the idea. These subsets are protected by the body of laws discussed in Sect. 5.3, namely, copyright, patent, trademarks, and trade secret laws. Under each of these subsets, different rights are protectable, as shown here:

1. *Copyrights*: Copyright laws protect all rights embodied within the copyrighted work by the copyright act of the particular country, including the right to use, transform, sale, copy, and modify.
2. *Patents*: Patent laws protect all rights embodied in the particular country's patent law.
3. *Trademarks*: Trademark laws protect all rights in the different trademark statutes depending on the state and country.
4. *Trade secrets*: Trade secret statutes and laws protect all rights within the different states, local authority, and the country's statutes.

Anything else outside these sets, except the various laws that protect personal identity, should be in the public domain and, therefore, is not protectable.

5.6.2 Source and Types of Protection

Because intellectual crimes have become global with the growing technological advances, there has been a realization that there must be protection of national interests. Thus, a number of national and global organizations have been put in place and national acts and international treaties signed to fight these crimes. In

in the United States, intellectual property rights are protected by the copyright and patent laws. Other intellectual property laws in United States include the following:

- The Antipiracy Act of 1976
- The Communication Act of 1984, The No Electronic Theft Act (NET Act)
- The Digital Millennium Copyright Act (DMCA)
- The Economic Espionage Act of 1996
- Money Laundering Act of 1956

These and a number of state statutes or local ordinances protect the IPR. But because neither federal nor state protection is extended outside U.S. borders, different organizations have over the years been set up to protect these rights. Among these:

- The World Trade Organization (WTO)
- Interpol
- The Universal Copyright Convention (UCC)
- Berne Convention
- The Trade-Related Aspects of Intellectual Property Rights (TRIPPS)
- The World Intellectual Property Organization (WIPO)

Remember that although intellectual property rights are protected by a body of laws, the burden of policing, detection, and prosecution in any country is squarely on the shoulders of the owner of the specific intellectual property rights protected.

5.6.3 Duration of Protection

As we saw in Sect. 5.3, the period during which intellectual property is protected depends on a number of factors, including the body of laws protecting your rights and your geographic region.

5.6.4 Strategies of Protection

The burden of safeguarding the intellectual property rights of an individual is with that very person owning the work. It is the duty of individual owners of copyrights, patents, trade secrets, and trademarks to devise strategies to safeguard these rights.

Various methods have been used by individuals and companies who hold these rights to defend themselves. Large companies and individuals have been known to use methods ranging from spying on competitors and suspected infringers using private undercover operatives to collaborating with government officials to check on imports and exports. Some companies call in their respective governments when they suspect foreign infringements. When governments step in, they negotiate joint policing within the respective countries and sign treaties to protect these rights. For example, the U.S. government has negotiated with and sometimes pressured

foreign governments on behalf of U.S.-based companies to observe the intellectual property rights of U.S. technology companies after these companies suspected infringement by individuals and companies in these countries.

Within the United States, some corporations, especially software companies and computer chip manufacturers, have started using local law enforcement agencies to raid suspected infringers both in the United States and in other countries. Another approach used by private companies is a blitz of publicity about suspected countries and counterfeit products. Education campaigns via the mass media are also being used as a viable and effective protection strategy.

5.7 Protecting Computer Software Under the IP

We know that algorithms and ideas are not classified as intellectual property and, therefore, are not protected in any way. Ideas and algorithms belong to everybody, and no one can put a claim on them. Software, although it comes from and strictly follows an algorithm, is not considered an algorithm but rather a manifestation, an expression of that algorithm. Many people may have many different ways of expressing, and therefore representing, the same algorithm and hence have different programs that are considered clear intellectual property and are, therefore, protectable. But for computer software, there are no guidelines one can use to claim that because software is considered a derivation of an algorithm, it is, therefore, protectable. Computer products, in particular computer software, are more elusive and thus have been presenting many problems for those seeking protection under the intellectual property rights law. The difficulty with software protection comes from the difficulty in categorizing it. As we said earlier, software can be a product, a service, or a mixture of both.

5.7.1 Software Piracy

Discussing the intellectual property rights (IPR) one cannot fail to think about the modern wonder of technology, the computer software, and its relationship with IPR. The biggest problem concerning computer software and IPR is software piracy. Generally speaking, we can define software piracy as the act of copying, distributing, or using proprietary software. This act is and has been illegal ever since software began to be protected by law after software manufacturers started filing for patents and copyrights for their products and creations. However, this has not always been the case. In the early days computers, mainly mainframe, came with software preloaded. There were few computers and few users, nobody cared about the software, the least understood component, let alone knowing how to use it and when to use it. With the miniaturization and widespread use of computers together with the high costs of production and purchase costs of software, this changed. A demand for software was created that lead to the piracy problems. We

come around to this issue when we discuss the transnational software issues in the next section.

The issue of software piracy is a complex one. Several other issues complicate software piracy. Some people use illegal software without knowing that the copies they have are illegal. Others use it with the full knowledge that the copies they are using are illegal but they go ahead anyway. Others are confused by the software terminology that includes freeware, shareware, and commercial software. Yet others, especially those in educational institutions, are confused by the IPR principle of fair use. They cannot tell how much is fair. There is also a large percentage of illegal software users who do it purposely to ‘get even’ with software manufacturers that frequently upgrade software versions making older versions of the product obsolete. We discuss these and other issues in the coming section.

Is there a solution to this problem? Yes and no. Yes, in that software companies and governments are working together in efforts to eliminate or downgrade the problem.

5.7.2 Protection of Software Under Copyright Laws

Computer software, along with its documentation, can be protected under the copyright laws. According to Section 101 of the 1980 U.S. Copyright Amendment, a computer program is defined as “a set of statements or instructions to be used directly or indirectly in a computer in order to bring about a certain result” [10]. This statement automatically implies that such a set of instructions or statements is a written creative work by someone, hence a literary work. Therefore, copyright laws that protect creative works also protect computer programs, including technical manuals and documentation. The developer of a program thus has protective rights to authorize others of his or her choice to reproduce the format of the program en masse at any time, to update and upgrade the program, and distribute it for free, or to sell or transfer ownership. The copyright laws in this case protect the source code, the object code, and the executable codes, including the manuals and documentation, from illegal copying and piracy.

Although the developer has such rights, there are limitations or exceptions in some instances. For example, a buyer of such a program has some protected rights also. The buyer has the right to make another copy as a backup, provided that a copy is not used on another machine. Software registration for copyright, however, does not stop a software developer’s worries because by registering software, the developer is opening up the secrets of the whole project as part of the requirements of the copyright issuance. According to Neitzke, there are two roadblocks in copyright registration in some countries:

1. Some courts have taken the position that copyright registration precludes maintaining software as a trade secret.

2. In some countries registration requires submitting at least the first and last 25 pages of the program for public inspection.

So before a developer goes ahead with a copyright application he or she should weigh the pros and cons of copyright protection.

5.7.3 Protection of Software Under Patent Laws

In Sect. 5.3.2 we defined a patent as the protection of the manifestation of an idea on condition that the patent owner discloses the methodology for the manifestation and workings of the product from the idea. In contrast to the copyright laws, however, patents protect the processing of the idea; they also protect the implementation of the idea as a machine, a manufacturing process, or a composition of matter or materials. Under these conditions, computer hardware components, by their very nature, are protected under the patent laws. Software also may be protected under the patent laws under certain circumstances. Under these conditions, how can software be protected?

This is a difficult question, and we must answer it first by explaining that patent issues for computer programs are not yet settled. There are various reasons for the debate, among which are the following:

1. The requirement of the patent system for total disclosure of all information pertaining to the application of the patent is still a big issue. Given that the patent protection lasts for 17 years and 2-year-old software is as old as software can get and still be really viable, requiring developers to disclose the secrets of their software before the 17-year deadline opens them up to stiff competition and better variations of the applicant's software format.
2. Most of the computer programs on the market are simple one-person ventures. Such persons, many of them independent-minded individuals, may not support yet let alone be able to afford the expense of the patent process application.
3. It has been and still is very difficult, as we saw earlier in this section, to prove to courts and patent offices that algorithms are processes and therefore a form of manifestation of an idea and not mere mental gymnastics that any human being can do and, therefore, not patentable because mental steps and mathematical formulas are not patentable items.

Although computer programs are not suited for patent protection, there have been successful applications that have received patents. For specific examples of some of these cases, see Gervaise Davis's book *Software Protection* [11].

5.7.4 Protection of Software Under Trademarks

In Sect. 5.3.4 we defined a trademark as a symbol or mark with financial value that helps customers connect to the product. The main purpose of a trademark is to make the product stand out against the competitors. All hardware companies and a few software concerns such as Microsoft have their trademark protected under the trademark laws and statutes. But how is a mark or symbol be used to protect computer programs?

The protection of computer programs by trademarks is achieved through self-realization by the infringer that it is not easy to copy, change, or redistribute copies of well-known software works. For example, it is not easy to make copies of Windows 9X, NT, or any other Windows product, and resell them, although there have been instances of this. So for big-name software developers, this realization by would-be infringers works far better than law enforcement. But the trick does not work all the time, especially in countries where this sort of realization does not have as much appeal because of lack of publicity of the products. Apart from these measures, software developers do include their symbols and marks within the product so that during use, the symbols and marks are displayed automatically. However, there are no effective global trademark laws to protect computer programs.

5.7.5 Protection of Software Under Trade Secrets

So far we have defined a trade secret as information one has about a manifestation of an idea and that no one should disclose or use for the benefit of themselves or a competitor. As pointed out in Sect. 5.3.3, there are basic laws to protect trade secrets. How do these laws help protect computer products, especially software?

The manifestation of an idea into a computer program usually starts with the blueprint and flowchart, as we saw earlier. This process then goes through the remaining stages of object code and executable code. One's knowledge of the process anywhere during these stages forms a trade secret and should not be revealed for personal gain or to a competitor.

It is generally known to computer programmers and software developers, as it is known to all hardware engineers, that once the blueprint and flowchart of a computer program are known, it is easy to develop a program. It is, therefore, of the utmost importance that at the early software development stages, the blueprint and flowchart not be known outside design circles. Typically the trade secret laws require an infringer, if caught, to stop, return the material under dispute to the rightful owner, and pay damages. But there are difficult cases, such as when former employees leave their employers without written material but with years of acquired knowhow of product development. Here the law is difficult to apply. Some companies make the employees sign nondisclosure contracts for a specific number of years after leaving the company. This method works to some extent, but it is very difficult to enforce except in high-profile and rich companies.

5.8 Transnational Issues and Intellectual Property

A number of studies concerning the international IP system show that there is an extensive and growing number of losses being incurred by businesses in the developed world as a result of nonenforcement of IP laws in the developing world. The developed world is charging that the lack of IP legislation in some development countries and the absence of enforcement in others are amounting to sanctioning pirates and leading to losses amounting to tens of billions of dollars' worth of goods of multinational corporations every year. Developing countries, however, are not amused with the charges lobbed on them. They argue that:

- The IP system, if instated in full in their countries, results in significant social costs on that country; this may include developing the cost of acquiring and maintaining the IP rights and defending those rights whenever there are international legal disputes.
- Country memberships in the present IP system are costly and exacerbate the costs of enforcement.
- Loosely enforcing the IP laws will speed their industrialization and development by enabling them to copy state-of-the-art technologies.
- The IP protection is not as profitable as was touted by developed countries. There is evidence that the innovation for developing countries is not visible. For example, the introduction or strengthening of patent protection for pharmaceutical products has not increased national or foreign direct investment, production, or R&D in developing countries. On the other hand, the Indian pharmaceutical industry became a global producer of active ingredients and medicines in the absence of patents on such products, which was only introduced in January 2005, at the expiry of the transitional period allowed by the TRIPS Agreement [1].
- The industrialized world, when in the process of development, did not depend on the patent system but rather the lack of the IP system, which promoted innovation.

So, the developing world is reluctant to accept the IP system wholesale without concessions unless the industrialized countries guarantee them greater access to their markets for their goods and agricultural products. There are other issues pertinent to the IP system, but we do not go into those here.

Discussion Issues

1. Do you think the developing world has relevant issues in this discussion?
2. Is the developing world being misled by a few powerful countries within their ranks?
3. What kind of concessions should the developed world make?

Exercises

1. Discuss the problems faced by software developers trying to apply for protection under trade secret statutes.
2. Why is it difficult to apply patent laws to software?
3. Why is it possible to apply patent law to software?
4. Is it possible to trademark software?
5. Discuss the ethical and legal issues surrounding software ownership.
6. There is a move to do away with the current copyright law. Why?
7. Why is the copyright law, in its present form, considered to be unenforceable?
8. What changes would you suggest in the current copyright laws to make it enforceable in cyberspace?
9. Has the Internet made software protection easier or more difficult? Why or why not?
10. There is a movement (that includes hackers) that is advocating for free software! Discuss the merits of this idea, if any.
11. Because of income disparities between north and south, and haves and have-nots, fair pricing of computer products is impossible. Discuss.
12. Most copyright violations are found in developing, usually poor, countries. Why?
13. Does the high price of software marketing in developing countries justify the high rate of software piracy in those countries? Why?
14. What do you think is the cause of the rising cost of software?
15. Is globalization a means through which the developed, usually northern, countries will enforce the copyright laws?

References

1. WIPO—World Intellectual Property Organization, WIPO Arbitration and Mediation Center. *Administrative Panel Decision Hugo Boss Trade Mark Management*. GmbH & Co. KG, Hugo Boss AG v. Irfan Butt Case No. D2016-1123
2. H. Nasheri, *Addressing Global Scope of Intellectual Property Law*. <http://www.ncjrs.gov/pdffiles1/nij/grants/208384.pdf>
3. D.J. Johnson, *Computer Ethics*, 4th edn. (Pearson Education, Inc., Upper Saddle River, NJ, 2009)
4. J. Prince, Negligence: liability for defective software. *Oklahoma Law Review* **33**, 848–855 (1980)
5. G.A. Gow, *Copyright Reform in Canada: Domestic Cultural Policy Objectives and the Challenge of Technological Convergence*. http://www.academia.edu/11974855/Copyright_Reform_in_Canada_Domestic_Cultural_Policy_Objectives_and_the_Challenge_of_Technological_Convergence
6. W.E. Strong, *The Copyright Book: A Practical Guide*, 6th edn. (MIT Press, Boston, 2014)
7. F.W. Neitzke, *A Software Primer* (Van Nostrand Reinhold, New York, 1984)
8. Personal Identity Theft on the Rise. *USA Today*, Tech Report. 09/14/00
9. D.A. Burge, *Patent and Trademarks: Tactics and Practice*, 2nd edn. (Wiley, New York, 1984)
10. M.D. Scott, *Computer Law* (Wiley, New York, 1984)
11. G.G. Davis, *Software Protection* (Van Nostrand Reinhold, New York, 1985)

Further Reading

12. R. Davis, A new view of intellectual property and software. *Commun. ACM* **39**(3), 21–30 (1992)
13. E. Oz, Protecting software as intellectual property, in *Ethics for the Information Age* (Business and Education Technologies, Barr Ridge, 1994), pp. 273–285
14. P. Samuelson, Information and property. *Cathol. Rev.* **38**, 365–410 (1989)
15. P. Samuelson, Is information property? *Commun. ACM* **34**(10), 15–18 (1991)
16. P. Samuelson, Copyright law and electronic compilations of data. *Commun. ACM* **35**(2), 27–32 (1992)
17. P. Samuelson, Regulation of technologies to protect copyrighted works. *Commun. ACM* **39**(7), 17–22 (1992)
18. J. Suapper, Intellectual property protection for computer software, in *Computer Ethics and Social Values*, ed. by D. Johnson, H. Nissenbaum (Prentice Hall, Englewood Cliffs, 1995), pp. 181–190



Social Context of Computing

6

Abstract

This chapter considers social issues in computing including the digital divide, workplace issues like employee monitoring, health risks due to computer use, and how these issues are changing with the changing computer technology. The chapter also covers a detailed discussion on a number of obstacles to overcoming the digital divide through digital inclusion within countries and globally. On workplace issues, the discussion focuses on the best practices to deal with the changing workplace issues resulting from the growing army of home-based workers and measuring employee productivity.

Learning Objectives

After reading this chapter, the reader should be able to

1. Interpret the social context of a particular software/hardware implementation.
2. Identify assumptions and values embedded in a particular computer product design, including those of a cultural nature.
3. Evaluate a particular computing tool implementation through the use of empirical data.
4. Describe positive and negative ways in which computing alters the modes of interaction between people.
5. Explain why computing/network access is restricted in some countries.
6. Learn the impact of the digital divide.
7. Understand how income, geography, race, and culture influence access to information technology and technology in general.
8. Analyze the role and risks of computing in the implementation of public policy and government.

9. Articulate the impact of the input deficit from diverse populations in the computing profession.

Scenario 5: Electronic Surveillance and the Bodyguard

Jon Kiggwe is a young aggressive entrepreneur, with a bright future. With several businesses doing well and a few start-ups with promising financial status, Jon is on his way to making a million dollars before his 25th birthday. Jon's business meetings take him into tough neighborhoods. So, that he may feel secure, Jon uses a team of professional security bodyguards to shadow him almost 24 h a day.

In his big 10-million-dollar home, Jon receives a stream of guests, including both business associates and friends. His bodyguards, besides keeping an eye on him, also see to the orderly arrival and departure of the guests. Because of this, the bodyguards keep a permanent office and sleeping quarters at Jon's mansion.

Without informing them, Jon installed video recording and listening gadgets in the guards' office and sleeping quarters to record their every conversation and movement. He feels safe that way!

Discussion Questions

1. *Is Jon violating any law?*
2. *Do the bodyguards have any right to privacy on Jon's premises?*
3. *Does Jon have a right to know what the bodyguards are doing in his house?*

6.1 Introduction

In the past 5 years or so, we have witnessed an invasion of computers and computer-related equipment in workplaces, homes, and schools. The advent of the Internet, wireless communication, and mobile computer technology has considerably expanded this invasion into planes, trains, and automobiles. The widespread use of computers and computer technology in its present form has also resulted in a shift in computer usage. The computer started as a utilitarian tool but has now also been embraced as a social tool. Probably because of the popularity of the Internet, both young and old have found solace in computing devices everywhere. Playing this double role as a utility and an entertainment tool, the computer has become an integral part of our social fabric.

However, in the meantime, two worlds have been created for humanity: the unreal world of entertainment and a real computer technology-driven world, which augments our familiar environment and makes our daily activities easier and more enjoyable. This development in turn has led to an influx of computer technology into the workplace, schools, and the home. Indeed, the home has turned into a

hub of technology. No one knows, as yet, the social, psychological, and intellectual implications that may result from this. Predictions abound that this will enhance our intelligence and improve our performance at whatever we do. This belief alone has been a driving force for the computerization of schools and homes, with parents hoping to produce young geniuses.

These beliefs about the value of technology, whether or not supported by scientific research, are not new. Ever since the beginning of the industrial age, when technology started entering the workplace and homes, the aim has been to utilize it to help us be wiser and more productive. It is, therefore, no wonder that as technology has developed, progress and fundamental changes have been taking place almost daily. Our focus in this chapter is on both the social and ethical effects of computer technology on people, whether we are at home, school, or work. We focus on the social and economic dimensions of computing as a result of the “digital divide,” the workplace, workplace monitoring of employees, and the well-being of employees.

6.2 The Digital Divide

The technological inequalities among people in one country and between countries, commonly known as the digital divide, arose from the landmark 1994 U.S. Commerce Department’s National Telecommunications and Information Administration (NTIA) report, “Falling Through the Net,” commonly referred to as NTIA I. The NTIA I report used the Information and Communication Technologies (ICT) *access* indicator, one of the many digital divide indicators, to highlight sectors of the U.S. population that were technologically deprived. Since then, the digital divide debate has been raging, centered on a number of key critical issues including the following:

- Whether there is such a thing as a digital divide
- Indicators that should be used to measure such a divide if it exists, and
- The best ways to close such a divide

Much of the debate is the result of a lack of understanding about the digital divide—its origins, inputs, and responses to inputs. In general, in a broader sense, the study of the digital divide involves the study of the impact of the digital divide indicators. These indicators concern communication technologies such as radio, television, the press, fixed and cellular telephones, fax machines, computers, and connectivity to the Internet and participation in cyber activities for all members of a society. However, in its most basic definition, it is a discrepancy in access to information technology. What causes it? Why does it exist? Answers to these two questions can fill as many as two large books. There is a multitude of causes and enablers, and so long as these exist in any society, the digital divide will exist. Study after study, since the inception of the concept, have pointed to *social, economic, and geographic* factors as influencing the digital divide. More specifically,

the following are enablers of the digital divide: *access, relevant technology, humanware (human capacity), infrastructure, and enabling environment*. These enablers fuel the following causes of the digital divide: *geography, age, education, income, race, and ethnicity*.

6.2.1 Access

Access is a crucial component in the digital divide: it involves obstacles that exist even if all the other remaining indicators are in place. Such obstacles may include, but are not limited to, costs involved in acquiring the technologies, availability of free or low-cost facilities in the neighborhood, the ability to travel to places where there are low-cost access points, such as libraries and community centers, and having the capacity needed to utilize the technologies. These obstacles can broadly be grouped into five categories: geography, income, ethnicity, age, and education.

6.2.1.1 Geography

According to the UN Human Development Report of 2011, there is a large digital divide between the rich, industrialized countries of the Northern Hemisphere and the poor, less industrialized countries in the Southern Hemisphere. The poor developing countries, geographically in the Southern Hemisphere and mostly in the southern axis of development, are more deprived of access to information, although mobile technology has improved this situation greatly in the last few years.

ITU World Telecommunications/ICT databases (WTI) and UNDP for years 2000–2008 show us the digital divide that exists between countries. For example, in the highest ranked 30 or so countries of the HDI (the very high group), Internet users represent an average of 61.4% of the population, whereas they represent an average of 1.8% for the 20 or so lowest ranked countries classified as low human development (ITU 2009; UNDP 2008) [1].

Focusing on information communication technology (ICT), the main driver among the indicators of the digital divide, the picture, although improving some, remains the same in mobile cellular, mobile broadband, fixed broadband, and Internet technology (Fig. 6.1).

According to Notari [2], the status of global digital inclusion leaves much to be desired. For example, of the approximately 7 billion inhabitants of the earth (2011 estimates):

- 65% are not digitally connected.
- 69% of people in the developed countries have access to the Internet.
- 21% of the people in developing countries have access to the Internet.

The divide is not only between the Northern and Southern Hemisphere nations, it also exists within individual nations. For example, within the U.S., Kruger and

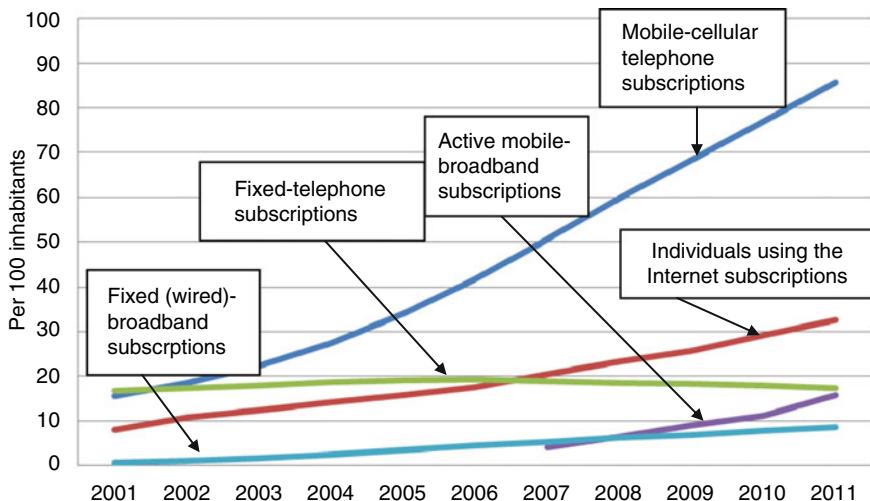


Fig. 6.1 World Information and Communication Technologies (ICT) Indicators. *Source* World Telecommunication/ICT Indicators Database, <http://www.itu.int/ITU-D/ict/statistics/>

Gilroy [1] report that although the number of new broadband subscribers continues to grow, the rate of broadband deployment in urban areas appears to be outpacing deployment in rural areas: 13 recent surveys and studies have indicated that, in general, rural areas tend to lag behind urban and suburban areas in broadband deployment. Consider the following surveys [1]:

- The Department of Commerce's "Exploring the Digital Nation" report found that although the digital divide between urban and rural areas has lessened since 2007, it still persists with 70% of urban households adopting broadband service in 2010, compared to 57% of rural households.
- Data from the Pew Internet & American Life Project show that the percentage of all U.S. adults with broadband at home is 70% for non-rural areas and 50% for rural areas.
- Data from the National Broadband Map (2011) indicate that 99.7% of the population in urban areas have access to available broadband speeds of at least 3 Mbps (download)/768 kbps (upload), as opposed to 84.0% of the population in rural areas.

However, this north–south technological divide is constantly changing for the better. Data from Table 6.1 show that although there is still a substantial rift between the Northern and Southern Hemispheres, it is rapid narrowing in at least in these four technologies.

Table 6.1 Key statistical highlights: ITU data release June 2012*Mobile cellular:*

Total mobile-cellular subscriptions reached almost 6 billion by end 2011, corresponding to a global penetration of 86%

Growth was driven by developing countries, which accounted for more than 80% of the 660 million new mobile-cellular subscriptions added in 2011

In 2011, 142 million mobile-cellular subscriptions were added in India, twice as many as in the whole of Africa, and more than in the Arab States, CIS, and Europe together

By end 2011, there were 105 countries with more mobile-cellular subscriptions than inhabitants, including African countries such as Botswana, Gabon, Namibia, Seychelles, and South Africa

Countries where mobile-cellular penetration increased the most in 2011 include Brazil, Costa Rica, Kazakhstan, Lao P.D.R., and Mali

Mobile broadband:

By end 2011, there were more than 1 billion mobile-broadband subscriptions worldwide

Mobile broadband has become the single most dynamic ICT service, reaching a 40% annual subscription growth in 2011

Although developing countries are catching up in terms of 3G coverage, huge disparities remain between mobile-broadband penetration in the developing (8%) and the developed world (51%)

In Africa there are fewer than five mobile-broadband subscriptions per 100 inhabitants, whereas all other regions have penetration levels above 10%

By end 2011, there were more mobile-broadband subscriptions than inhabitants in the Republic of Korea and Singapore. In Japan and Sweden, active mobile-broadband penetrations surpassed 90% by end 2011

In 2011, 144 million mobile-broadband subscriptions were added in the BRICS (Brazil, the Russian Federation, India, China and South Africa), accounting for 45% of the world's total subscriptions added in 2011

Fixed (wired) broadband:

By end 2011, there were 590 million fixed (wired)-broadband subscriptions worldwide

Fixed (wired) broadband growth in developed countries is slowing (5% increase in 2011), whereas developing countries continue to experience high growth (18% in 2011)

Fixed (wired)-broadband penetration remains low in some regions, such as Africa and the Arab States, with 0.2% and 2%, respectively, by end 2011

In 2011, 30 million fixed (wired)-broadband subscriptions were added in China, about half of the total subscriptions added worldwide, and fixed (wired)-broadband penetration reached 12% in the country

Top performers, such as France, Denmark, the Netherlands, Norway, the Republic of Korea, and Switzerland, had fixed (wired)-broadband penetrations above 35% by end 2011

(continued)

Table 6.1 (continued)

Countries where fixed (wired)-broadband penetration increased the most in 2011 include Bahrain, Costa Rica, Ecuador, Mauritius, and Uruguay. However, among these, only Bahrain and Uruguay surpassed the 10% fixed (wired)-broadband penetration by end 2011

Internet:

The percentage of individuals using the Internet continues to grow worldwide and by end 2011 2.3 billion people were online

In developing countries, the number of Internet users doubled between 2007 and 2011, but only a quarter of inhabitants in the developing world were online by end 2011

The percentage of individuals using the Internet in the developed world reached the 70% landmark by end 2011

In Iceland, the Netherlands, Norway, and Sweden, more than 90% of the population are online

By end 2011, 70% of the total households in developed countries had Internet, whereas only 20% of households in developing countries had Internet access. Some outstanding exceptions include Lebanon and Malaysia with 62% and 61% of households with Internet, respectively

Source ITU World Telecommunication/ICT Indicators Database. http://www.itu.int/ITU-D/ict/statistics/material/pdf/2011%20Statistical%20highlights_June_2012.pdf

6.2.1.2 Income

According to the most recent Pew Research Center study of the U.S. population, household income is the greatest predictor of the use of the Internet and other ICT technologies. Households earning more than \$75,000 a year significantly outpace lower-earning households, particularly those making less than \$30,000 a year [3]. In his Pew Research Center Report, “The Better-Off Online,” Jansen [4] reports that the analysis of several surveys conducted by the Pew Research Center’s Internet & American Life Projects revealed significant key differences between those who live in households making \$75,000 or more relative to those in lower-income households. The key findings in Jansen’s report in three information communication technologies, namely, broadband at home, Internet use, and mobile cell phone ownership are as follows [4]:

- Broadband at home:

< \$30,000	40%
\$30,000 to < \$50,000	79%
\$50,000 to < \$75,000	79%
> \$75,000	87%

- Regularly use the Internet

< \$30,000	57%
\$30,000 to < \$50,000	80%
\$50,000 to < \$75,000	86%
> \$75,000	95%

- Cell phone ownership

< \$30,000	75%
\$30,000 to < \$50,000	90%
\$50,000 to < \$75,000	93%
> \$75,000	95%

6.2.1.3 Ethnicity

According to NTIA 2000 [5], one's ethnicity has a great influence on ICT access. For example, in the United States, blacks and Hispanics, the two main U.S. minority groups, are twice as likely as their white counterparts not to have a computer and access to the Internet.

Although there has been no comprehensive study of global ICT access based on ethnicity and race, there have been limited but revealing national studies. Nearly all these national studies point to results similar to those in the original NTIA 2000 report.

Since the NTIA 2000 report [5], however, there have been interesting changes in the issue of ethnicity and access to ICT technologies. These dramatic changes have been brought about by the rapid changes in modern communication technologies, more specifically Internet-able mobile communication technologies. Jesse Washington [6] reports in “For minorities, new ‘digital divide’ seen,” that Latinos and blacks are now more likely than the general population to access the Web by cellular phones. Today, mobile technology has become an equalizer of sorts, in some aspects, as it brings computers and the ability to access the Internet and the Web, at the same rate as whites, to the hands of minorities such as Latinos and blacks. However, because these minorities have limited options to access the Internet and the Web, they are more likely than the general population to access the Web by cellular phones, and they use their phones more often to do more things. Smith [7] reports on a continuing a trend, first identified in 2009, that minority Americans lead the way when it comes to mobile access, especially mobile access using handheld devices. Nearly 64% of African-Americans and 63% of Latinos are wireless Internet users. It is also emerging that minority Americans are significantly more likely to own a cell phone than their white counterparts because 87% of blacks and Hispanics own a cell phone, compared with 80% of whites. Additionally, black and Latino cell phone owners use their cell phones in a wider array of functions compared to white cell phone owners.

Does this mean that the divide is over? Probably not! There is wide agreement that Latinos and blacks are becoming more challenged as they access the new technologies because the computing powers and functionalities of the current mobile technologies are still very limited. Thus, overreliance on them is creating a sort of new “digital divide” for these groups. For example, it is tough to fill out a job application on a cell phone. Also, blacks and Latinos are increasingly using their mobile power more for entertainment than for empowerment.

Discussion Topic

In what ways does mobile technology disadvantage minorities?

6.2.1.4 Age

There is a myth that young people use computers and the Internet far more than any other age group. There is also conventional wisdom that young people under age 18 do more surfing of the Internet than any other age group. However, this is not the case. There are consistent data from NTIA 2000 and the UCLA Internet Report showing that the highest usage of computers and the Internet is among people between the ages of 18 and 49 (Table 6.2). NTIA 2000, the UCLA Internet Report (Table 6.3), and the British Office of National Statistics report (Table 6.4) are consistently showing that older people and those under 10 years of age use computers and online access far less than any other age group. Also, the latest figures from Media Metrix show similar global patterns with a bell-shaped curve pattern of usage peaking between 18 and 54 years [8].

More than 12 years after the NTIA 2000 report, advances in technology have changed the digital inclusion landscape. Now, instead of talking about the use of computers, cell phones, and Internet access, it is more about mobile phones and

Table 6.2 U.S. households with computer and online access by age of inhabitants

Category (years)	%
Under 8	15.3
9–17	53.4
18–24	56.5
25–49	55.4
50 +	29.8

Table 6.3 Age and internet usage (UCLA report)

Age group (years)	12–15	16–18	19–24	25–35	36–45	46–55	56–65	65 +
Average hours per week	5.6	7.6	9.7	11.3	9.4	10.3	8.5	6.8

Table 6.4 Internet users and non-users, UK, 2011, Q1–Q4

	Used internet				Never used internet			
	2011 Q1	2011 Q2	2011 Q3	2011 Q4	2011 Q1	2011 Q2	2011 Q3	2011 Q4
All	82.2	82.3	82.9	83.5	17.5	17.4	16.8	16.3
<i>Age (years)</i>								
16–24	98.8	98.8	98.6	98.7	0.9	0.9	1.1	1.0
25–34	97.5	97.7	97.8	98.0	2.1	2.1	2.0	1.8
35–44	95.4	95.4	95.6	95.9	4.3	4.3	4.1	3.9
45–54	89.5	89.9	90.2	90.5	10.2	9.8	9.5	9.5
55–64	79.0	79.2	79.9	81.1	20.8	20.6	19.8	18.7
65–74	57.1	57.6	58.7	59.8	42.6	42.1	41.2	40.0
75 +	23.8	23.6	27.3	29.0	76.1	76.3	72.4	70.8

Source British Office of National Statistics, <http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-250549>

wireless access. So, the discussion now focuses on use of the Internet and mobile devices to access the Internet.

Although we do not have comprehensive data for global digital inclusion based on age, we can discuss data from the U.S. and Britain. According to Ian Clark [9], in the United Kingdom, the 2011 statistics from the Office for National Statistics (ONS) of Internet use by age reveal the highest percent of use for the 16- to 24-year-old group than any other age group. This figure then levels off as age increases. It is worrying to see that 71% of those more than 75 years of age and 40% of the 65- to 74-year-old age group have never used the Internet (see Table 6.3). There are a number of reasons why this is the case.

The data do not change very much when it comes to the U.S. Some 92% of Americans aged 18–29 are online, according to the Pew Internet and American Life Project [7]. Again, in a similar fashion, the rate falls as the ages of users increase, showing 87% for those aged 30–49 years and 79% in the age range of 50–64 years, down to a low of 42% for those over 65 years of age.

On wireless communication, the picture becomes more interesting. Smith [7] reports that nine in ten persons of 18–29 years old own a cell phone, and these young cell phone owners are significantly more likely than those in other age groups to engage in all the mobile data applications, as follows:

- 95% send or receive *text messages*
- 93% use their phone to take pictures
- 81% send photos or videos to others
- 65% access the Internet on their mobile device
- 64% play music on their phones
- 60% use their phones to play games or record a video
- 52% have used their phone to send or receive e-mail
- 48% have accessed a social networking site on their phone

- 46% use instant messaging on their mobile device
- 40% have watched a video on their phone
- 33% have posted a photo or video online from their phone
- 21% have used a status update service such as Twitter from their phone
- 20% have purchased something using their mobile phone
- 19% have made a *charitable donation* by text message

There is growing evidence that this love for mobile devices is also growing fast among those 30–49 years old.

6.2.1.5 Education

Ever since the NTIA I report showed that the higher the education level one achieves, the more likely one is to use a computer and, therefore, the Internet, study after study have shown the same thing. Data from NTIA 2000 and the UCLA Internet Report (Tables 6.5 and 6.6) show the same trend. For example, the very highly educated with advanced degrees, reported in both NTIA 2000 and the latest UCLA Internet Report, show 69.9% and 86% Internet usage, respectively, compared to 11.7% and 31% usage, respectively, for those with less than a high school diploma.

As we observed earlier, more than 12 years since the NTIA 2000 report, the rapid advances in technology have changed the digital inclusion landscape. When we talk about digital inclusion, understanding has shifted from using computers, cell phones, and Internet access, to having an Internet-able mobile device. Based on this thinking, in the past 12 years since the NTIA 2000 study, the situation has changed considerably but has remained the same in that digital inclusion still

Table 6.5 U.S. households with computer and online access by education level

Category	Computer (%)	Online (%)
Elementary	18.2	11.7
High school diploma	39.6	29.9
Some college	60.3	49.0
College diploma	74.0	64.0
Postgraduate	79.0	69.9

Table 6.6 Internet use and level of education (UCLA report)

Education level attained	Less than high school	High school graduate	Some college	College graduate	Advanced degree
Percent using Internet	31.2	53.1	70.2	86.3	86.3

favors high education. Look at the data from the Pew report 2010 on the U.S. population [7]:

- Of all people with less than a high school education, 38% have access to a wireless Internet-able mobile device.
- Of all people with a high school diploma, 48% have access to a wireless Internet-able mobile device.
- Of all people with some college education, 68% have access to a wireless Internet-able mobile device.
- Of all people who are college graduates, 76% have access to a wireless Internet-able mobile device.

6.2.2 Technology

The computer-driven technological revolution has brought the countries of the world closer together. In their study of the digital divide, Rodriguez and Wilson observed that all developing countries, including the poorest, are improving their access to the use of ICT]. In fact, technological progress in developing countries between the 1990s and 2000s has been very strong, outpacing that in developed countries by 40–60%, according to data from the World Bank report, “Global Economic Prospects 2008: Technology Diffusion in the Developing World.” The percentage change in technological achievements between the 1900s and the 2000s is given here [10]:

High income	≈75% change
Upper middle income	≈110% change
Lower middle income	≈102% change
Low income	≈160% change

But the gap between rich and poor countries is still very wide [11]. As Figs. 6.2 and 6.3 show, there is still a large, persistent gap between the industrialized north on one hand and the predominantly developing south on the other. This state of affairs is the result of a lack of broad-based technological skills and know-how. The acquisition of technological skills and, therefore, the development of a good technological base, depends a great deal on relevant inputs that include investment capital, infrastructure, and humanware (human capacity). However, the situation with technology input and output is no better. New technological innovations require huge amounts of money to be invested in research and development. Unfortunately, not enough capital investment is done in developing countries. According to the UN Human Development Report 1999, although developed countries have 21% of the \$(US) 21,000 billion GDP in 1999 invested, the least developed

countries had 20% of the \$(US) 143 billion GDP invested [12]. Because capital investment in technology is usually in form of hardware and software, let us focus on those here.

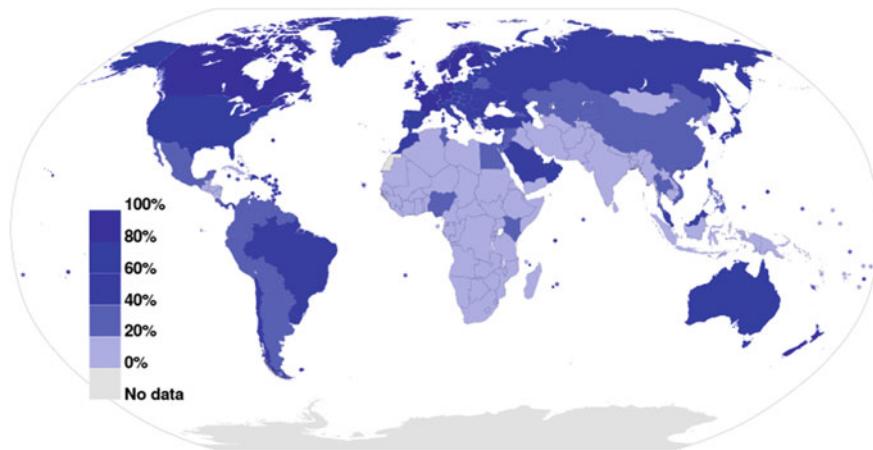


Fig. 6.2 Global technological divide [2]. Source http://en.wikipedia.org/wiki/Global_digital_divide

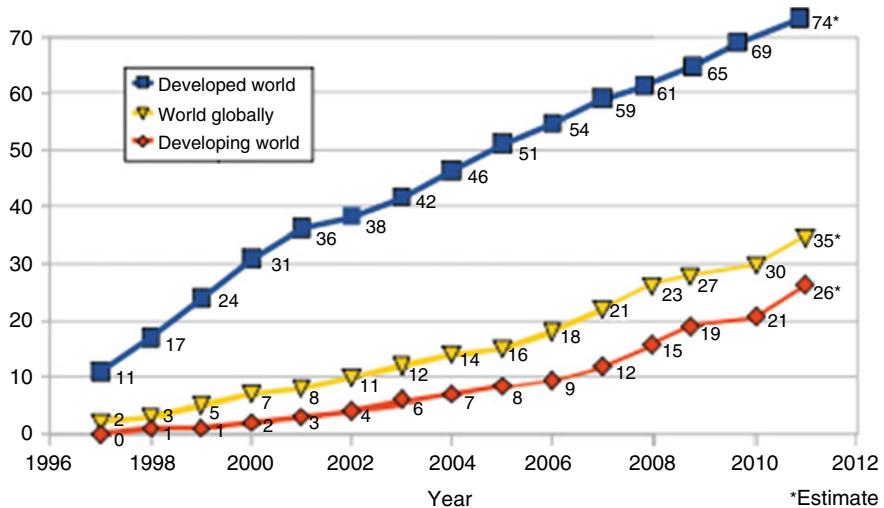


Fig. 6.3 Internet users per 100 inhabitants [2]. Source http://en.wikipedia.org/wiki/Global_digital_divide

6.2.2.1 Hardware

Although there has been a steady increase in the number of computers, telephones, and other modern communication technologies in almost all countries of the world in the last couple of years, as noted by Rodriguez and Wilson, the quantity, quality, and maintenance of these technologies is still a big problem that is challenging the narrowing of the ICT digital divide. There is a serious regression in hardware acquisition and maintenance. Computer components for example, are being acquired, but they are being disposed of at probably the same rate as they are acquired. Uncertain and unreliable power supplies contribute to the shorter lifespans of ICT products in developing countries; probably many of the unusable ICT products are so because they were hit by a power surge.

Replacement of the bad parts is hampered by the price of new ones. ICT parts are very expensive in a number of developing countries because governments either levy high tariffs on imports to raise local revenue or impose luxury taxes because these are classified as luxury items.

ICT products are also expensive because most outlet owners are not indigenous people; they are foreign investors who usually raise prices to cover their local and infrastructure expenses plus profits. In addition, similar to all equipment and software produced in developed countries and imported into developing countries, by the time such items arrive in the developing world, their prices have been inflated three to four times their original value.

6.2.2.2 Software

The problems presented by hardware are accompanied by the problems of software. For ICT equipment to be helpful, it must have good and relevant software. Countries that have seen benefits from ICT, such as those in OECD, either produce their own software or have enough financial capacity to source software with few problems. This is not the case in many developing countries. There is very limited humanware to have software locally produced. In addition, they do not have enough money to source the software from their northern counterparts where it is produced. Even if they have some money to buy software from developed countries, as we pointed out earlier, by the time software arrives in developing countries its list price is much higher. The end result, at least for the time being, is that most ICT software in developing countries comes with the bulk of the donated ICT equipment. However, the software shipped on donated company computers rarely meets the needs of the recipients. Consequently, the local people end up using a product that in most instances produces outputs that have very little value to them, irrespective of the inputs. In such a situation, that equipment ends up not benefiting the locals.

6.2.3 Humanware (Human Capacity)

In this section we consider the complex issues related to human capacity development. Availability and easy access to ICT does not always solve the digital

divide problem. As Rodriguez and Wilson pointed out, it is only a partial solution to a complex problem. Even if we were to provide everyone on the globe with first-class ICT equipment, the reality would remain that only a few would be able to maintain and gainfully use the provided technology. This situation is likely to persist until there is a corresponding degree of technical capacity and knowledge acquired by the people intending to use the technologies so that they can maintain the equipment and derive value-laden outputs. The first problem is lack of humanware in developing countries to maintain the equipment. There is a shortage of teachers, technicians, and institutes to train them. The next challenge is to ensure that people can gainfully use ICT to add value to local inputs. People will take ICT seriously when it meets and serves their own local needs. Human capacity development is a complex multifaceted endeavor consisting of many aspects, including:

- Creating awareness of the potential for ICT to meet one's needs
- Creating, developing, and strengthening capacity to use information and ICT effectively, using local inputs
- Building capacity to produce and package information so that it adds value to local inputs
- Ensuring ongoing technical capacity development and developing a format for sharing knowledge and information
- Preventing the local capacity from being drained to other, usually developed countries

The challenge, therefore, in tackling human capacity development is to take care of each of these issues so that the local persons using ICT may find useful answers to their local problems. ICT capacity development should take into account equity, fairness, and cultural and other contextual factors at the local levels.

6.2.4 Infrastructure

As noted by many, the digital divide infrastructure is related to access in many ways: both present obstacles to gaining access to ICT. For us, infrastructure will mean fixed communication structures. In those countries with good fixed communication structures such as electricity, telephones, good roads, and airports, ICT development is much faster. Lack of such resources hinders the development of ICT.

The availability of these resources helps to speed up the development of ICT structures such as Internet cafes. ICT access enablers such as personal computers, personal assistants, Internet-enabled cellular phones, and other miniature Internet-enabled gizmos in developed countries and the urban areas of developing countries, together with civic centers in developed countries and telecenters in developing countries, have all been hailed in advancing global communication. For these to work, however, there must be a basic communication infrastructure in place. So if

digital communication is to be developed in the developing world, ICT-accessible points such as telecenters, civic centers, and Internet or cyber cafes must be opened up where there are none and expanded where there are only a few.

6.2.5 Enabling Environments

As Rodriguez and Wilson [13] noted, many countries with similar levels of per capita incomes and economic structures exhibit widely varying ICT performances. There are no good explanations for this except for the existence, or lack thereof, of enabling environments. An ICT-enabling environment is an environment in which ICT can thrive. Several things can bring about such an environment, including politics, public policy, and management styles.

6.2.5.1 Politics

According to Rodriquez and Wilson, ICT thrives in a good political environment that ensures:

- A climate of democratic rights and civil liberties conducive to ICT adaptation
- Respect for the rule of law and security of property rights
- Investment in human capacity
- Low levels of government distortions.

One sure way of creating such environments in the different political systems that make up the landscape of the developing world is for the two leading nongovernmental organizations, the G8 Dot Force and the Commonwealth Expert Group on Information Technology, and other development organizations working toward the elimination of the digital divide, to develop a policy that charges governments in individual countries with the task of creating the enabling environments for ICT to thrive. One approach is to develop a Leadership Mobilization Strategy (LMS) to target and educate first the top political leadership at all levels within and outside government about the benefits of ICT to the country's development. These officials must understand the need for ICT and then articulate this need to others: this is crucial to convince leaders to mobilize people and bring ICT awareness to the general population. Although politics is conducive to a good ICT environment, much of ICT development also depends on good public policy and management styles.

6.2.5.2 Public Policy and Management Styles

Governments must put in place streamlined regulatory policies for the importation and licensing of ICT technologies. Laws must be enacted and enforced uniformly so that nongovernmental organizations (NGOs) and other organizations interested in investing in ICT economic activities can do so with ease.

In many developing countries, there are currently ICT-related laws and policies on the books that are not enforced. Such policies must be updated where necessary and enforced strictly and fairly. New competitive policies such as the liberalization of the telecommunication and energy sectors must be developed and the sectors must be staffed with competent managers with appropriate expertise. These ICT regulatory policies need to be efficient, predictable, and easy to understand. Licensing bodies need to be efficient and staffed with professionals. In addition, there must be government support for taxing policies that grant favors such as tax holidays to ICT equipment and investment firms. Finally, there must be transparency in government to create a moral bar for the rest of the country.

6.3 Obstacles to Overcoming the Digital Divide

Based on a number of studies and data, including those of Felix Bankole, Farid Shirazi, and Irwin Brown [14] titled “Investigating the Impact of ICT Investments on Human Development” and that of Kim et al. [15], indicating that digital inclusion is one of the agents of development, countries and policy makers are making every effort to expand the digital inclusion and thus decrease the digital divide within countries and across the global. However, minimizing the digital divide requires considerable effort and a plan in addressing the following types of access [2]:

- Physical Access: individuals need to be able to obtain access to computers, landlines, and networks so they can access the Internet.
- Financial Access: the means to meet the costs of ICT devices, traffic, applications, technician and educator training, software, maintenance, and infrastructures.
- Political Access: the political environment that enables a faster growth of the Internet and other digital inclusion technologies.
- Cultural Access: availability of images and language to carry over the digital inclusion across different cultural lines.

6.4 ICT in the Workplace

The automation of the workplace has been the most vigorously pursued concept since the Industrial Age. Despite the original fear that workplace automation would mean the end to human work, except in a few areas workplace automation has proceeded hand in hand with increases in employment numbers [16]. This is, of course, not to deny that automation has caused some human displacements in the workplace. But overall numbers are steady, and according to the International Labor Office report, the introduction of computers into offices did not bring about any significant dismissal of personnel, nor did it result in a decline in the general

level of employers [16]. Among all the different technologies that have thus far entered the workplace, computer technology has entered at an astonishingly high rate of speed.

6.4.1 The Electronic Office

We can define an electronic office as a technology-augmented office with knowledgeable employees. The technology in the environment may include computers and computer-driven devices that help in interpersonal oral and electronic communication, distribution, and receipt of correspondence; telecommunication devices with text-processing and storage capabilities to enable the office staff to design, develop, edit, and store material electronically; and other office support equipment to streamline decision-making tasks. The evolution of the electronic office began with industrialization but took giant steps beginning in the 1950s with rapid advances in computer technology and telecommunications. Since then the workplace has been undergoing a rapid transformation of its own. Gone are notepads, typewriters, large cabinets filled with manila folders, the rotary telephone, and rotary fans. Computers have replaced most of the filing cabinets, the files, and typewriters. Electronic notepads, automatic answering systems, office intercoms, copiers, and fax machines have moved in. Living plants and air-conditioning have become standard. Increasingly, office job descriptions at all levels and in all professions are being transformed to incorporate computer and telecommunication skills.

Two factors have been and are still fueling the growth of the electronic office. The first is the increasing productivity of office employees, both clerical and professional, to counter the rising costs of office operations, which according to Olson and Lucas [17] have been increasing faster than office employee productivity. The second is the acquiring of technology necessary to handle the ever-increasing complexity and modernization of office communication and decision-making processes.

6.4.2 Office on Wheels and Wings

As electronic gadgetry has been invading the office and the overall workplace, workers have been leaving the office in droves, a few of them replaced by the new technology, others transplanted by it, but many for the experience of working outside the confines of their original office.

The advent of laptop computers, tablets, cellular phones, and personal digital assistants (PDAs) have accelerated the mobility of the office. Busy executives, white-collar workers, and, this time around, blue-collar workers, especially those in the service industry, can be seen in airports, hotel lobbies, restaurants, and aboard airplanes and in trains, keying in data and putting in a day's work as they would have done previously in their offices.

Mail and package service company drivers are also keying in their locales and speed, transmitting the data to the company computers so a package can be continuously traced from the time of departure to within minutes of the estimated time of arrival. Many companies are embracing this new ‘office on the go.’ Among the industries that have found the edge in this phenomenon is the home service industry, which is utilizing the new office technology to improve services and of course increase business. Others include delivery services, home repair, and heating and air-conditioning services to keep workers on location and in the office in constant contact.

6.4.3 The Virtual Workplace

With the latest developments in telecommunication and computer technology, the virtual workplace is home to an increasing type of employees who work very briefly in their corporate workplaces, are mostly on the road, and often telecommute using personal or company-provided equipment. This breed of worker is rarely in a fixed workplace, but nevertheless he or she performs a full day’s work even if at the beach.

According to Snizek [18], the most important element of the virtual workplace is the use of computers and other telecommunication devices to link employees with the massive worldwide databases of vital information and other human resources. As computer and telecommunication technologies improve and bandwidth and computer miniaturization increase, this will not only lead to more workers opting for the virtual workplace but will also increase vital information flow into the company and corporate offices, which may lead to companies gaining a higher level of vital data, employee expertise, and experience from colleagues around the globe. The increasing popularity of the virtual workplace is mainly the result of recent changes in computer and telecommunication technology and organizational changes coming from corporate downsizing and outsourcing. For example, for corporations to keep the same level of overall effectiveness and efficiency and even to surpass it sometimes with fewer employees, companies are increasingly encouraging virtual offices, and the trend is likely to continue [18].

There are other benefits of the virtual office in overhead savings and other costs. With virtual employees rarely in their offices, a number of other employees can share their office space and other office resources, thus saving millions of dollars in facilities and equipment costs. The company may no longer need to have a large workforce on a permanent full-time basis. Companies can now use a limited staff and seek out contract expertise on a case-by-case basis as the situation arises.

In addition to the transformation of traditional workers, the virtual office is also opening doors to a new group of workers such as the disabled, the homebound, and the elderly who have traditionally been left out of the work force.

It is probably too early to talk about the long-term effects of the virtual office on both employees and employer, but some difficulties are already visible in both the employee and employer communities. Because most employee time is spent

outside the physical office, employees rarely meet face to face, so there is a lack of collegiality and of community spirit. Also, because most employees, especially experts, are not full-time employees of the corporation, there is a lack of belonging that affects employees and eventually undercuts their loyalty and hence their effectiveness. A company built on a transient work force lacks the moral force and legitimacy in the community in which its operations are based.

6.4.4 The Quiet Revolution: The Growth of Telecommuting

As workers have changed their work habits from working 40 h per week in the workplace environment to sharing those 40 h between being at the workplace and commuting to the workplace, the 9-to-5 time schedule for the majority of workers has started to crumble, with many moving their work locales outside the normal confines of time and space. Studies show that the largest number of workers doing their work outside their primary place of work do so in their homes. According to figures reported by Kraut [19], the percentage of home office workers or telecommuters in the total U.S. workforce by 1960 was 3%, but numbers have been on the rise ever since. It is estimated that by year 2020 close to 30% of the American workforce will be telecommuting [6]. This significant rise can be attributed to the growth in U.S. information-related work. In fact, JALA International [20], an international group of consultants in telework, telecommuting, and applied futures research, projects that more than 60% of the U.S. workforce will be information related by 2020. The growth of telecommuting is also driven by advances in office technology and the plummeting of prices for computers and telecommunication devices, the diminishing sizes of communication devices, and the increase in speed and bandwidth of communication devices.

As office technology improves, a large number of workers outside the self-employed professions of artists, writers, and craftspeople are potentially able to work at home. The advances in technology are making many types of jobs that used to require a worker to stay in an office environment more mobile. This trend is being helped further by the shift in global economies from manufacturing based to information based.

6.4.4.1 Categories of Telecommuters

There are three categories of telecommuters. The first category of telecommuters consists of workers who use their homes as an adjunct to their conventional office jobs. These workers are usually in white-collar jobs in areas such as management, research, market studies, and education. They are highly motivated. For them, occasional work-at-home is a flexible alternative used most in cases of critical work that can best be done at home to avoid the office environment. Kraut [19], reporting on a study he did in 1983 and 1984 at AT&T Bell Laboratories, states that these workers put in a full day's work at the conventional office in addition to taking some of their office work home. They additionally work at home on weekends and in the evenings.

The second category of telecommuters consists of workers who use their homes as the base for their businesses. The majority of these are in telemarketing, small start-up companies, and human services such as child care and elderly care. In contrast to the first category, these individuals are less educated and less likely to use a fully equipped electronic home office. Others in this category are the dispatchers in the home service industry, who are more likely to use a telephone and a computer without much data transmission.

The third category of telecommuters consists of those who have full-time jobs with large companies, but prefer through their own initiative to work from home. This category includes computer programmers, sales specialists, editors, writers, and those whose work depends on a high degree of creativity such as artists, musicians, and composers. This third category is a mixed bag of highly educated, independent, and specialized workers, and those who are not so highly educated but are very talented and skilled.

As computers and telecommunication technology become cheaper, and people obtain more access to smaller more portable computers, and other communication devices become more readily available, the home is becoming more and more a place of refuge for conventional office workers. Although it is not possible to predict the future direction of the home office, it is likely that if the technology that has caused the increase in the home office keeps on track, the number of telecommuters is likely to continue growing, with the majority of workers remaining in home offices for economic benefits and convenience.

6.4.4.2 Company Role in Telecommuting

To many, the home office is a revisit to the cottage industry of the fifteenth through eighteenth centuries in Europe: Raw materials were dropped off at the workers' cottages and finished products later picked up for market. Ever since industrialization, factories and later companies have been using home workers. Thus, company experimentation with their employees telecommuting is not a new idea.

The home office has always been prompted by new advances in technology and by the need of businesses to become more productive with minimum expenditures. As the Internet and globalization open up new international competition and as new technologies make telecommuting more acceptable to employees, company-sponsored telecommuters will increase.

By the 1960s, according to Kraut [19], telecommuters accounted for 3.6% of the U.S. workforce, and a small portion of this was company sponsored. But by the late 1970s and early 1980s, big companies such as IBM and AT&T were deeply involved with telecommuting experiments. Kraut estimates that by 1983 IBM had more than 8000 employees telecommuting. These big companies and other smaller ones spent money in this experiment, expecting a benefit in return. The experiments were also meant to provide information on the classification of work suitable for the home, to identify individual workers who could work at home, and to suggest how workers were to be monitored as they worked from their homes.

Although no study has yet reported big monetary benefits for companies from these experiments, some studies on other related issues have provided some results. For example, on the issue of remote supervision, classification of jobs fit for telecommuting, and identifying individuals better suited for telecommuting, a study by Olson and Lucas [17] provided some partial answers. On the issue of classification of work, the study found that work with possible measurable milestones is most suited for telecommuting. On the issue of identifying individuals most suited to telecommute, the study found that people who usually need less supervision at the office and those who do volunteer work are the most suited to telecommute. These conclusions are also influenced by the nature of the work, gender, age, and labor supply. The study also highlighted difficult issues such as the effect of telecommuting on the promotability of employees because visibility is key to promotion. There was also some clarification of the issue of pay. Telecommuters tend to be paid less because their pay is based on output, which makes output the real mechanism of monitoring telecommuters [17].

6.4.4.3 Effects and Benefits of Telecommuting

Whenever there is a change in the environment of workers, there are always some social, psychological, and financial effects on both employee and employer. If the effects are financial, they become benefits. However, if they are psychological, they become health issues; if they are social, they become organizational issues. In this section we concentrate on social and financial issues.

An employer–employee-arranged home office is supposed to reap benefits for both parties. Let us start by qualifying our discussion to include only those telecommuters who are company employed, have traditional offices at the company premises, and through mutual arrangements with their companies have decided to work from their homes. This group truly exemplifies the benefits, if there are any, for both the employer and the employee. Because these workers have a choice of either staying at the office or working from home, they can only work only from their homes if they experience a benefit and the companies can only let them work from their homes if the companies expect a benefit from the arrangement. For those working at home with no choice except to work at home, such as those in the majority in the second category (see earlier), the benefits are already clear. Defining benefits for telecommuters is not easy because each participant in the arrangement perceives the benefits the way they would like them to be. For example, the company may see the benefit as savings on office space so that other workers can use the space, or as savings in office supplies, or a reduction in the likelihood of employee risks while on company premises. The employee may see benefits as spending more quality time with their loved ones at home, or spending less time in traffic commuting to and from work, or the flexibility and independence in decision making concerning the work the employee has to do.

The value of benefits from this arrangement depends on individual circumstances as discussed by Kraut [19] and reported as follows:

1. *Gender*: Women have traditionally given care to children and the elderly, the two groups most homebound; women would therefore draw maximum benefits from telecommuting arrangements with their employees, if their primary objective for telecommuting is to take care of their families.
2. *Nature of work: managerial, clerical, sales, or service*: The nature and type of work one does also influences the kind of benefits one obtains. For example, clerical work tends to be more supervision intensive than managerial and professional work. In those types of work where supervision is not so intensive, there is a high degree of latitude for one to make decisions. However, jobs that are supervision intensive are less likely to be moved into home environments. If such jobs are to be moved to a home environment, chances are that the company may not garner any benefits, but employees may benefit by getting more freedom and flexibility in their work routine and in decision making.
3. *Labor supply*: When there is a limited supply of one type of worker, companies try to find innovative ways of attracting and keeping workers in those limited-supply areas. For example, in 1981, IBM, anticipating a demand in programmers and engineers, started a telecommuting program to attract young talented programmers and engineers. Members of such groups usually garner great benefits with such an arrangement.
4. *Age*: Age may be a factor in home office productivity. For example, in sales, young people are more productive outside of offices than older workers. In management, older people are more productive in offices than outside offices. Women in their childbearing years are more productive when they telecommute than when they work in company offices. So, using the age factor, both employer and employee can benefit from a home office.

The U.S. Department of Transportation summarizes the benefits of telecommuting for both employees and employers as follows [21]:

- An individual benefits from telecommuting because he or she immediately eliminates the time, trouble, and expense of physically commuting to work. This change gives the average person an extra hour per day, right off the top, to use for the thinking, writing, telephoning, planning, and reporting work that keeps the business organization moving forward.
- The benefits of telecommuting also translate directly and immediately into more discretionary time, less stress, and general health improvements.
- More autonomy in work decisions and having more control over time and more flexibility in job variations.
- Decreased commuting expenses for the individual.
- More quality time with family with less or no frustration at home.
- Employers benefit from the extra productivity, reported to be consistently at 10–15% in many studies in the past two decades. Employers also save on expenses

through having fewer employees on company premises. Such savings come from the daily need for offices, desks and chairs, restrooms, copy machines, parking spaces, heating and lighting, and all the rest.

- In addition, telecommuting helps the best and satisfied employees stay longer, thus saving on recruiting and training costs.
- Society benefits from telecommuting through benefits to the environment.

However, the overall benefit to employers of home office workers is evaluated through such measures as the productivity of the employee. According to the report of the National Academy of Sciences and Electronic Services Unlimited [19] conducted in 1984, productivity among both clerical and managerial telecommuters of the 24 projects evaluated in the report increased about 15–25%. Reductions in expenses come from office space, equipment purchases and wear, and office supplies. Businesses also reduce costs by hiring contract workers as the need arises, hence reducing the overall worker benefit expenses. Beside the reductions in overhead costs, by employing contract home workers, companies tap into the scarce resources of the professional expertise they would not otherwise find.

Telecommuting is not all positive, however. Among the issues that negatively affect the company image are employee morale and alienation. Because of the lack of professional contacts, employee morale may suffer and they may feel abandoned by the company. If this happens, productivity falls. Another negative impact is the public's perception of the company when they see an employee mowing the lawn at 3 p.m. on a workday.

6.4.5 Employee Social and Ethical Issues

Mention of the phrase *office automation* used to conjure up nightmarish images of less control, helplessness, joblessness, and the stagnation of humanity. Within the context of office automation, the concept implies the idea of massive layoffs because offices with intelligent machines may require fewer people. Besides the fear of layoffs, workplace automation has also been plagued with the issue of *deskilling*, meaning stripping an employee of job skills as a result of changes in either job content or procedures. Deskilling, according to Attewell and Rule [22], can either be intraoccupational, in which case the skill content of the job decreases over time, or inter-occupational, in which very few people gain the skills needed for the job, causing either low-paying jobs or layoffs. Driscoll [23] expressed the fear of deskilling in a more sarcastic way by saying that the office of the future would “leave people in only two roles: bosses and garbage collectors.” But so far these horrific fears of deskilling have not been realized, even with the heavy office automation of the past 10 years.

There have been some layoffs and deskilling of employees, but the numbers have been very small. Several factors such as the following have prevented this from happening:

1. The willingness of employees to retrain and use the newly acquired technology, which, of course, has led to the upgrading of skills in the workplace. In fact, according to Attewell and Rule, computerization has led to reskilling of employees rather than deskilling.
 2. The historical patterns show that more efficient production techniques lead to expanded operations and added growth, which leads to more hiring rather than firing of existing employees.
 3. In anticipation of automation, more employees are usually hired to cope with the new technology and to handle the expanded work capacity.
-

6.5 Employee Monitoring

In the past decade, most of the large industrialized economies have been shifting from a heavy manufacturing base to an information management base. Along with this shift has been stiff competition resulting from globalization. Competition is coming from not only large economies but also from upcoming developing countries. These developing economies with their cheap labor costs are making this competition more costly for a number of older, more established and mature economies.

This shift in the economies and the stiff competition have resulted in a shift in management styles to bring more efficiency and quality in the established economies. This is not the first time such management styles have shifted. Styles in management have been changing with shifts in economies since the dawn of the Industrial Revolution. In those early days, management followed a style now commonly known as Theory X, after Douglas McGregor. Theory X management, with all the trappings of the industrial era, was characterized by a top-down autocratic style of management in which the manager—literally from the top floor—commanded the activities of the factory workers on the factory floor with almost omniscient and demeaning power.

As economies grew larger and employees became more elite, a new management style started to evolve that became known as Theory Y. Theory Y put more faith and empowerment in the hands of the employees. The style was hierarchical with the employee ranks broken down into small semi-independent units. Each unit was headed by a supervisor. The supervisors themselves formed another top-down hierarchy ending with the top management. Theory Y, or scientific management, as this management style is commonly known because of its hierarchical structure, gave more flexibility and partial decision-making powers to employees at different levels of the management hierarchy. The workers themselves were more removed from the top management, but at the same time they were closer to management

decisions and control from the smaller units. Scientific management has been in effect for years.

With the recent shifts and globalization of world economies, however, scientific management has been slowly giving way to a new style in which management is trying to wrest control of the work process away from the workers and slowly bring back the techniques of Theory X. Given the technological advances of recent years and the abundance of educated and highly skilled workers, however, it would be unwise for today's management to bring back these techniques. So, a new technique in the works is called "fear management." It is aimed at keeping workers in line, just like all other management styles, but with "voluntary" compliance by workers to company management policies and practices they would normally have questioned or challenged.

Differing from theories X and Y, which achieved worker control through autocratic and supervisory unit means, fear management uses both worker surveillance and control as enforcement means. Fear is transmitted to workers through policies such as "downsizing," "contingent workforce," and "outsourcing." To workers these policies spell disaster and fear of losing job security and being replaced by part-time, temporary, and contract workers. According to Karen Nussbaum [24], temporary workers now make up one-third of the U.S. workforce; less than one-half are covered by any pension, and many have no health insurance.

Management is using a wide array of surveillance gadgets and techniques, which include employees taking polygraph tests if they are suspected of a breach of any kind. Although compulsory use of the lie detector is banned in the United States, it is still used on a voluntary basis. Drug testing is widely used by many companies and is required of all U.S. government employees in some categories. Handwriting analysis, the honesty test, electronic monitoring, mind control, and many other techniques are also being used.

6.5.1 Workplace Privacy and Surveillance

The electronic office or workplace has provided management with a bonanza of new possibilities for monitoring employees in their drive to reduce ever-increasing workplace costs. The issue of employee monitoring is not new because of advances in computer technology. Ever since the Industrial Revolution, workers have been monitored for performance evaluation because its performance has been used as the basis for pay and for decisions about employee advancement. Monitoring has also been employed to control employees and impose overall discipline in the workplace. But before the advent of surveillance gadgets, workplace monitoring was done through human eyes—those of the supervisor.

As workplace modernization picked up speed with advances in technology, the techniques and debate surrounding employee surveillance intensified. The battles were fought on two fronts: those who see monitoring as good management control tools with plausible reasons such as increased production, more accurate assessment of employee performance, greater organizational control over employees,

immediate feedback on individual employees (which can lead to high motivation), and more flexibility in work location, and those who see surveillance as an outright transgression of employee privacy, causing problems such as stress, decreased job satisfaction, and an affront to human dignity. The replacement of the human eye with an electronic one, on guard 24 h a day, 7 days a week, without taking a break, and easily concealed, started the real erosion of employee privacy.

Employers collect information from employees through two channels. The first is the voluntary channel in which employees surrender the information through forms, interviews, worker sessions, and worker get-togethers. The first work-related information collected from the employee by the prospective employer is collected from the job application, followed by more information surrendered by the prospective employee during the interviewing process. Most of the time this information is given voluntarily because the person wants to get a job and of course employers need employees they can trust. After being hired, especially during the first few days at the job, the new employee usually fills out myriad forms for an employee folder so the employer can pay for the new employee's benefits, taxes (part of them anyway), and salary.

The second channel is the private information the employer gathers through surveillance. The degree, rate, and method of surveillance depend on the employer and how much information is needed from the employee and the value of that information to the employer. The information collected is supposedly used solely for managerial decision making regarding employee work assignments, individual feedback, pay increases, bonuses, promotions, and other benefits, and, of course, termination. If most of this information, legitimately collected or otherwise, was used solely for employee benefits, very few would complain. But sometimes it is not, which is when employee privacy issues arise. For example, how much personal information is needed by the employer for employee benefits before it becomes an invasion of the employee's personal privacy? Are there restrictions on the use of that information? Does the employee have the right to view any information collected on him or her? Is employee surveillance legal, and if so, what legal avenues does an employee have?

According to Adler et al. [25], there are no general explicit constitutional rights to privacy in the United States except in a few states. The U.S. Privacy Act of 1974 has limited applicability, mostly to federal employees. Private employees are not adequately covered by this act; they are only covered by a threat to sue for libel, discrimination, and ethical consideration. In light of the limitation of both the U.S. Federal Privacy Act and state statutes, courts started to recognize independent torts, according to Adler et al. [25]. This judgment may be true in many other countries.

Is employee surveillance an invasion of employee privacy? That depends. Notice that invasion of privacy does not mean collection of information on an individual without the individual's knowledge but rather the disclosure of collected information on an employee without legitimate reason or interest. An employer can gather information from the employees with whatever means so long as that information is not used maliciously. For example, an employer can collect information from an individual employee through covert actions such as electronic monitoring

and use that information solely to further business interests without disclosure to the employee. According to Adler et al. [25], this procedure is legal, and most courts have recognized it as a legitimate business interest and have sided with employers.

Adler et al. cite a case that further clouds the employee privacy issue. An employer was requested by a court to provide an employee's records. In such a case the employee may not have any legal rights regarding the information the employer has. If the employer refuses the request, the employer can be cited for contempt of court. But if the employer obliges, he or she may be charged with violating the employee's privacy rights.

Why are the employee privacy issues becoming so important? As the U.S. and many other economies shift toward information-based economies, the value of owning information for economic advantages becomes even greater. Many companies are trying to obtain information on individuals to market their products, to approve loans, to offer audits, and many other revenue sources. Because companies such as insurance, banks, loan assurance, and legal investigations want the information on their clients to be as accurate as possible (their businesses depend on it), information-gathering companies see the employer as their best source of such accurate and reliable information.

Individual information has become valuable not only to banks and insurance companies that want security for their money but also to a cross section of manufacturing and service companies. These companies want to find new markets for their products. To do that, they need a source from which to launch their marketing and get a foothold in specialized markets. This kind of information can best be acquired from employers. Once a company has gathered that information about individuals, it can model its market strategies around the characteristics exhibited by these individuals. Such information may include income levels, leisure activities, foods, favorite wines and beers, whether one eats chili, and so on.

In this rush for personal information, the employer takes center stage as the best source of such intriguing tidbits. Statistics show that the workplace is second only to the home as a place where we spend most of our time. It is common sense, therefore, that the workplace should be the next best place to look for information on an individual.

6.5.2 Electronic Monitoring

Electronic monitoring is generally the monitoring of employees using electronic devices such as video cameras, computer equipment, audio devices, and many other concealed gadgets. In most cases it measures the quality and usually the quantity of work and the ability and effectiveness of the worker. In other cases it also measures the worker's habits on and off the work premises because some employers believe these habits have a great bearing on employee performance. For example, if the employee is a drug user, the effects of drugs will eventually affect the quality of that employee's work.

Electronic monitoring of employees is characterized by workers' ignorance that they are being monitored, fear of the ever-watching eyes of the supervisor, and fear of how much that supervisor knows about them. Let us illustrate these fears by two short examples from Nussbaum [26]. She first cites the case of *Mary Williams v. United Airways* in which Mary Williams was first disciplined for her remarks to a coworker, sent to a psychiatrist, and subsequently fired from her work at United Airlines because she confided to a coworker about an obnoxious customer while management was listening. In another example in the same paper, Nussbaum cites a New York data processor whose boss kept flashing the message "you are not working as fast as the person next to you" on her computer screen.

There are thousands of cases similar to these two arising from employee monitoring. Although there are no comprehensive studies on the spread of electronic monitoring in the workplace, it is generally believed that electronic monitoring of employees is on the rise and is already enshrined in the banking, insurance, and airline industries, to name but a few.

As technology becomes cheaper, therefore more affordable, smaller, and easier to conceal, the trend is likely to pick up momentum as the pressure for quality, quantity, and standards increases because of global competition. This pressure is likely to force more companies to resort to electronic monitoring as a way to control employees to extract more performance, compliance, and probably more money. In fact, in some sectors of employment the percentages are already high. For example, according to Grant et al. [27], in the United States 25–35% of all clerical workers are electronically monitored for work performance.

6.5.2.1 Effects of Electronic Monitoring on Employees

Recently, I watched a British television comedy in which the theme was employee monitoring. The setting was a department store. The managers of the store found out they were losing merchandise and decided the employees were the most likely culprits, so they hired a security guard to check all employees' bags and pockets at the end of each day as the employees left the premises. They also installed video cameras throughout the store, including the restrooms. With the cameras in place, all human movements could be monitored in the comfort of the manager's office. Employee morale and performance declined considerably because employees, already aware of cameras watching their every move and carefully recording their every word to customers, were more concerned about being seen sweet-talking their customers and looking smart than actually working. Employees neglected those parts of the store where they could not be seen "working" by management. Also there were fights between employees to take those strategic places. Funny as the television episode was and indeed as it was intended to be, it illustrates a number of issues research has shown to exist among electronically monitored employees.

In research conducted by a North American insurance company reported by Grant et al. [27], results similar to those portrayed in the television comedy were observed. The research studied a monitored group of the group claims-processing division of the insurance company and a non-monitored group of the

same division. In these two groups, the researchers included some supervisors and department managers. The monitored group was responsible for entering and paying claims using an automated claim-processing system and interacting directly with subscribers answering their questions and settling their disputes so far as payments were concerned. Their work included printing checks and talking to customers on phones.

The computer “monitor” counted the number of checks produced by an individual on a daily basis. According to the findings of the research, the group that was monitored considered the number of checks printed as the most important part of their work. In fact, they thought that talking to subscribers was an impediment to their job of printing checks. These employees, just like those in the British comedy, focused on those parts of their jobs they thought were being monitored and neglected all other essential parts. Although the monitored group did their work this way, the researchers found that the employees in the non-monitored group had a different perception of their work. This group thought that interacting with customers was the most important part of their work.

Another research project conducted by Irving et al. [9] compared two groups of employees working in insurance companies, financial institutions, and government. One group was electronically monitored and the other was not. The researchers considered the effects of monitoring on issues such as job satisfaction, what employees consider as a measure of performance, amount and usefulness of feedback, and relationships among employees and between employees and supervisors. The results of the study were very similar to those of Grant’s study and the British television comedy. Employees put much more emphasis on quantity as a measure of performance; there was no significant usefulness in the timely individual feedback; rewards were based on electronic performance evaluations in the monitored group, and those in the monitored group felt, and rightly so, that they were more supervised than any other group in the company. From these studies two important issues emerged.

1. Very often an intended goal of a monitoring program may be clouded by a different goal perceived by the monitored group. Therefore, without a well thought out electronic monitoring program, the intended goal of the company may be lost in the resulting perceptions of the employees.
2. The psychological effects on the monitored employees may be more severe than previously thought and anticipated. The philosophy that “if it isn’t counted, it does not count” should not be allowed to flourish among employees. Besides what has been observed in Grant’s study and the British comedy, there are social, ethical, and mental effects on the monitored employees.

6.5.2.2 Consequences of Electronic Monitoring

The most devastating effect of electronic monitoring on employees is fear of losing their jobs. For many of us a job is the only source of livelihood and any sign of

losing it triggers fear. In addition to fear of job loss, electronic monitoring also causes the following problems:

Reduced task variety: The type of work monitored most is of a low-skilled, repetitive nature. In these jobs employees take the quota to be the measure of work and usually cannot afford to take a break, let alone slow down, thus increasing the monotony of their activities.

Lack of individual initiatives: Most monitored jobs do not require personal creativity because they are of a low-skilled, repetitive nature. The employee usually is not allowed to vary the procedures but must follow them to the letter.

Reduced or no peer social support: Monitored groups are always given separate stations where gadgets can monitor them in full view. Thus, an employee must remain where he or she can be “seen.”

Lack of self-esteem: The isolation, the monotony of work, the lack of creativity, and the lack of freedom to vary job steps lower employee morale and consequently self-esteem.

Lack of interest in the job: With low self-esteem, many people definitely lose interest in their jobs.

Lack of trust among workers, between workers and supervisors, and between supervisors and management: This lack of trust can result in low company morale, and later production levels may begin to fall. As employee morale plummets and dislike of the employer rises, workers turn to litigation, filing privacy suits against their employers. Nussbaum reports that in the United States there were twice as many lawsuits of workplace privacy filed between 1984 and 1987 as between 1978 and 1980. In the courts, although workers’ privacy rights have not been successful in the number of lawsuits filed, there is a growing recognition of workplace privacy rights. A number of states in the United States and indeed in other countries have been working on legislation to protect workers. The trade union movement has also been actively negotiating languages in worker contracts to help curb unnecessary workplace monitoring.

Alienation: Sociologists define the concept of worker alienation as lack of worker freedom and control, purpose and function, and self-involvement in their work. Alienation, according to Shepard [28], is lower among workers in industries with automated technologies.

6.6 Employee Health and Productivity in the Workplace

Productivity of workers depends on the quality of their physical and mental state. Employers have always striven to make their workers happy, healthy, and productive.

There is now a movement to improve employee work environment as companies start to add facilities such as employee gyms, cafeteria, daycare centers, and worker facilities for their employees. For example, currently Google, Inc., is

cited by many reports to be the top company in this movement. There has always been a feeling of powerlessness among employees to control the state of working conditions because they lack freedom and control. According to Shepard [28], a worker has freedom and control at work if he or she can vary steps involved in doing the job, determine work methods and workload, and increase or decrease speed at which the work is done. With the changing work environment caused by advances in computer technology, employers are finding themselves achieving what has eluded them for years, offering their employees happiness, healthier environments, and high productivity through empowerment.

Human beings always want to feel they are in control of their work and other aspects of their lives. The changing work environment gives the workers the choice to work either in a traditional office or from home. Choice brings commitment and obligation. When people make a choice of their own, they tend to commit to the requirements of their choice. The commitment to work always translates into higher productivity quotas. Although computer technology has given workers more control in decision making, it has also given them new dangers in the workplace. These dangers are collectively discussed next as ergonomics.

6.6.1 Ergonomics

Ergonomics is an applied science concerned with designing human–machine interactions that offer and maintain a safe, comfortable, healthy, and habitable work environment. With the increasing automation of the workplace, our dependence on machines is on the rise, and the field of ergonomics is correspondingly expanding. It now covers a wide array of work environments and factors that influence the employee's health and wellness through prevention of occupational diseases. In particular, ergonomics studies the design of human work and production because when the demands for human performance of a task exceed human capacity then ergonomic injuries start to occur and human wellness declines.

An ergonomic injury results when the demand on a person to perform a task exceeds that person's working capacity. Examples of ergonomic injuries include work accidents that occur because of the overwhelming demand for performance and all work-related musculoskeletal disorders such as back pain, neck, and shoulder pains, and repetitive strain injuries (RSI), with most studies now focusing on RSI.

6.6.1.1 Repetitive Strain Injuries

RSI is a set of work-related musculoskeletal disorders caused by repeated and prolonged body movement resulting in damage to the fibrous and soft body tissues including tendons, nerves, and muscles. Some RSI conditions are well known in medical communities, but a number of others are still very obscure and difficult to diagnose because they present with different and very often unpredictable

patterns. RSI as a disease is not new; it has been affecting people performing repetitive motions, such as cashiers, musicians, and assembly and data entry workers for years; it has just recently gained prominence because of the rise in computer availability and widespread computer use. Recent studies have isolated some of the main causes of RSI as repetitive motion, forced gripping, performance stress, alienation, static loading, fixed posture, deviated wrists, and boredom. Computer users of keyboards, mouse, tracking balls, touchscreens, and the foot mouse are among the groups most prone to RSI. RSI attacks those body parts such as tendons, wrists, shoulders, nerves, and arms, and sometimes the neck, that receive tremendous stress exerted by body movements. This condition, which has come to be known by a string of names such as occupational overuse syndrome (OOS), cumulative trauma disorder (CTD), carpal tunnel syndrome (CTS), and upper limb disorder (ULD), causes serious pain and if not treated early may even cause permanent disability. As a result of the damage to the nerves, wrists, arms, tendons, and muscles, the disease also causes eyestrain, fatigue, headaches, usually back pain, tingling, coldness, hand numbness, and stiffness and discomfort in body movement, especially of the fingers, arms, and head. When RSI is caught in time, it can be cured with proper care and prescriptions that emphasize changes in individual work styles and techniques. Among the suggested changes in work styles and techniques are the following:

1. *Use ergonomically correct work equipment.* Such may include chairs, tables, and computer equipment like new keyboards, monitors, new software, and new lightning in the workplace.
2. *Use a light touch on the keyboard to place less stress on body parts.* Also keep the wrists straight in line with your arms.
3. *Take frequent breaks from your work.* Do not work for long hours without a break. Once you get a break, walk around and do some stretching exercises.
4. *Educate yourself about RSI.*
5. *If necessary, reduce the time you spend at the computer terminal.*

Improvements in the design of human work and occupational environments can result in benefits to the employee and the employer. Among such benefits are the following:

- Reduced medical bills
- A higher level of self-esteem
- Increased productivity because of fewer employee errors. High attendance rate and retention skills increase per capita output

Studies have shown dramatic increases in the range of 20–50% in increased productivity after effective ergonomics remedies were implemented for people working with visual display units (VDU) [29].

6.6.1.2 Stress

Besides RSI, stress has also recently drawn public attention as a work hazard. Similar to its counterpart RSI, stress has been targeted to explain much worker discomfort and frustration that may lead to poor job performance, strained interpersonal relationships, and erratic employee behavior. Stress is believed to have its origins in environmental inputs, and it appears through symptoms such as fear, anxiety, and anger. Anything that increases the stress level of an individual ultimately endangers that individual's health.

In the work environment stress is mainly caused by a variety of factors including impending deadlines, long hours at work, uncooperative colleagues, lack of management support and understanding, constantly changing requirements, and lack of job capacity because of changes in either work procedures or the workplace environment. Stress onset affects individuals differently depending on the environment they are in, and different individuals react differently to stress. For example, Ivancevich et al. [24] report that under stress women consume less coffee than men, shout less but consume more aspirin, and visit doctors more frequently than men.

Employers can significantly reduce employees stress by enhancing the overall work environment, keeping consistent work schedules, giving fewer deadlines, and making fewer management demands. Health awareness and knowledge of the causes of stress are always the first step in controlling it.

Exercises

1. Discuss the effects of telecommuting on family life.
2. If there are benefits to the employer for telecommuting, why is it that not many companies have embraced telework?
3. Ergonomics is turning into a multimillion-dollar business. Discuss.
4. Electronic monitoring has more negative effects on both employers and employees. Discuss.
5. Work productivity is directly related to the well-being of the employee. Discuss.
6. Has automation caused any massive worker layoffs?
7. Has technology in the workplace created jobs or trimmed job rolls?
8. There has been a debate on the existence of the digital divide. What is your opinion? Is there one or not?
9. Is there anything like equal access to computing technology? If otherwise, is it achievable in any society?
10. The concept of telecommuting has not realized its earlier potential and, therefore, has not been successful. Is this a true statement? Why or why not?
11. Has the Internet, together with new developments in telecommunications, increased the value of telecommuting?

12. Have the Internet and related technologies helped to lessen the problems experienced by the telecommuter?
13. Discuss the social implications of telecommuting.
14. What, if any, are the ethical implications of telecommuting?
15. What are the benefits, if any, of employee monitoring?
16. What are the social and ethical problems resulting from employee-mandated drug and polygraph testing?
17. Why do you think employee monitoring is such a hot issue?
18. There are benefits to employee monitoring! Discuss.
19. Should employees sacrifice their privacy because of fear of losing a job?

References

1. L. Kruger, A. Gilroy, Broadband internet access and the digital divide: federal assistance programs, in *Congressional Research Service*. <https://digital.library.unt.edu/ark:/67531/metadc958717/m1/1/>
2. C. Notari, *What is the Status of Global Digital Inclusion*. <http://intelligentinclusion.com/2012/05/what-is-the-status-of-global-digital-inclusion/>
3. T. Wayne, *Digital Divide is a Matter of Income*, er 12, 2010. The New York Times, 12 Dec 2010. http://www.nytimes.com/2010/12/13/business/media/13drill.html?_r=1
4. J. Jansen, *The Better-off Online*. Pew Research Center Internet & American Life Project, 24 Nov 2010. <http://pewresearch.org/pubs/1809/internet-usage-higher-incomeamericans>
5. National Telecommunications and Information Administration, Technical Report 2000 (NTIA), *Falling Through the Net: Toward Digital Inclusion* (2000). <https://www.ntia.doc.gov/files/ntia/publications/fttn00.pdf>
6. J. Washington, *For Minorities, New 'Digital Divide' Seen*. Associated Press, 1 Nov 2011. <http://www.rolandsmartin.com/blog/index.php/2011/01/11/for-minorities-new-digital-divide-seen/>
7. A. Smith, *Mobile Access 2010*. Pew Research Center Internet & American Life Project, 7 July 2010. <http://www.pewinternet.org/Reports/2010/Mobile-Access-2010/Summary-of-Findings.aspx>
8. A. Rickert, The Dollar Divide: Web Usage Patterns by Household Income. Media Metrix, August 2000. <http://www.mediametrix.com/data/MMXI-USHHI-0600.pdf>
9. R.H. Irving, C.A. Higgins, F.R. Safayeni, Computerized performance monitoring systems: use and abuse. Commun. ACM **29**(8), 794–801 (1986)
10. World Bank Global Economic Prospects 2008: Technology Diffusion in The Developing World. [http://siteresources.worldbank.org/INTGEP2008/Resources/4503313-119947339970/Technological-progress-\(large\).gif](http://siteresources.worldbank.org/INTGEP2008/Resources/4503313-119947339970/Technological-progress-(large).gif)
11. World Bank Global Economic Prospects 2008: Technology Diffusion in the Developing World. http://en.wikipedia.org/wiki/Global_digital_divide
12. United Nations Human Development Report, United Nations Development Program (1999). http://hdr.undp.org/sites/default/files/reports/260/hdr_1999_en_nostats.pdf
13. F. Rodriguez, E.J. Wilson, *Are Poor Countries Losing the Information Revolution?* Info DEV (The World Bank, Washington, DC, 1999)
14. F. Bankole, F. Shirazi, I. Brown, Investigating the impact of ICT investments on human development. Electron. J. Info. Syst. Dev. Ctries **48**(8), 1–19 (2011)

15. Y.J. Kim, H. Kang, G.L. Sanders, S.T. Lee, Differential effects of IT investments: complementarity and the effect of GDP level. *Int. J. Inf. Manag.* **28**(8), 508–516 (2008)
16. C.C. Gottlieb, A. Borodin, *Social Issues in Computing* (Academic Press, New York, 1973)
17. M. Olson, H. Lucas, The impact of office automation on the organization: some implications for research and practice. *Commun. ACM* **25**(11), 838–847 (1982)
18. W. Snizek, Virtual office: some neglected considerations. *Commun. ACM* **38**(9), 15–17 (1995)
19. R. Kraut, Predicting the use of technology: the case of telework, in *Social Issues in Computing: Putting Computing in its Place*. ed. by C. Huff, T. Finholt (McGraw-Hill, New York, 1994), pp.312–334
20. JALA US Workforce. <https://www.jala.com/usworkers.php>
21. Benefits of Telecommuting. U.S. Department of Transportation's Departmental Office of Human Resource Management. http://dohr.ost.dot.gov/Telecommuting/benefits_of_telecommuting.htm
22. P. Attewell, J. Rule, Computing and organization: what we know and what we don't know. *Commun. ACM* **27**(12), 1184–1193 (1984)
23. J. Driscoll, Office automation: the dynamics of a technological boondoggle, in *Emerging Office Systems*, ed. by R.M. Landau, J.H. Bair, J.H. Siegman (Ablex, Norwood, 1982)
24. J. Ivancevich, A. Napier, J. Wetherbe, Occupation stress, attitudes, and health: problems in the information systems professions. *Commun. ACM* **26**(10), 800–806 (1983)
25. P.A. Adler, L.K. Parsons, S.B. Zolke, Employee privacy: legal and research developments and implications for personal administration, in *Social Issues in Computing: Putting Computing in its Place*. ed. by C. Huff, T. Finholt (McGraw-Hill, New York, 1994), pp.312–334
26. K. Nussbaum, Computer monitoring: a threat to the right to privacy, in *Ethical Issues in Information Systems*, ed. by R. Dejoie, G. Fowler, D. Paradice (Boyd & Fraser, Boston, 1991)
27. R. Grant, C. Higgins, R. Irving, Computerized performance monitors: are they costing you customers?, in *Social Issues in Computing: Putting Computing in its Place*. ed. by C. Huff, T. Finholt (McGraw-Hill, New York, 1994), pp.312–334
28. J. Shepard, *Automation and Alienation* (MIT Press, Cambridge, MA, 1971)
29. E. Grandjean, *Ergonomics in Computerized Offices* (Taylor & Francis, London, 1987)

Further Reading

30. L. Bailyn, Towards a perfect workplace. *Commun. ACM* **32**(4), 460–471 (1989)
31. L. Flynn, They are watching you: electronic surveillance of workers raises privacy concerns. *San Jose Mercury News*, 13 June 1993, p. 1F (1993)
32. M. Payser, *When e-mail is Oops-Mail: Think Your Private Messages are Private? Think Again*. *Newsweek*, 16 Oct 1995, p. 82
33. S. Sauter, M. Gottlieb, K. Jones, V. Dodson, K. Rohner, Job and health implications of VDT use: initial results of the Wisconsin NIOSH study. *Commun. ACM* (1983)



Software Issues: Risks and Liabilities

7

Abstract

Software Issues: Risks and Liabilities focuses on the issues that arise out of the relationship between the developer and the buyer, including claims, user expectations, and the legal ramifications that may follow an unhealthy relationship. The discussion touches on standards, reliability, security, safety, quality of software, quality of service of software products, causes of software failures, developer and buyer protection, and techniques for improving software quality. Causes of software failures or poor performance of a software product are discussed, attributing the causes to a variety of reasons but most notably human error, the nature of software itself, and the environment in which software is produced and used. Finally, historic examples of software caused accidents are given including the Therac-25, the Space Shuttle Challenger, the Indian Bhopal chemical accident and the Chernobyl Nuclear Power accident. Both consumer protection and techniques for improving software quality are also discussed.

Learning Objectives

After reading this chapter, the reader should be able to

1. Explain the limitations of software testing as a means to ensure correctness.
2. Describe the differences between correctness, reliability, and safety.
3. Discuss the potential for hidden problems in reuse of existing software components.
4. Describe current approaches to manage risk and characterize the strengths and shortcomings of each.
5. Outline the role of risk management in software systems design.

Scenario: Who Will Pay the Price for Flawed Software?

Peter Efon works as a programmer for a major software company. The company, Cybersoft, is launching itself to be a major Internet-based platform developer and it is soon to launch a web initiative. Peter is involved in the development of a crucial component of the initiative. The company has trust in Peter for he has worked for it since he left college 15 years ago. Since his arrival at the company, Peter has pioneered a number of major software development projects. Peter has followed, and is very much aware of, the losses suffered by other businesses due to defective software. He even knows that in 2000, US companies suffered a whopping \$100 billion loss due to bad software. He and his company, Cybersoft, are determined to target quality as the major focus of their new web initiative. Peter dreams of the success of the web initiative and the recognition it might bring both to his company and him. However, a few days before the launch of the much-awaited initiative, as Peter makes his final quality checks, he discovers a flaw in the core component of the initiative whose magnitude he could not determine. To do so would mean a few weeks delay at best, a major blow to the company's efforts.

The company had mounted an advertising blitz on all major media outlets. Even a few weeks delay would cause major financial losses and the public's loss of confidence in the right company. This must never happen. Peter decides to see to it.

Discussion Questions

1. Is Peter Efon wrong?
2. What damage would Cybersoft have suffered had there been a delay?
3. What do you think would have been the right course of action for Peter and Cybersoft?
4. Can you estimate the damage?

7.1 Definitions

Software is a set of computer programs made up of a sequence of short commands called instructions that tell the computer what to do. Normally, software is in one of two forms, either built into the computer's more permanent memory, called ROM (read-only memory), or loaded on demand in the less permanent but more volatile memory called RAM (random access memory). A *software producer*, or *developer*, creates or develops a set of programs to meet the specifications of a user, if there is a contract, or of a specific problem if it is general software. Developers are either individuals working alone or companies such as Microsoft, which employs hundreds of software engineers including analysts and programmers. *Software buyers*, or *customers*, obtain the finished software from the developer to satisfy a need,

basing their decision on developer claims. The buyer may be an individual or a company.

In this chapter, we focus on the issues that arise from the relationship between the developer and the buyer, including claims, user expectations, and the legal ramifications that may follow an unhealthy relationship. The discussion touches on standards, reliability, security, safety, quality of software, quality of service of software products, causes of software failures, developer and buyer protection, and techniques for improving software quality. Let us begin by defining these terms.

7.1.1 Standards

Software developers must convey to buyers' satisfaction that their products are of high quality. The buyer, however, has little leverage in disputing the claims of the developer in these areas because there is no single universally acceptable and agreed-upon measure of software standards. But there are universal basic standards that a software product must meet. Such standards include the mutually agreed upon criteria and expectations of the buyer. In this case, the law imposes such standards, and if the product does not live up to them, the buyer has the right to pursue legal action. There is no one criterion that can be used to measure software standards but rather a collection of criteria such as development testing, verification and validation of software, and the programmer's professional and ethical standards.

7.1.1.1 Development Testing

According to Hamlet [1], “programs are complex, hard to understand, hard to prove, and consequently often riddled with errors.” But might not a small set of tests on a program pinpoint problems? Answering yes to this question has been the driving force behind testing, which helps reveal the discrepancies between the model being used and the real situation. Testing tries to assure that the program satisfies its specifications and it detects and prevents design and implementation faults. But testing is limited by an exponential number of states, which makes exhaustive testing very expensive and unworkable for large projects. Thus, a number of other selective testing techniques are being used. One such technique is *development testing*, which consists of a series of random tests on the software during the development stage. However, the use of mathematical techniques in developmental testing, which seems to offer good assurances and is widely used, does not ensure error-free code. Neither does refocusing verification of code to the underlying algorithm and basic computation, because not all errors may be in these areas. So, testing alone does not eliminate all the bugs.

7.1.1.2 Verification and Validation

The process of V&V involves static formal mathematical techniques such as proof-of-correctness and dynamic techniques such as testing to show consistency

between the code and the basic initial specifications. It works from the specifications of the software and develops tests that can show that software under review is faulty. Tests are randomly chosen. But as any programmer will tell you, as the level of programming gets lower and lower toward machine code, software bugs get harder and harder to detect, and no amount of V&V is able to prevent those bugs from falling through the cracks.

7.1.2 Reliability

In contrast to hardware products whose reliability is measurable from age and production quantities, software reliability cannot be measured by wear and tear, nor can it be measured by copies produced at manufacture time, although experience has shown that it exhibits some degree of stochastic properties on unpredictable input sequences. A software product can fail to deliver expected results because of an unexpected input sequence. Reliability of software can, therefore, be defined in relation to these input sequences. According to Parnas et al. [2], reliability of software is the probability that such software does not encounter an input sequence that leads to failure. A software product, therefore, is reliable if it can continue to function on numerous unpredictable input sequences. Other measures of reliability include the number of errors in the code. But this also is difficult to take as a good measure because a program with fewer errors is not necessarily more reliable than one with many. Because no system can be certified as error free, including software systems, there have been, and will continue to be, numerous cases in which systems have failed and will fail the reliability standards.

Consider the example of the Denver International Airport baggage system [3]. When the city of Denver, CO, wanted to replace Stapleton International Airport, they contracted an automated baggage company, BAE Automated Systems of Dallas, to design and build a baggage delivery system. When BAE delivered the system, it failed all initial tests. Bags flew out of carts, and jams were frequent. After a number of failed test runs, and knowing they were running out of time, city officials hired another firm, which recommended a smaller, less expensive, but working manual system to run as a stand-alone alongside the automated system. When it opened, the airport was \$2 billion over budget because of the delay caused mostly by this system.

In his book *Computer-Related Risks*, Neumann gives numerous examples of system failures caused by unreliable products [4]. Similar to standards, reliability is another very difficult concept for a buyer or customer to understand because there are no universally accepted criteria for ascertaining the reliability of a product.

7.1.3 Security

In Sect. 5.3 we discussed the general concepts of system security including information security. In this section we focus on software security. As computer

technology makes giant advances, our dependence on it increases, and so do our security concerns as more and more of the vital information that used to be secured under lock and key is now on giant computer disks scattered on numerous computer systems.

Software is an integral part of a computer system, and the security of such a system depends on its hardware but even more so on the software component. There are more security attacks on systems through software “holes” than hardware, mainly through piracy, deletion, and alteration of programs and data. Computer system software is secure if it protects its programs and data—in other words, if it does not contain trapdoors through which unauthorized intruders can access the system.

According to Neumann [5], improper encapsulation, inheritance of unnecessary privileges, and inadequate enforcement of polymorphism are the most common sources of software security flaws. Polymorphism is a state or a condition of passing through many forms or stages. Software development passes through many different forms. In addition to these as common causes of system insecurity is the human element. Computer system software can be protected from undetected modification through strong and sound design principles, enforcement of proper encapsulation, separation of all privileges, and ethical education of system developers and users about security issues.

The human and probably ethical side to system security, according to Davis [6], is that most computer crimes are not committed by hackers but by trusted employees, programmers, managers, clerks, and consultants in the company who know and can manipulate the working of the software. If Davis’ observation is true, then computer security and hence system software security greatly depends on the education of system developers and knowledgeable users.

7.1.4 Safety

Recent advances in computer technology have resulted in wider computer applications in previously unthinkable areas such as space exploration, missile and aircraft guidance systems, and life-sustaining systems. In these areas the safety of software has become one of the most prominent components of the whole security system. Such a system cannot afford an accident or an error because of software failure without dire consequences to human life, property, and the environment.

A software system is unsafe if a condition is created whereby there is a likelihood of an accident, a hazard, or a risk. The function of software safety in system safety is that software executes within a prescribed context so as not to contribute to hazards or risk either by outputting faulty values and timing or by failing to detect and respond to hardware failures that may cause a system to go into a hazardous state.

According to Leveson [7], software safety depends on the design and environment in which such software is used. Thus, software that is considered safe in one environment may be unsafe in another. Because software is designed and produced

by different people in different environments and used in different applications in a variety of environments, no one software product can conform to all requirements in all environments; in other words, one cannot assume that because a software product is hazard free in one environment it is hazard free in all environments. For example, according to Littlewood and Strigini [8], although the requirement for rate of occurrence of failures as a dependability measure is appropriate in systems that actively control potentially dangerous processes, the same measure is not as appropriate for life-critical processes in which the emphasis is on failure-free survival.

In the final analysis, good and safe software depends on good programming practice, which includes control techniques, application of various types of safety analysis during the development cycle, and evaluation of the effectiveness of these techniques. Whether these techniques are enough depends on the chosen and acceptable risk level, which tends to vary with the application environments [9]. For other dependability measures, consult Littlewood's article.

7.1.5 Quality

The emergence of a global software market, the establishment of powerful software development warehouses in different countries, and the improving standards of global software have all brought software quality to the forefront of software issues. A software product has quality if it maintains a high degree of excellence in standards, security, safety, and dependability. Many software vendors are starting to develop and apply quality improvement techniques such as total quality management (TQM).

A TQM technique that tries to improve software quality through a software development process known as the software quality function development (SQFD) represents a movement from the traditional techniques of TQM to the software development environment by focusing on improving the development process through upgrades in the requirement solicitation phase [10]. This technique focuses on this phase because software problems occur when user requirements are misunderstood, which causes overruns of development costs. Introducing design techniques that focus on user specification in this early phase leads to fewer design changes and reduces transfer errors across design phases.

7.1.6 Quality of Service

For a product, and in particular, a software product, quality of service (QoS) means providing consistent, predictable service delivery that will satisfy customer application requirements. The product must have some level of assurance that the customer's service requirements can be satisfied. For example, in the case of the Internet, QoS would mean that the network elements such as routers and hosts

expect a high level of assurance that its traffic and service requirements can be satisfied. This requirement and expectations are important because the working and the architecture of the Internet are based on the “dumb” network concept, which at its simplest involves two smart end routers, one transmitting and one receiving, and no intelligence in between. Then, datagrams with source and destination addresses traverse a network of routers independently as they move from the sender to the receiver. Internet Protocol (IP) provides only an addressing mechanism and nothing else. It provides no guarantees of the delivery of any independent datagram in the network. So, QoS is needed in network protocols.

7.2 Causes of Software Failures

Failure or poor performance of a software product can be attributed to a variety of causes, most notably human error, the nature of software itself, and the environment in which software is produced and used.

7.2.1 Human Factors

In the human factor category, poor software performance can be a result of the following:

1. *Memory lapses and attentional failures*: For example, someone was supposed to have removed or added a line of code, tested, or verified, but did not do so because of simple forgetfulness.
2. *Rush to finish*: The result of pressure, most often from management, to get the product on the market either to cut development costs or to meet a client deadline, rushing can cause problems.
3. *Overconfidence and use of nonstandard or untested algorithms*: Algorithms are put into the product line before they are fully tested by peers because they seem to have worked on a few test runs.
4. *Malice*: Software developers, like any other professionals, have malicious people in their ranks. Bugs, viruses, and worms have been known to be embedded and downloaded in software, as is the case with Trojan horse software, which boots itself at a timed location. As we will see in Sect. 7.4, malice has traditionally been used for vendetta, personal gain (especially monetary), and just irresponsible amusement. Although it is possible to safeguard against other types of human errors, it is very difficult to prevent malice.
5. *Complacency*: When either an individual or a software producer has significant experience in software development, it is easy to overlook certain testing and other error control measures in those parts of software that were tested previously in a similar or related product, forgetting that no one software product can conform to all requirements in all environments.

7.2.2 Nature of Software: Complexity

Both software professionals and nonprofessionals who use software know the differences between software programming and hardware engineering. It is in these differences that many of the causes of software failure and poor performance lie. Consider the following:

1. *Complexity*: In contrast to hardwired programming in which it is easy to exhaust the possible outcomes of a given set of input sequences, in software programming a similar program may present billions of possible outcomes on the same input sequence. Therefore, in software programming one can never be sure of all the possibilities on any given input sequence.
2. *Difficult testing*: There will never be a complete set of test programs to check software exhaustively for all bugs for a given input sequence.
3. *Ease of programming*: The fact that software programming is easy to learn encourages many people with little formal training and education in the field to start developing programs, but many are not knowledgeable about good programming practices or able to check for errors.
4. *Misunderstanding of basic design specifications*: This lack affects the subsequent design phases including coding, documenting, and testing. It also results in improper and ambiguous specifications of major components of the software and in ill-chosen and poorly defined internal program structures.

As we already discussed in Sect. 7.1.4, the environment in which a software product is produced and tested has a great bearing on its safety.

7.3 Risk

The first step in understanding the nature of software is to study the concept of risk, software risk in particular. However, before we define risk, let us define *hazard*. A hazard is a state or set of conditions of a system or an object that, together with other conditions in the environment of the system, or object, will lead inevitably to an accident [7]. According to Leveson, hazard has two components; severity and likelihood of occurrence. These two form the hazard level. Risk is a hazard level together with the likelihood of an accident to occur and the severity of the potential consequences [7]. Risk can also be defined in simpler terms as the potential or possibility of suffering harm or loss; danger, in short. Neumann defines risk as a potential problem, with causes and effects [4]. Risk can be both voluntary, with activities that we knowingly decide to undertake, or involuntary, with activities that happen to us without our prior consent or knowledge as a result of nature's actions such as lightning, fires, floods, tornados, and snowstorms. As our focus here is on the big picture of the dangers of software in particular and computer systems in general, we leave the details of the definitions at that.

How does risk function in software? Because we have defined risk as a potential problem with causes and effects, software risks, therefore, have causes and effects. Among the causes of software risks are poor software design, a mismatch of hardware software interfaces, poor support, and maintenance. Others include these [11]:

- Personnel shortfalls
- Unrealistic schedules and budgets
- Developing the wrong functions and properties
- Developing the wrong user interface
- Continuing stream of requirement changes
- Shortfalls in externally furnished components
- Shortfalls in externally performed tasks
- Real-time performance shortfalls
- Straining computer science capabilities.

Because computers are increasingly becoming a part of our lives, there are numerous ways computers, and computer software in particular, affect our lives. In many of these encounters, risk is involved. For example, computers are used in medical care and delivery, in power generation and distribution, in emergency services, and in many other facets of life. So wherever we are, be it at work, on the way to or from work, or in our own homes, where there is direct or indirect use of computer software there is always a risk that an accident can occur.

For example, there is no way for a system manager to predict how and when a system failure or attack by hackers or viruses will occur. As our world become increasingly engulfed with computer and telecommunication networks, network-related threats by hackers, viruses, system overloads, and insider misuse are increasing to such a level that the risks involved are shaping the way we work. Appropriate and effective measures are needed to manage risk. Let us look at some here.

7.3.1 Risk Assessment and Management

Risk management is a process to estimate the impact of risk. It is an approach for system managers to measure the system's assets and vulnerabilities, assessing the threat and monitoring security. For software, we consider risk management both during the design phase and during use. Risk is an important aspect of the design process. Because it is so important, two constituent components must be included: assessment and control. To implement these two components, there must be a requirement that no software project may be delivered or accepted until and unless a risk assessment or risk control evaluation has been carried out. There must be documentation of the probability and consequences of hazards and accidents to help determine what the risks are and what to do about them.

The assessment aspects in the documentation should involve a list of all the potential dangers that are likely to affect the project, the probability of occurrence and potential loss of each item, and how each item ranks among all the listed items.

The control component in the documentation should consist of the following [11]:

- Techniques and strategies to mitigate the highest ordered risks
- Implementation of the strategies to resolve the high-order risk factors
- Monitoring the effectiveness of the strategies and the changing levels of risk throughout the design process.

After the design process, when the software is in use, risk management then involves the following phases: assessment, planning, implementation, and monitoring.

7.3.1.1 Assessment

This step involves identifying the software's security vulnerabilities and may consist of a variety of techniques including question and answer, qualitative assessment, or methodology and calculation. A simple equation for calculating risk is

$$\text{Risk} = \text{Assets} \times \text{Threats} \times \text{Vulnerabilities}$$

7.3.1.2 Planning

Planning involves outlining the policies for security management.

7.3.1.3 Implementation

A good implementation may seek to match the security needs of the system with all available security tools.

7.3.1.4 Monitoring

Risk management is an ongoing process that needs constant monitoring: this helps to determine the necessary changes and new security applications to the system. The monitoring tools must be chosen based on the nature and applications of the system being protected. For example, if the system being protected is a network, the tools may include a firewall as well as intrusion detection and network forensics software.

7.3.2 Risks and Hazards in Workplace Systems

The workplace is second only to our homes in the amount of time we spend there. For most people with nine-to-five work schedules, work comprises about 40 h of

the 168-h week. When you figure in commuting to and from work and other work-related activities, we spend on the average 84 h a week at home. Because we spend so much time outside our homes and in close contact with people from all walks of life and most often work with workplace machinery and people, which we call workplace systems, there is always a high risk associated with these systems, as well as with the commute to and from work.

In a workplace environment, accidents resulting from this three-faceted model of hardware, software, and humanware are caused by the intertwining of the components whereby each part affects the others. According to Leveson [7], an accident is then a coincidence of factors related to one another through this intertwining. Each component's contribution to system accidents depends on the environment of the system. Different environments may cause different types of accident. In some accidents, software may contribute more than the other two, whereas in others, humanware may contribute more, especially in cases where there is lack of effective training of the human component. There is a perception that humanware is more prone to errors in workplace systems than either hardware or software. According to Leveson, most workplace accidents are caused by what she calls a safety culture based on humanware—a general attitude and approach to safety consisting of overconfidence, complacency, placing low priority on safety, and accepting flawed resolutions of conflicting goals. To these we also add poor employee training and poor employee morale. In workplace systems where there is a transient human component, overlooking the human component for a critical safety decision-making process may result in high-risk system safety.

This perception is enhanced by the credibility problem and the myth about computers. People still hold the computer dear, that it is more reliable and creates less risk, that software testing eliminates software errors, that increased software reliability automatically implies increased safety, and that reusing software increases its safety level. All these are myths. Software safety is as unpredictable as its counterpart, the humanware.

For those with such perceptions, there is good news. The development of intelligent computer technology and communication devices may lessen the human component in the workplace. However, this does not mean that workplace systems will be error free. It will, however, shift the burden onto software because hardware errors are more readily predictable than those by humanware and software.

Hardware errors can easily be located and fixed. Software errors, on the other hand, may take many hours before they are found, and fixing them may take even longer. Yet software systems are becoming even more complex, with complicated codes and tight delivery schedules.

7.3.3 Historic Examples of Software Risks

In the maiden days of the “Wonder Machine,” risk and vulnerability of both the computer user and data were not a problem. Software was unknown, the way we know it today, because it was embedded. Also, the computing system consisted

more of hardware than software, and projects were small. As systems became smaller and less dependent on hardware, software came out of the hardware, and projects became bigger, more complex, and more dependent on software and humanware. Then, the problems of risk and vulnerabilities set in. Ever since then, major system mishaps in hardware, software, and humanware have been recorded that have given us a glimpse of the development of computer systems and the long road that system safety, vulnerability, and risk have taken.

In his book *Computer-Related Risks* [4], Neumann, for many years the moderator of the online Internet group, “The Risk Forum,” and contributor to ACM’s “Inside Risk,” has documented a wide collection of computer mishaps that address problems in reliability, safety, security, and privacy issues in day-to-day computer activities.

Numerous other authors have written about hundreds of incidents that have made headlines in their day. We cannot list them all. But we can look at the major history-making system safety incidents, a few among many that have dotted the computing landscape.

7.3.3.1 The Therac-25

The Therac-25 is a computer-controlled electronic–accelerator radiation-therapy system developed by Atomic Energy of Canada, Ltd. (AECL). Between 1985 and 1987 the system was involved in a number of accidents, some resulting in deaths because of radiation overdose.

The machine works by creating a high-energy beam of electrons targeted to the cancerous tumor, leaving the healthy tissue surrounding the tumor unaffected. The Therac-25 was not supposed to malfunction, but like all systems, there are many possibilities for errors. Therac-25 accidents did not occur until after 6 months of use, thus creating a high degree of confidence. And when malfunctions occurred, they were very irregular, and the system successfully worked on hundreds of patients in between malfunctions. Whenever malfunctions occurred, the Therac-25 could send through the patient readings in the range of 13,000–20,000 rads instead of the normal 200 rads. Anything over 500 rads can cause death. The Therac-25 used a software upgrade of the older model of the Therac-6. The manufacturers of the Therac-25, sure of the safety record of Therac-6, paid little attention to software. They were overconfident that it worked very well. So they simply upgraded it, adding in more parameters with few changes. In addition to endangering patients, the Therac-25 also endangered operators because of the stress that resulted from the situation. For the full account of the investigation into the Therac-25 accident, the reader is referred to the paper “An Investigation of the Therac-25 Accident” by Nancy G. Leveson and Clark S. Turner (*Computer*, vol. 26, #7, July 1993, pp. 18–41).

7.3.3.2 The Space Shuttle Challenger

On January 28, 1986, the US National Aeronautical and Space Administration (NASA) flight of mission STS 51-L using the *Challenger* spaceship burst into flames 72 s after takeoff. Flight 51-L of the *Challenger* spacecraft was scheduled

originally to fly in July 1985, then it was postponed three other times until this fateful day. The accident left millions of people in shock, and it was a great setback for NASA and the prestige of the space program. The combination of these and other matters surrounding the accident, including problems within NASA, forced President Ronald Regan to appoint a commission of inquiry into the accident and the working of NASA so that similar future accidents could be avoided. The commission, chaired by William P. Rogers, former secretary of state under President Nixon (1969–1973) and attorney general under President Eisenhower (1957–1961), was expected to

- (i) [R]eview the circumstances surrounding the accident to establish the probable cause or causes of the accident; and (ii) develop recommendations for corrective or other action based upon the commission's findings and determinations.

In its deliberations, the commission interviewed more than 160 individuals, held more than 35 formal panel investigative sessions, and examined more than 6300 documents, totaling more than 122,000 pages, and hundreds of photographs.

On June 6, 1986, the commission handed their findings and recommendations to the President. In its executive summary report, the commission and other investigative agencies found that the loss of the *Challenger* was the result of a failure in the joint between the two lower segments of the right solid rocket motor. More specifically, the seals that prevent hot gases from leaking through the joint during the propellant burns of the rocket motor were destroyed, thus causing the joints to fail. Following are the commission's findings [12].

1. A combustion gas leak through the right Solid Rocket Motor aft field joint initiated at or shortly after ignition eventually weakened and/or penetrated the External Tank, initiating vehicle structural breakup and loss of the Space Shuttle Challenger during STS Mission 51-L.
2. The evidence shows that no other STS 51-L Shuttle element or the payload contributed to the causes of the right Solid Rocket Motor aft field joint combustion gas leak. Sabotage was not a factor.
3. Evidence examined in the review of Space Shuttle material, manufacturing, assembly, quality control, and processing on nonconformance reports found no flight hardware shipped to the launch site that fell outside the limits of Shuttle design specifications.
4. Launch site activities, including assembly and preparation, from receipt of the flight hardware to launch, were generally in accord with established procedures and were not considered a factor in the accident.

5. Launch site records show that the right Solid Rocket Motor segments were assembled using approved procedures. However, significant out-of-round conditions existed between the two segments joined at the right Solid Rocket Motor aft field joint (the joint that failed).
 - (a) While the assembly conditions had the potential of generating debris or damage that could cause O-ring seal failure, these were not considered factors in this accident.
 - (b) The diameters of the two Solid Rocket Motor segments had grown as a result of prior use.
 - (c) The growth resulted in a condition at time of launch wherein the maximum gap between the tang and clevis in the region of the joint's O-rings was no more than 0.008 in. and the average gap would have been 0.004 in.
 - (d) With a tang-to-clevis gap of 0.004 in., the O-ring in the joint would be compressed to the extent that it pressed against all three walls of the O-ring retaining channel.
 - (e) The lack of roundness of the segments was such that the smallest tang-to-clevis clearance occurred at the initiation of the assembly operation at positions of 120° and 300° around the circumference of the aft field joint. It is uncertain if this tight condition and the resultant greater compression of the O-rings at these points persisted to the time of launch.
6. The ambient temperature at time of launch was 36 °F, or 15° lower than the next coldest previous launch.
 - (a) The temperature at the 300° position on the right aft field joint circumference was estimated to be $28^\circ \pm 5$ °F; this was the coldest point on the joint.
 - (b) Temperature on the opposite side of the right Solid Rocket Booster facing the sun was estimated to be about 50 °F.
7. Other joints on the left and right Solid Rocket Boosters experienced similar combinations of tang-to-clevis gap clearance and temperature. It is not known whether these joints experienced distress during the flight of 51-L.
8. Experimental evidence indicates that as the result of several effects associated with the Solid Rocket Booster's ignition and combustion pressures and associated vehicle motions, the gap between the tang and the clevis will open as much as 0.017 and 0.029 in. at the secondary and primary O-rings, respectively.
 - (a) This opening begins upon ignition, reaches its maximum rate of opening at about 200–300 ms, and is essentially complete at 600 ms when the Solid Rocket Booster reaches its operating pressure.
 - (b) The External Tank and right Solid Rocket Booster are connected by several struts, including one at 310° near the aft field joint that failed. The effect of this strut on the joint dynamics is to enhance the opening of the gap between the tang and clevis by about 10–20% in the region of 300°–320°.

9. O-ring resiliency is directly related to its temperature.
 - (a) A warm O-ring that has been compressed will return to its original shape much more quickly than will a cold O-ring when compression is relieved. Thus, a warm O-ring will follow the opening of the tang-to-clevis gap, whereas a cold O-ring may not.
 - (b) A compressed O-ring at 75 °F is five times more responsive in returning to its uncompressed shape than is a cold O-ring at 30 °F.
 - (c) As a result it is probable that the O-rings in the right solid booster aft field joint were not following the opening of the gap between the tang and clevis at time of ignition.
10. Experiments indicate that the primary mechanism that actuates O-ring sealing is the application of gas pressure to the upstream (high-pressure) side of the O-ring as it sits in its groove or channel.
 - (a) For this pressure actuation to work most effectively, a space between the O-ring and its upstream channel wall should exist during pressurization.
 - (b) A tang-to-clevis gap of 0.004 in., as probably existed in the failed joint, would have initially compressed the O-ring to the degree that no clearance existed between the O-ring and its upstream channel wall and the other two surfaces of the channel.
 - (c) At the cold launch temperature experienced, the O-ring would be very slow in returning to its normal rounded shape. It would not follow the opening of the tang-to-clevis gap. It would remain in its compressed position in the O-ring channel and not provide a space between itself and the upstream channel wall. Thus, it is probable the O-ring would not be pressure actuated to seal the gap in time to preclude joint failure caused by blow-by and erosion from hot combustion gases.
11. The sealing characteristics of the Solid Rocket Booster O-rings are enhanced by timely application of motor pressure.
 - (a) Ideally, motor pressure should be applied to actuate the O-ring and seal the joint before significant opening of the tang-to-clevis gap (100–200 ms after motor ignition).
 - (b) Experimental evidence indicates that temperature, humidity, and other variables in the putty compound used to seal the joint can delay pressure application to the joint by 500 ms or more.
 - (c) This delay in pressure could be a factor in initial joint failure.
12. Of 21 launches with ambient temperatures of 61 °F or greater, only 4 showed signs of O-ring thermal distress, that is, erosion or blow-by and soot. Each of the launches below 61 °F resulted in one or more O-rings showing signs of thermal distress.
 - (a) Of these improper joint sealing actions, one-half occurred in the aft field joints, 20% in the center field joints, and 30% in the upper field joints. The division between left and right Solid Rocket Boosters was roughly equal.
 - (b) Each instance of thermal O-ring distress was accompanied by a leak path in the insulating putty. The leak path connects the rocket's combustion

- chamber with the O-ring region of the tang and clevis. Joints that actuated without incident may also have had these leak paths.
- 13. There is a possibility that there was water in the clevis of the STS 51-L joints because water was found in the STS-9 joints during a destack operation after exposure to less rainfall than STS 51-L. At time of launch, it was cold enough that water present in the joint would freeze. Tests show that ice in the joint can inhibit proper secondary seal performance.
 - 14. A series of puffs of smoke were observed emanating from the 51-L aft field joint area of the right Solid Rocket Booster between 0.678 and 2.500 s after ignition of the Shuttle Solid Rocket Motors.
 - (a) The puffs appeared at a frequency of about three puffs per second. This rate roughly matches the natural structural frequency of the solids at liftoff and is reflected in slight cyclic changes of the tang-to-clevis gap opening.
 - (b) The puffs were seen to be moving upward along the surface of the booster above the aft field joint.
 - (c) The smoke was estimated to originate at a circumferential position of between 270° and 315° on the booster aft field joint, emerging from the top of the joint.
 - 15. This smoke from the aft field joint at Shuttle liftoff was the first sign of the failure of the Solid Rocket Booster O-ring seals on STS 51-L.
 - 16. The leak was again clearly evident as a flame at approximately 58 s into the flight. It is possible that the leak was continuous but unobservable or nonexistent in portions of the intervening period. It is possible in either case that thrust vectoring and normal vehicle response to wind shear as well as planned maneuvers reinitiated or magnified the leakage from a degraded seal in the period preceding the observed flames. The estimated position of the flame, centered at a point 307° around the circumference of the aft field joint, was confirmed by the recovery of two fragments of the right Solid Rocket Booster.
 - (a) A small leak could have been present that may have grown to breach the joint in flame at a time on the order of 58–60 s after liftoff.
 - (b) Alternatively, the O-ring gap could have been resealed by deposition of a fragile buildup of aluminum oxide and other combustion debris. This resealed section of the joint could have been disturbed by thrust vectoring, Space Shuttle motion, and flight loads induced by changing winds aloft.
 - (c) The winds aloft caused control actions in the time interval of 32–62 s into the flight that were typical of the largest values experienced on previous missions.

In conclusion, the commission stressed that the *Challenger* accident was the result of failure of the pressure seals in the aft field joint of the right Solid Rocket Booster. The commission also concluded that the failure, therefore, was a result of a faulty design unacceptably sensitive to a number of factors that include temperature, physical dimensions, character of materials, the effects of reusability, processing, and the reaction of the joint to dynamic loading.

During the commission's hearing, information emerged indicating that engineers at Morton Thiokol, Inc., the Utah company that designed the Rocket Booster joints in the *Challenger*, warned management against the launch of the space shuttle because of the predicted low temperatures. They feared that the predicted low temperatures would stiffen the O-rings.

Against their company's guidelines to give "yes" or "no" answers to the commission's questions, three engineers, Allan McDonald, Arnold Thompson, and Roger Boisjoly, broke ranks with management to reveal the warning. The three, led by Roger Boisjoly, told the commission that they warned management that the temperature of 18 °F (-8°C) predicted the morning of the launch may make the booster O-ring stiff, preventing them from sealing the gases properly. They presented evidence to the commission to show that at 53 °F, in one of the past launches, one of the two redundant joints had not sealed. It was learned that although Morton Thiokol's management had not previously approved any launch at temperatures below 53 °F, on this occasion, management changed their position under duress from NASA, after previously postponing the *Challenger* launch four times. NASA argued that there were never any such data on the booster joint acceptable range of temperatures and they were, therefore, ready to go. Up to the last moment of launch, engineer Allen McDonald, the Morton Thiokol resident engineer at the Kennedy Space Flight Center, fought NASA to postpone the launch, but he did not succeed and the launch went ahead—at least for 27 s [13].

7.3.3.3 The Indian Bhopal Chemical Accident

The Union Carbide industrial accident in Bhopal, India, illustrates many of the elements of this safety culture. In December 1984, an accidental release of methyl isocyanate killed between 2000 and 3000 people and injured tens of thousands of others, many of them permanently. The accident was later blamed on human error. The official report stated that water was let into the storage tank of methyl isocyanate through an improperly cleaned pipe [7, p. 40]. According to Leveson, Union Carbide management, including scientists, believed that because of the modern technology they had at the plant, such an accident could not happen there. It did.

7.3.3.4 The Chernobyl Nuclear Power Accident

The 1986 Chernobyl nuclear power accident in northern Ukraine, then a republic of the USSR, was the worst nuclear accident that has ever occurred. For a number of days after the accident, the Soviet government kept the world guessing at what was happening. But when details started to spill out, it was discovered that things started going bad on April 26, 1986, when during an experiment to determine the length of time the turbine and the generator could supply the emergency cooling system with electricity if an accident were to occur, the experiment went haywire and the operators started to notice a decline in the power output.

On noticing the decline, the operators turned off two automatic systems that were supposed to activate the controller rods in an emergency. At the same time they pumped more water into the reactor tank. When the water in the reactor tank

stopped boiling, they then decreased the freshwater flow into the reactor tank—a bad mistake.

This action resulted in an unprecedented power upsurge in a very short time when the water in the reactor tank started to boil again. This overwhelming power, generated in only a couple of seconds, overheated the nuclear fuel, and a third of the core exploded from the inside. The quick upsurge in power and the subsequent explosion resulted from the fact that the steam from the boiling reactor tank water reacted with the graphite in the reactor and formed carbon dioxide and hydrogen, generating high steam pressure that lifted the lid off the reactor tank and quickly reacted with the air outside to cause the huge explosion. Immediately after, radioactive emissions were blown by the wind and quickly covered the surrounding areas and threatened western Europe [14].

7.4 Consumer Protection

Asset purchasing is a game of wits played between the buyer and the seller. Any time you make a purchase, remember that you are starting at a disadvantage because unlike the seller you do not have all the cards to win the game; the seller does. He or she always has more information about the item for sale than you, the buyer. As the game progresses the seller picks and chooses the information to give to the buyer.

In the case of software purchases, the buyer needs to be even more careful because many software products do not always work as the seller claims they do, or at least as the buyer would like them to do. Software products may not work as expected because the buyer has unrealistic expectations about the product, the environment in which the product is supposed to work is inadequate, the seller exaggerated the capacities of the software, or the software is simply faulty. So what can buyers do if the product just purchased does not live up to expectations? It usually depends on how much buyers know about their rights. Without claiming to be lawyers, let us begin this section by defining the legal jargon buyers need to press for their rights and to take legal action if nothing else works. Legal action should be the last resort, however, because once filed, a lawsuit takes on a life of its own in expense, time, and outcome.

7.4.1 Buyers' Rights

What are our rights as purchasers of a software product that does not live up to our expectations? The first step is to review the available options by contacting the developer of the product. If the developer is not the seller, then start with the vendor from whom you bought the product. Sometimes the vendor or seller may replace the product with a new one, depending on the elapsed time and warranties, or may refer you to the developer.

When talking to the developer, explain specifically and clearly what it is that you want, why you are not satisfied with the product, and what you want to accomplish. Although developers claim to help unsatisfied customers, computer software is more difficult to handle once it has been opened, and you may have to do more than you would with a different kind of product to convince both the vendor and the developer that their product is not satisfactory. Developers typically have technical teams to help customers with problems, and most of the problems are solved at this level. However, if you are still not satisfied with the service, other options are open to you, such as the following:

- *Product replacement:* You may demand a product replacement if you think it will solve the problem. Most developers usually do replace faulty products.
- *Product update:* When the product is found to have a fault that the provider was not aware of at the time of shipping the product to market, the producer may fix the fault by providing a patch or an upgrade of the product that can either be downloaded or shipped to all customers who report that fault (e.g., the Netscape case and the Intel Pentium chip debacle).

In the Netscape case, a serious flaw in the then just released Netscape Communications Corporation's browser was uncovered by a small Danish software company called Cabocomm. The bug made it possible for web site operators to read anything stored on the hard disk of a PC logged on the web site. Netscape acknowledged the error and offered to send upgrades to its customers [15].

The Intel Pentium chip situation was very much like that of Netscape, Inc., except that for Intel it was a hardware problem. A mathematics professor using a Pentium-based office PC found that at a high level of mathematical computation, the chip froze. He reported his discovery via e-mail to a colleague, and the word spread like wildfire. But unlike Netscape, Inc., which immediately admitted fault, Intel did not at first admit it until the giant IBM and other small PC companies threatened not to use the chip in their line of products. Intel then accepted responsibility and promised to send upgrades to all its customers [16].

If none of the options already discussed proves viable, the next and probably last step is legal action. A liability suit is filed in civil court against the producer of the product for damages. In some cases, if the product has resulted in a casualty, a criminal suit against the producer can also be filed if the customer believes there was criminal intent. If you decide to file a civil liability suit, two avenues are open to you—the contract and/or tort options (see Sects. 7.4.3 and 7.4.4). For a successful outcome in a software case, you need to proceed with care to classify what was purchased either as a product or a service. The decision of the courts in a purchase lawsuit depends heavily on this classification.

7.4.2 Classification of Computer Software

As we explained earlier, computer software falls into three categories: product, service, and a mixture of both service and product.

7.4.2.1 What Is a Product?

A product must have two fundamental properties. First it must have a tangible form, and second it must have a built-in intrinsic value for the buyer. Look at a few examples. Consider a lottery ticket you have just bought for which the jackpot is \$100 million. Suppose you paid \$1 for your ticket. The ticket has a form that everyone can see and touch; it is tangible. Now suppose your grandmother, who is visiting with you, lives in a state with no lottery. She finds your ticket lying on the kitchen table. To her, although the ticket has a tangible form because she can see and touch it, it has no value to her. To you, however, the ticket not only has tangible form, it also has an intrinsic value worth millions of dollars.

For a second example, suppose you have a splitting headache, which your physician tells you can be cured by a certain tablet. For you this tablet has a tangible form because you can see and touch it, but it also has an intrinsic value because you believe it will cure your headache. To somebody else who does not have a headache, the tablet simply has a tangible form but no value beyond that. For software to be considered a product, it must have both a tangible form and an intrinsic value. Many software packages have these two properties and can therefore be considered as products. For example, when you buy a U.S. tax preparation package and you live in the United States, to you the package has both a tangible form and an intrinsic value. But to somebody else living in another country, the package, although it has tangible form, does not have any intrinsic value at all.

7.4.2.2 What Is a Service?

A service, in contrast to a product, has intrinsic value, but it does not have a tangible form. Because it has no tangible form, whoever wants a service must describe it. A service most often involves a provider-client or provider-customer relationship: The provider in this equation is the person offering the service and the client is the person receiving the service. For professionals, the relationship is always provider-client, where there is an imbalance of power in favor of the provider (e.g., an attorney-client relationship or a doctor-patient relationship).

In nonprofessional situations, it is often a provider-customer relationship and the power play here is in favor of the customer because customers must always get what they want, and the customer must always be satisfied. What the provider and customer receive in this relationship, however, has no tangible form—but it does have intrinsic value. The customer gets satisfaction with the service and this satisfaction is the value of the service. The provider in turn gets paid, and again that is the value of the service.

7.4.2.3 Is Software a Product, a Service, or a Mixture?

Now that we have differentiated between a product and a service, let us tackle the problem of classification of computer software. According to Johnson [17], courts do not always equate products with what is tangible. If we accept this line of reasoning, then software with no tangible form can be considered a product and therefore can be protected by patent laws that protect products. But we have to be very careful with this line of argument not to conclude too hastily that software is going to be accepted as a product because there are items with no tangible forms that even the courts cannot accept as products.

If we define software as a set of instructions describing an algorithm to perform a task that has intrinsic value for the buyer, this definition classifies software as a service. For example, suppose you want a software item to perform a certain task for your company but you cannot find the appropriate type on the market. You describe what it is that you want done to a software developer, and the software developer comes up with what you want. What has been produced can be considered a service performed; it has intrinsic value for you, but no tangible form.

But suppose you want a program to do a task for you, and instead of describing it to a software developer, you decide to go to your discount store where you know such software item is sold, and you buy a box containing that item. What you have just paid for is a product no different from that box of potato chips you picked up at the same store. Suppose further that when you open your potato chips you find them crushed to a powder, or suppose when you eat the potato chips they make you sick, and later you find they were contaminated. Legally you cannot sue the producer of the chips for malpractice or negligence, but you can sue for product liability. Similarly, then, you cannot sue for malpractice or negligence if the contents of a software box do not work properly. You should sue for product liability because in this case software seems to be an indisputable product.

There are interesting yet mixed views concerning the classification of software. Prince [18] defines three categories of software classes using the producer–market–customer relationship:

1. *Canned software*: Off-the-shelf software for a general market customer. Tax preparation software packages fall in this category.
2. *Customized software*: The software produced for a customer after the customer has described what he or she specifically needs to the software producer.
3. *Hybrid software*: Canned software that is customized to meet certain customer needs, but cannot be used to perform the whole task without focused modifications.

Prince argues for a product classification of all software falling in category 1 on the basis of three principles:

1. The product was placed in mainstream commerce by the producer with the purpose of making a profit, and the producer therefore should be responsible for the implied safety of the product.

2. The producer has a better understanding of the product than the buyer, and therefore is in a better position to anticipate the risks.
3. The producer can easily spread the burden of the cost of product liabilities from injuries over the entire range of product customers without the customers knowing, thus minimizing the costs and risks to him or her [17].

With software in category 1, strict liability for a bad product, including negligence, can be raised by the customer seeking benefits from the producer. For software in category 2, customers can seek benefits from the producer resulting from injuries by using malpractice laws because software in this category is a service.

With these two categories, the distinction between a product and a service is straightforward. But this is not the case with software in category 3. Some elements in this category belong to a product classification (e.g., placing the item in the mainstream of commerce in a tangible form).

Also, because the software can be changed to suit individual needs, the principle that the producer can spread the burden of the cost of product liability because of injuries over all product customers does not apply anymore. This is what Johnson calls a mixed classification case because it belongs in two categories: the canned category and the customized category. Johnson suggests that such software should be treated in the following way. If there is an error in the canned part, then it can be treated like a product. And if it develops an error in the customized part, then it should be handled as a service. The problem with this line of thinking, however, is that for an average software user it is almost impossible to tell in which part the error originated.

As technology advances, new categories will definitely emerge, and ideally new laws will be enacted to cover all these new categories. We cannot keep on relying on old laws to cope with the ever-changing technological scene.

When you have successfully classified your software, then you can pursue the two possible options open to you: contract or tort.

7.4.3 The Contract Option

Lawyers define a contract as a binding relationship between two or more parties. A contract need not be in a physical form like a document; it can be oral or implied. For a relationship to be a contract, it must satisfy several requirements including mutual consent. Mutual consent is a meeting of the minds on issues such as the price bargained or agreed upon, the amount paid or promised to be paid, and any agreement enforceable by law.

In contract laws, a producer/developer can be sued for breach of contract. Contract laws also cover express and implied warranties, third-party beneficial contracts, and disclaimers. Warranties are guarantees that the product or service will live up to its reasonable expectations. Some warranties are not specifically written down but are implied, whereas others are merely expressed either orally or in some other form.

7.4.3.1 Express Warranties

Express warranties are an affirmation of a fact, a promise, or a description of goods, a sample, or a model made by the seller to the buyer relating to the goods and as a basis for payment negotiations. Express warranties are entered into between the customer and the producer when a producer agrees to supply the product to the customer. They also involve promises made by the producer through sales representatives and written materials on packaging attesting to the quality of the product and guidelines buyers must follow to get detectable errors corrected by the producer. These warranties are also included in the U.S. Uniform Commercial Code (UCC) and, unless specifically excluded by the seller, express warranties become enforceable immediately upon application of the UCC transaction.

Producers usually limit their liability on products by stipulating a timeframe on warranties and contracts. But in most cases, time limits do not apply, especially in cases of express warranties because of advertising and description of the product capacity on or in packages [19].

7.4.3.2 Implied Warranties

Implied warranties are enforced by law according to established and accepted public policy. For example, in the nonideal world we live in, we cannot expect a contract to contain everything the buyer and producer may want. Remember that the process of buying and selling is a game in which there is a winner, a loser, or a draw. In this game, as we pointed out earlier, the seller has more cards than the buyer.

On the buyer's side are quite a number of things they do not have to negotiate because they do not know as much and need time to learn the product. The law protects buyers so they do not have to negotiate every small detail of the product conditions. Implied warranties make such conditions always part of the package agreement even if they are not specifically written down in the contract. An implied warranty guarantees that a product is of average quality and will perform no less than similar products and that it is fit for the intended use. For buyers to benefit from implied warranties, proof must be given that the contract did not exclude some features and there is no time limitation for reporting defects; some companies, however, stipulate a timeframe in which defects must be reported. Implied warranties are advantageous to buyers because they enforce a degree of discipline on the producers and vendors to sell standard products for the intended purposes. They are also useful to the producer and vendors because they create a degree of confidence and trust in buyers, hence increasing sales of products. However, there is a downside to implied warranties; they tend to make software expensive because the producer anticipates the cost of the lawsuits that might be generated and passes such costs on to the customers.

7.4.3.3 Third-Party Beneficiary Contracts

If a software product injures a user other than the buyer, under a third-party beneficiary contract the user may sue the producer for benefits from injuries or loss of

income resulting from the product. Third-party beneficiary contracts suits are not common because they are rarely found valid in courts.

7.4.3.4 Disclaimers

Producers try to control their liability losses by putting limits on warranties via disclaimers. Through disclaimers, producers preempt lawsuits from buyers by telling buyers in writing on the contracts the limits of what is guaranteed.

Many users see disclaimers as a way producers try to avoid responsibility. Producers see them as a way of informing the users of the risks before they buy the product, and they also like them because they put the burden of proof and risk taking squarely on the buyers: *caveat emptor* (the buyer beware), so to speak. Whether these disclaimers are recognized in courts depends on a number of factors including the belief that the disclaimers were made in good faith.

7.4.3.5 Breach of Contract

A contract entered into between two or more parties and not performed as promised by either party can be considered breached by the party not in compliance. If the complaint is not very serious, the breach may not cause the termination of the contract, but the breaching party may be asked to pay some damages. However, if the breach is considered serious by one of the parties, it may cause the termination of the contract. In this case the offended party may demand damages from the breaching party in the contract upon satisfactory proof that there were indeed damages resulting from contract breaching.

7.4.4 The Tort Option

If a buyer cannot seek benefits from the producer through contracts laws, another avenue of legal action is through tort. A tort is a wrong committed upon a person or property in the absence of a contract. A tort may include negligence, malpractice, strict liability, and misrepresentation. Torts fall into two categories: intentional and unintentional. For example, if you are passing by a construction site and somebody pours concrete on you, this act may be interpreted as intentional if the worker who poured the concrete knew it was you passing; otherwise, it is unintentional.

7.4.4.1 Negligence

Negligence can be used by the buyer to obtain benefits from the producer if there is provable evidence that the product lacked a certain degree of care, skill, and competence in the workmanship. Carelessness and a lack of competence may be proved from the design stage through the testing, installation, and user training stages of the product. For example, suppose that the buyer of a computer software product is a large hospital and the product is life-sustaining software. If it causes injury to a patient because the hospital personnel using the software were not adequately trained by the producer of the software, and this can be proved beyond a reasonable doubt, then the producer can be sued for negligence. In other

words, negligence in this case is holding the software producer party liable for the injuries he or she did not intend and even tried to avoid while making the software. Negligence cases apply mainly to services rendered.

7.4.4.2 Malpractice

Malpractice is a type of negligence. It is also applicable in cases involving services. For example, if you visit the doctor for a simple eye surgery and he or she cuts off your ear, you can sue the doctor for malpractice. Malpractice lawsuits are common in professional services. In the case of software, if it is taken as a service, then malpractice applies.

7.4.4.3 Strict Liability

Strict liability is a tort involving products. Any product sold in a defective condition that ends up endangering a person creates a case of strict liability to the seller of such a product even if the buyer did not make a direct purchase from the seller. In strict liability lawsuits, the burden of proof of negligence is shifted to the producer, and the costs incurred by defects in the product are squarely in the hands of the producer. Under strict liability, it is the product itself that is on trial. The product is examined and if it is found to be defective and dangerous, the buyer is awarded benefits. Strict liability laws are harsh. They ignore efforts made by the producer of the product to make the product safe for the reason that the producer was in a better position to know the risks [20].

7.4.4.4 Misrepresentation

Most computer software and other computer products are no longer sold by their original producers and developers but by third-party sellers and vendors. Examples of such third-party sellers are mail-order computer hardware and hundreds of software companies and many big brand name computer outlets. So, very few original manufacturers and software developers are selling directly to the buyer. To the buyer of a computer product, this situation may add another layer of bureaucracy and more problems. The problems are not one sided, however; indirect selling also causes additional problems for the producer. Many of the producer problems added by this layer actually come from misrepresentation of the product. Misrepresentation may be intentionally done by the sales representative to induce the buyer to buy the product or it may be just a genuine mistake. Consider car manufacturers, for example. Usually they buy back faulty new cars from customers when these cars have developed specific problems within a designated period of time. These cars are usually repaired and sent back to the dealers to be sold, not as new but as used products. Sometimes, however, car dealers sell these cars as new cars. Whether car manufacturers are aware of these sales practices or not, customers always end up suing the car manufacturers.

Before you sue the producer, however, determine first whether it was an intentional misrepresentation called fraudulent misrepresentation. To prove fraudulent misrepresentation, you need to prove that the vendor was aware the facts given were not true or that the vendor would have known the true facts but opted not

to inform the buyer accordingly. You also need to show, and be believed, that you as a buyer relied on that information to buy the product. And finally you need to show that the product resulted in damage. If you can establish all these facts and be believed by the courts, then you have a case [17].

7.5 Improving Software Quality

The problem of software quality cannot be solved by courts alone. Software producers must themselves do more to ensure software quality and hence safety.

7.5.1 Techniques for Improving Software Quality

Reputable software standards, reliability of software, and software safety depend greatly on the quality of the software. If the quality is low, software is prone to errors, is therefore not reliable, and hence has poor standards. In Sect. 7.1.1, we stated that software can be enhanced by techniques such as developmental testing, V&V, and programming standards. But the quality of software cannot be assumed by looking at these factors only. According to Linger et al. [20], software cannot be made reliable by testing alone. Software quality can be improved through the following innovative new review techniques:

- *Formal review*: Presentation of the software product by a person more familiar with the product to others with competent knowledge of that product so they can critique the product and offer informed suggestions.
- *Inspection*: Involves checking the known specific errors from past products and establishing additional facilities that may be missing in the product to bring the product up to acceptable standards.
- *Walk-through*: Requires code inspection line by line by a team of reviewers to detect potential errors. Each review session is followed by a discussion of the findings by the members of the review team, usually with the creators of the code present.
- *Phased inspection*: Technique developed by Knight and Mayers [21]. It is an enhanced method combining the previous three methods by putting emphasis on the limitations of those methods. It consists of a series of coordinated partial inspections called phases during which specific properties of the product are inspected.

If care is taken by the software developer to improve the development process of software by improving validation, verification, and the survivability of the software, the liability on their part will be minimized, and software safety will be greatly improved. If software developers paid more attention to software quality using many of the techniques cited here during development, there would be little need to discuss consumer protection.

7.6 Producer Protection

In Sect. 7.4 we outlined user rights and protection mechanisms in the case of sub-standard software. In this section we focus on the other side of the same coin: the software producer's rights and protection mechanisms. Software producers need to protect themselves against piracy, illegal copying, and fraudulent lawsuits. But because of the high costs, time, and the unpredictability of the outcome of lawsuits, it is not good business practice for software producers to sue a single person making a copy of the software. It only makes sense to go after big-time and large-scale illegal copying. Software producers should be prepared to seek protection from the courts to protect the software itself from illegal copying, piracy, and also from lawsuits from customers. In addition, producers should be prepared to protect themselves from lawsuits filed by consumers. For this kind of protection, producers are advised to use the courts as much as possible and ask for advice from lawyers and business colleagues. There is no one single magic bullet approach.

Exercises

1. Discuss the difficulties faced by software producers.
2. Discuss ways software customers can protect themselves from substandard software products.
3. Discuss how the following are used to protect a software customer:
 - Implied warranty
 - Express warranty
 - Disclaimer
 - Strict liability
4. It has been said that software problems are a direct result of its differences from hardware. Discuss.
5. Discuss the elements of software quality. How can software quality be a solution to many of software's liability problems?
6. Do safe software systems imply reliable systems? Why or why not?
7. Software reliability, especially in critical systems, is vital. Discuss the necessary conditions for reliability.
8. With the development of scanning and snooping software, computer systems cannot be assured of security. Discuss what steps need to be taken to ensure system safety.
9. We discussed in this chapter that risk can be both voluntary and involuntary. Give examples in each case.
10. Why do we always take risk as a crucial design component? What two aspects must be considered?
11. In carrying out the risk assessment and control of a software product, what aspects must be considered and why?
12. Why is there a myth that computers (hardware and software) do not make errors?
13. Does the myth about computers complicate the safety issue?

14. How does humanware affect system safety?
15. If workplace systems were all automated, could this eliminate workplace system risks? Would it reduce it?
16. Why is software safety so difficult to attain? Can it be guaranteed?

References

1. R. Hamlet, Special section on software testing. *Commun. ACM* **31**(6), 662–667 (1988)
2. D. Parnas, J. van Schouwen, S. Kwan, Evolution of safety-critical software. *Commun. ACM* **33**(6), 636–648 (1990)
3. J. Taylor, America's loneliest airport: Denver's dreams can't fly. *Kansas City Star*, 25 Aug 1994. NewsBank, Transportation, fiche 43, grids D12–14 (1994)
4. P. Neumann, *Computer-Related Risks* (ACM Press, New York, 1995)
5. P. Neumann, The role of software engineering. *Commun. ACM* **36**(5), 114 (1993)
6. A. Davis, Employee computer crime on the rise. *Creative Computing*, June 1985, p. 6
7. N. Leveson, *Safeware: System Safety and Computers* (Addison-Wesley, Reading, 1995)
8. B. Littlewood, L. Strigini, Validation of ultrahigh dependability for software-based systems. *Commun. ACM* **36**(11), 69–80 (1993)
9. D. Ritchie, Reflections on trusting trust. *Commun. ACM* **27**(8), 761–763 (1984)
10. S. Haag, M.K. Raju, L.L. Schkade, Quality function deployment usage in software development. *Commun. ACM* **39**(1), 41–49 (1996)
11. B.W. Boehm, *Software Risk Management: Principles and Practices* (IEEE Computer Society Press, New York, 1989)
12. President's commission on the challenger accident report. <https://science.ksc.nasa.gov/shuttle/missions/51-l/docs/rogers-commission/table-of-contents.html>
13. K. Fitzgerald, Whistle-blowing: not always a losing game. *IEEE Spectr.* **26**(6), 49–52 (1990)
14. Chernobyl Accident 1986. <http://www.world-nuclear.org/information-library/safety-and-security/safety-of-plants/chernobyl-accident.aspx>
15. S. Young, Netscape bug uncovered. CNNfn, 12 June 1997
16. Computer stock tumble over chip flow. *New York Times*, 4 Dec 1994, section D
17. D. Johnson, *Computer Ethics*, 2nd edn. (Prentice Hall, Englewood Cliffs, 1994), p. 134
18. J. Prince, Negligence: liability for defective software. *Okla Law Rev.* **33**, 848–855 (1980)
19. F. Neitzke, *A Software Law Primer* (Reinhold, New York, 1984)
20. C. Linger, H.D. Mills, B. Witts, *Structured Programming: Theory and Practice* (Addison-Wesley, Reading, 1979)
21. J. Knight, A. Mayers, An improved inspection technique. *Commun. ACM* **36**(11), 51–61 (1994)

Further Reading

22. R. Bunker, S. Datar, C. Kemerer, D. Zeneig, Software complexity and maintenance costs. *Commun. ACM* **36**(11), 81–94 (1993)
23. J. Fetzer, Program verification: the very idea. *Commun. ACM* **31**(9), 1048–1063 (1988)
24. D. Gelperin, B. Hetzel, The growth of software testing. *Commun. ACM* **31**(6), 687–690 (1988)
25. R. Grady, Practical results from measuring software quality. *Commun. ACM* **36**(11), 50–61 (1993)
26. J.-C. Laprie, B. Littlewood, Probabilistic assessment of safety-critical software: why and how? *Commun. ACM* **35**(2), 13–21 (1992)
27. N. Leveson, Software safety in embedded computer systems. *Commun. ACM* **34**(2), 34–46 (1991)



Computer Crimes

8

Abstract

Computer Crimes surveys the history and examples of computer crimes, their types, costs to society, and strategies of detection and prevention. In the discussion, it is noted that a great number of computer attacks fall into two categories: penetration and denial of service attacks. And these are discussed in depth. Attack motives are also discussed. Are nations, businesses, and individuals prepared for computer attacks? Are they ready to pay the price? We look for answers to these questions as we ponder the costs and consequences of computer crimes. We note also that although it is difficult to estimate the actual costs of e-attacks on physical system resources, progress is being made for better and more accurate estimates. An in-depth discussion of the social and ethical consequences that include psychological effects, moral decay, loss of privacy and loss of trust follows. We end the chapter with recommendations for educating the computing device users in computer ethics. The need to educate the user to be aware of possible sources of computer crimes and what to do if and when one becomes a victim of these crimes is stressed. It is noted that education can go a long way in reducing computer crimes if the users take crime preventive steps every time they use the computer and computer related technologies.

Learning Objectives

After reading this chapter, the reader should be able to

1. Describe trends in computer crimes and protection against viruses and denial-of-service attacks.
2. Understand techniques to combat “cracker” attacks.
3. Understand the history of computer crimes.
4. Describe several different cyber-attacker approaches and motivations.

5. Identify the professional's role in security and the tradeoffs involved.
6. Develop measures to be taken both by individuals themselves and by organizations (including government) to prevent identity theft.

Scenario 7: All in the Open, My Friend—Be Watchful for You Will Never Know the Hour!

Josephine Katu owns a company that manufactures women's cosmetics. She has loyal and dedicated employees, and each one of them works very hard. Josephine has been good to them too. She compliments them and rewards them handsomely when the occasion presents itself.

However, Josephine has become suspicious of some of the employees, without knowing which one(s) in particular. She is also not sure what it is that is not right, but she suspects something is going wrong somewhere in her company and she is not happy. So she decides to do something about it.

During the Christmas season, Josephine buys each of her 20 or so employees a laptop for his or her home and she promises to pay for their online expenses. In addition, she also promises to take care of all system maintenance, using the company technician, if they ever need it. Josephine writes a virus that she occasionally and selectively uploads to her employees' computer, which uploads the content of the machine.

The plan is working very well and Josephine is getting plenty of information whenever the virus is released. She is even planning on bringing in the press and the FBI.

Discussion Questions

1. Is Josephine right to release a virus on her employees' computers?
2. Do the computers belong to her or to her employees?
3. Are the employees' rights being violated? What rights?
4. What are the social and ethical implications of Josephine's little tricks?

8.1 Introduction

It is difficult to define a computer crime without getting tangled up in the legal terminology. We will try to make it simple for the rest of us nonlawyers. A computer crime is a crime like any other crime, except that in this case the illegal act must involve a computer system either as an object of a crime, an instrument used to commit a crime, or a repository of evidence related to a crime. With the Internet, the scope of computer crimes has widened to actually include crimes that would normally be associated with telecommunication facilities. Because of this, we want to expand our definition of a computer crime to be an illegal act that involves a computer system or computer-related system such as any mobile

device, microwave, satellite, or other telecommunication system that connects one or more computers or computer-related systems.

Acts using computers or computer-related technologies that fall within the limits that the legislature of a state or a nation has specified are considered illegal and may lead to forfeiture of certain civil rights of the perpetrator. In the United States, local, state, and federal legislatures have defined such acts to include such as the following:

- Intrusions into Public Packet Networks
- Network integrity violations
- Privacy violations
- Industrial or financial espionage
- Pirated computer software
- Computer-aided fraud
- Internet/e-mail abuse
- Using computers or computer technology to commit murder, terrorism, pornography, hacking, and many other crimes.

Computer crimes target computer resources for a variety of reasons [1]:

- Hardware such as computers, printers, scanners, servers, and communication media
- Software that includes application and special programs, system backups, diagnostic programs, and system programs such as operating systems and protocols
- Data in storage, transition, or undergoing modification

An attack on any one of these resources is considered a computer or computer-related attack. Some of these resources are more vulnerable than others and are, therefore, targeted more frequently by attackers. Most computer crimes on the resources just listed fall into the three categories following. Our focus in this chapter is on the last category [1, 2]:

- Human blunders, errors, and omissions that are usually caused by unintentional human actions. Unintended human actions are usually the result of design problems. Such attacks are called *malfunctions*. Malfunctions, although occurring more frequently than natural disasters, are as unpredictable as natural disasters.
- Intentional threats that originate from humans caused by illegal or criminal acts from either insiders or outsiders, recreational hackers, and criminals. For the remainder of this chapter, we are focusing on this last category.

8.2 History of Computer Crimes

As we look at the history of computer crimes, we focus on two aspects of such crimes: viruses and hacking. These two have been the source of almost all computer crimes. Sometimes they become one when hackers use viruses to attack computer systems, as we discuss next. The term virus is derived from the Latin word virus that means poison [3]. Until recently, the term had remained mostly in medical circles, meaning a foreign agent injecting itself into a living body, feeding on it to grow, multiply, and spread. Meanwhile, the body weakens and loses its ability to fight foreign invaders and eventually succumbs to the virus if not treated.

In contrast to a biological virus, however, a computer virus is a self-propagating computer program designed to alter or destroy a computer system's resources. Like its cousin, it follows almost the same pattern when attacking computer software. It attaches itself to software, grows, reproduces many times, and spreads in the new environment. It spreads by attacking major system resources including data and sometimes hardware, weakening the capacity of these resources to perform the needed functions and eventually bringing the system down.

The word virus was first assigned a nonbiological meaning in the 1972 science fiction stories about the G.O.D. machine, which were compiled in a book *When Harlie Was One* by David Gerrold. Later, Fred Cohen, then a graduate student at the University of Southern California, associated the term with a real-world computer program he wrote for a class demonstration [4]. During the demonstration, each virus obtained full control of the system within an hour. That simple class experiment has led to a global phenomenon that has caused nightmares for system administrators, security personnel, and cyberspace users.

Hacking, as a computer attack technique, utilizes the Internet working between computers and communication devices. So long as computers are not interconnected in a network, hacking cannot take place. So, the history of hacking begins with the invention of the telephone in 1876 by Alexander Graham Bell, which has made internetworking possible. However, there was a long gap between the invention of the telephone and the first recorded hacking activity in 1971 when John Draper, commonly known as "Captain Crunch," discovered that a toy whistle from a cereal box could produce the precise tone of 2600 Hz needed to make free long-distance phone calls [5]. With this act, "Phreaking," a cousin of hacking, entered our language. With the starting of a limited national computer network by ARPANET, in the 1970s, a limited form of system break-in from outsiders started appearing. The movie "War Games," which appeared in 1983, glamorized and popularized hacking. It is believed by many that the movie gave rise to the hacking phenomenon.

The first notable system penetration attack actually started in the mid-1980s with the San Francisco-based 414-Club. The 414-Club was the first national news-making hacker group. The group named their group 414 after the area code of San Francisco where they were. They started a series of computer intrusion attacks via a Stanford University computer which they used to spread the attack across the

country [6]. From that small but history-making attack, other headline-making attacks from Australia, Germany, Argentina, and the United States followed.

In the United States, these activities, although at a low level, started worrying law enforcement agencies so much so that in 1984 the Comprehensive Crime Control Act was enacted, giving the Secret Service jurisdiction over computer fraud. Also at around this time, the hacker movement was starting to get active. In 1984, *2600: The Hacker Quarterly*, a hacker magazine, was launched and the following year the electronic hacking magazine *Phrack* was founded. As the Internet grew, hacker activities increased greatly. Then, in 1986, the U.S. Congress passed the Computer Fraud and Abuse Act. Hacker activities that had only been in the United States started to spread worldwide. In 1987, the Italian hacker community launched *Decoder* magazine, similar to the United States' *2600: Hacker Quarterly* [5].

The first headline-making hacking incident involving a virus took place in 1988 when a Cornell graduate student created a computer virus that crashed 6000 computers and effectively shut down the Internet for 2 days [6]. Robert Morris's action forced the U.S. government to form the federal Computer Emergency Response Team to investigate similar and related attacks on the nation's computer networks. Law enforcement agencies started to actively follow the comings and goings of networks traffic and sometimes eavesdrop on the communications. This move did not sit well with some activists, who in 1990 formed the Electronic Frontier Foundation to defend the rights of those investigated for alleged computer hacking.

The 1990s saw heightened hacking activities and serious computer network “near” meltdowns, including the 1991 expectation of the “Michelangelo” virus that was expected to crash computers on March 6, 1992, the artist’s 517th birthday, but which passed without incident. In 1995, the notorious, self-styled hacker Kevin Mitnick was first arrested by the FBI on charges of computer fraud that involved the stealing of thousands of credit card numbers. In the second half of the 1990s, hacking activities increased considerably, including the 1998 Solar Sunrise, a series of attacks targeting Pentagon computers that led the Pentagon to establish round-the-clock, online guard duty at major military computer sites. Also, there was a coordinated attack on Pentagon computers by Ehud Tenebaum, an Israeli teenager known as “The Analyzer,” and an American teen. The close of the twentieth century saw heightened anxiety in both the computing and computer user communities about both the millennium (Y2K) bug and the ever-rising rate of computer network break-ins. So, in 1999 President Bill Clinton announced a \$1.46 billion initiative to improve government computer security. The plan intended to establish a network of intrusion detection monitors for certain federal agencies and encourage the private sector to do the same [5]. The year 2000 probably went down in history as one of the years that saw the most costly and most powerful computer network attacks: it included “Mel-lisa,” “Love Bug,” “Killer Resume,” and a number of devastating distributed denial-of-service attacks. The following year, 2001, the elusive “Code Red” virus was released. The future of viruses is as unpredictable as the types of viruses themselves.

The period between 1980 and 2001 saw sharp growth in reported incidents of computer attacks. Two factors have contributed to this phenomenal growth: the growth of the Internet and the massive news coverage of virus incidents.

8.3 Types of Computer Systems Attacks

A great number of computer system crimes are actually computer attacks. Major computer attacks fall into two categories: penetration and denial-of-service attacks.

8.3.1 Penetration

A penetration attack involves breaking into a computer system using known security vulnerabilities to gain access to a cyberspace resource. With full penetration, an intruder has full access to all that system's resources. Full penetration, therefore, allows an intruder to alter data files, change data, plant viruses, or install damaging Trojan Horse programs into the system. It is also possible for intruders—especially if the victim computer is on a network—to use it as a launching pad to attack other network resources. Penetration attacks can be local, wherein the intruder gains access to a computer on a LAN on which the program is run, or global on a WAN such as the Internet, where an attack can originate thousands of miles from the victim computer. Penetration attacks originate from many sources, including the following:

- (i) *Insider Threat.* For a long time, penetration attacks were limited to inhouse employee-generated attacks to systems and theft of company property. In fact, disgruntled insiders are a major source of computer crimes because they do not need a great deal of knowledge about the victim computer system. In many cases, such insiders use the system every day, which allows them to gain unrestricted access to the computer system, thus causing damage to the system and/or data. The 1999 Computer Security Institute/FBI report notes that 55% of respondents reported malicious activity by insiders [7].
- (ii) *Hackers.* Since the mid-1980s, computer network hacking has been on the rise, mostly because of the wider use of the Internet. Hackers penetrate a computer system for a number of reasons, as we discuss in the next section, including the thrill of the challenge, bragging rights in the hacker community, and illicit financial gain or other malicious purposes. To penetrate the system, hackers use a variety of techniques. Using the skills they have, they download attack scripts and protocols from the Internet and launch them against victim sites.
- (iii) *Criminal Groups.* Although a number of penetration attacks come from insiders and hackers with youthful intents, there are a number of attacks that originate from criminal groups, for example, the “Phonemasters,” a widespread international group of criminals who in February 1999 penetrated

the computer systems of MCI, Sprint, AT&T, Equifax, and even the FBI's National Crime Information Center. A member of the group in the United States, Calvin Cantrell downloaded thousands of Sprint calling card numbers. He later sold the numbers to a Canadian. From Canada, the numbers found their way back to America and on to Switzerland and eventually ended up in the hands of organized crime groups in Italy [7].

- (iv) *Hactivism.* Demonstrations have taken place in Seattle, Washington DC, Prague, and Genoa by people with all sorts of causes, underlining the new phenomenon of activism that is being fueled by the Internet. This activism has not only been for good causes, but it has also resulted in what has been dubbed *hactivism*—motivated attacks on computer systems, usually web pages or e-mail servers of selected institutions or groups by activists. A group with a cause overloads e-mail servers and hacks into web sites with messages for their causes. The attacks so far have not been harmful, but they still cause damage to services. Such groups and attacks have included the “Electronic Disturbance Theater,” which promotes civil disobedience online in support of the Zapatista movement in Mexico; supporters of Serbia, during the NATO bombing of Yugoslavia; electronically “ping”-attacked NATO web servers; and supporters of Kevin Mitnick, the famed computer hacker who while in federal prison, hacked into the Senate web page and defaced it [7].

8.3.2 Denial of Service

Denial-of-service attacks, commonly known as distributed denial of service (DDoS) attacks, are a new form of computer attacks. They are directed at computers connected to the Internet. They are not penetration attacks and, therefore, they do not change, alter, destroy, or modify system resources. However, they affect the system by diminishing the system’s ability to function; hence, they are capable of bringing a system down without destroying its resources. They first appeared widely in the summer of 1999. The year 2000 saw this type of computer attack become a major new category of attack on the Internet. Headlines were made when a Canadian teen attacked Internet heavyweights Amazon, Ebay, E*Trade, and news leader CNN.

Differing from penetration attacks, DDoS attacks typically aim to exhaust the network bandwidth, its router processing capacity, or network stack resources, thus eventually breaking the network connectivity to the victims; this is achieved by the perpetrator breaking into weakly secured computers. The victim computers are found by using freely available scan software on the Internet that pinpoints to well-known defects in standard network service protocols and common weak configurations in operating systems. Once the victims have been identified, the perpetrator breaks in and may perform additional steps that include the installation of software, known in the industry as a “rootkit,” to conceal the break-in trail and make the tracing of subsequent activities impossible.

When the perpetrator has several victim machines under its control, the controlled machines are then used to mount attacks on other machines in the network, usually selected machines, by sending streams of packets, as projectiles, to the secondary line of victims. For some variants of attacks like the Smurf attack (which is discussed shortly), the packets are aimed at other networks, where they provoke multiple echoes all aimed at the victim.

Similar to penetration electronic attacks (e-attacks), DDoS attacks can also be either local, where they can shut down LAN computers, or global, originating thousands of miles away on the Internet, as was the case in the Canadian-generated DDoS attacks. Attacks in this category include, among others, IP-spoofing, SYN-Flooding, Smurfing, Buffer Overflow, and Sequence Number Sniffing.

8.4 Motives of Computer Crimes

Hacking has many dubious motives. More recently, however, we have seen more cases of hacking for illicit financial gain or other malicious purposes. It is difficult to exclusively discuss all the motives, but let us look at the following major categories [2]:

- (i) *Political Activism.* There are many causes that lead to political activism, but all these causes are grouped under one burner—hactivism—as discussed in Sect. 8.3.1.
- (ii) *Vendetta.* Most vendetta attacks are for mundane reasons such as a promotion denied, a boyfriend or girlfriend taken, an ex-spouse given child custody, and other situations that may involve family and intimacy issues.
- (iii) *Joke/Hoax.* Hoaxes are warnings that are actually scare alerts started by one or more malicious persons, and are passed on by innocent users who think that they are helping the community by spreading the warning. Most hoaxes are viruses although there are hoaxes that are computer-related folklore and urban legends. Virus hoaxes are often false reports about nonexistent viruses that cause panic, especially to the majority of users who do not know how viruses work. Some hoaxes can become extremely widespread as they are mistakenly distributed by individuals and companies with the best of intentions. Although many virus hoaxes are false scares, there are some that may have some truth about them, but which often become greatly exaggerated, such as the “Good Times” and the “Great Salmon.” Virus hoaxes infect mailing lists, bulletin boards, and Usenet newsgroups. Worried system administrators sometimes contribute to this scare by posting dire warnings to their employees, which become hoaxes themselves.

(iv) *The Hacker's Ethics.* This is a collection of motives that make up the hacker character. According to Steven Levy, hackers have motivation and ethics and beliefs that they live by, and he lists six, as below [8]:

- Free access to computers and other ICT resources—and anything that might teach you something about the way the world works—should be unlimited and total.
- All information should be free.
- Mistrust authority; promote decentralization.
- Hackers should be judged by their hacking, not bogus criteria such as degrees, age, race, or position.
- You can create art and beauty on a computer.
- Computers can change your life for the better.

If any of these beliefs is violated, a hacker will have a motive.

(v) *Terrorism/Extortion.* Our increasing dependence on computers and computer communication has opened up a can of worms we now know as electronic terrorism. Electronic terrorism by individuals is targeting enterprise systems, institutions, and governments. But cyber-terrorism is not only about obtaining information; it is also about instilling fear and doubt and compromising the integrity of the data, which leads to extortion. In many countries, financial institutions, such as banks, brokerage firms, and other large corporations, have paid large sums of extortion money to sophisticated international cyber terrorists.

(vi) *Political and Military Espionage.* For generations, countries have been competing for supremacy of one form or another. During the Cold War, countries competed for military dominance. At the end of the Cold War, the espionage turf changed from military to gaining access to highly classified commercial information that would not only let them know what other countries are doing but might also give them either a military or commercial advantage without spending a lot of money on the effort. It is not surprising, therefore, that the spread of the Internet has given a boost and a new lease of life to a dying Cold War profession. Our high dependency on computers in the national military and commercial establishments has given espionage a new fertile ground. Electronic espionage has many advantages over its old-fashioned, trench-coated, sunglasses-wearing, and gloved Hitchcock-style cousin.

(vii) *Business and Industrial Espionage.* As businesses become global and world markets become one global bazaar, business competition for ideas and market strategies has become very intense. Economic and industrial espionage is on the rise around the world as businesses and countries try to outdo the other in the global arena. As countries and businesses try to position themselves and be a part of the impending global cutthroat competition, economic and industrial espionage is beginning to be taken seriously by company executives. The Internet has created fertile ground for cyber-sleuthing, and

corporate computer attacks are the most used business espionage technique. It usually involves physical system penetration for trophies such as company policy, as well as management and marketing data. It may also involve sniffing, electronic surveillance of company executive electronic communications, and company employee chat rooms for information.

- (viii) *Hate.* The growth of computer and telecommunication technology has unfortunately created a boom in all types of hate. There is growing concern about a growing rate of acts of violence and intimidation motivated by prejudice based on race, religion, sexual orientation, or ethnicity. Hate is being given a new and very effective and global forum.
- (ix) *Personal Gain/Fame/Fun.* Personal gain motives are always driven by the selfishness of individuals who are not satisfied with what they have and are always wanting more, mostly financially.

8.5 Costs and Social Consequences

Are nations, businesses, and individuals prepared for computer attacks? Are they ready to pay the price? The answers to both these questions at the moment are probably no. It is not that we are not aware of it. It is not that we do not talk about it. And it is not that it has not happened before. It has. In fact, there have been heated and sometimes furious debates about it. There have been newspaper reports, and television and congressional discussions about the United States' preparedness for a national electronic attack. Yet not enough is being done beyond discussions. Because there are not enough data collection and analysis by U.S. intelligence agencies, or business and financial communities that would have provided lead information, assessment, and preparedness of the nation for an electronic attack on the national information infrastructure, a good credible policy cannot be formulated. In fact, during 1996 Congressional hearings on "Intelligence and Security in Cyberspace," a senior member of the intelligence community in charge of collecting such data compared the efforts in place at the time to a "toddler soccer game where everyone just runs around trying to kick the ball somewhere" [9]. We have come a long way since that time. Now both the U.S. Congress and the President are committed to protecting the nation's cyber infrastructure and are making resources available for this purpose.

This problem is not limited to the United States only; country after country around the globe is facing similar problems. Very few countries, if any, have assessed and analyzed any information on their information infrastructure, on how an electronic attack can affect not only their national security but also other essential infrastructures such as businesses, power grids, and financial and public institutions. There are various reasons for this lack of information [2]:

- In nearly all countries there is no required reporting mechanism in government agencies, even the private sector, to detect intrusions and report such intrusions.

- In the private sector, there is very little interest in the reporting of any system-related intrusions, a result of the fear of marketplace forces that would expose the management's weaknesses to the shareholder community and competitors.
- The insider effect. Various reports point to a blank picture about the effects of insider intruders on the overall detection and reporting of electronic attacks or e-attacks. It is reported in some studies that a majority of all e-attacks are generated and started by inside employees, which makes the job of detection and reporting very murky. It is like having an arsonist working in the fire department.
- Many nations have no required and trained security agencies to fight e-attacks.

The danger is real. The ability to unleash harm and terrorize millions of people, thus causing widespread panic, is possessed by many. The arena to play the game is global, and there is no one who can claim a monopoly on such attacks. In the United States, and probably in other countries, most attacks originating from outside the country are directed, for the moment, toward military and commercial infrastructures, for obvious reasons. Although most reporting of attacks seem to come from government and public sources, there is a similar rate of attempt and probably success in the private sector. The good news is that private industry is beginning to become a partner with the public sector in reporting.

The universality of cyber attacks creates a new dimension to cyberspace security. In fact, it makes it very difficult to predict the source of the next big attack, let alone identify trouble spots, track and apprehend hackers, and put a price on the problem that is increasingly becoming a nightmare to computer systems administrators, the network community, and users in general.

Every survey of computer crime and computer attacks indicates a rising trend. There are several reasons to which we can attribute this rather strange growth of cybercrimes [2].

- (i) *Rapid technology growth.* The unprecedented growth and merging of both the computer and telecommunication industries has enabled access to the Internet to balloon into billions of users. The growing wireless technology and mobile devices have made Internet access easier because people can now log on to the Internet anytime, anywhere. But this easy access has also made hiding places plentiful. From Alaska's snowcaps to the Sahara desert to the Amazon and Congo forests, cyber access is as good as in London, New York, or Tokyo, and the arena of possible cyber attacks is growing.
- (ii) *Easy availability of hacker tools.* There are an estimated 30,000 hacker-oriented sites on the Internet advertising and giving away free hacker tools and hacking tips [9]. As the Manila-generated "Love Bug" demonstrated, hacking prowess is no longer a question of affluence and intelligence but of time and patience. With time, one can go through a good number of hacker sites, picking tips and tools, and come out with a ready payload to create mayhem in cyberspace.

- (iii) *Anonymity*. The days when computer access was only available in busy, well-lit, public and private areas are gone. Now as computers become smaller and people with these small Internet-able gizmos become more mobile, hacker tracing, and apprehension have become even more difficult.
- (iv) *Cut-and-paste programming technology*. This phase has removed the most important impediment that prevented many would-be hackers from trying the trade. Historically, before anybody could develop a virus, one had to write a code for it. The code had to be written in a computer programming language, compiled, and made ready to go. This means, of course, that the hacker had to know or learn a programming language! Learning a programming language is known to be more than a 1-day job. It takes long hours of studying and practicing. Well, today this is no longer the case. We're in an age of *cut-and-paste programming*. The pieces and technical know-how are readily available from hacker sites. One only needs to have a motive and the time.
- (v) *Communications speed*. With the latest developments in bandwidth, high volumes of data can be moved in the shortest time possible. Thus, intruders can download the payload, usually developed by cut-and-paste offline, very quickly log off, and possibly leave before detection is possible.
- (vi) *High degree of internetworking*. Global networks are becoming more and more connected in every country. Nearly all these networks are connected on the Internet. In many countries, with readily available and cheap Internet-able mobile devices, Internet access is available.
- (vii) *Increasing dependency on computers*. The ever increasing access to cyberspace, together with increasing capacity to store huge quantities of data, increasing bandwidth in communication networks to move huge quantities of data, increased computing power of computers, and plummeting prices on computer equipment have all created an environment of human dependency on computers. This, in turn, has created fertile ground for hackers.

8.5.1 Lack of Cost Estimate Model for Cyberspace Attacks

As the prices of computers and Internet-able mobile devices plummet and Internet accessibility becomes global, cyber attacks are likely to skyrocket. Cost estimating cyber attacks in this changing environment is becoming increasingly very difficult. Even in a good environment, estimates of cyber attack crimes are difficult. The efforts to develop a good cost model is hindered by a number of problems, including the following [2]:

- (i) It is very difficult to quantify the actual number of attacks. Only a tiny fraction of what everyone believes is a huge number of incidents are detected, and even a far smaller percentage of that is reported. In fact, as we noted in the previous section, only one in 20% of all system intrusions is detected, and of those detected only one in 20% is reported [10].

- (ii) Even with these small numbers reported, there has been no conclusive study to establish a valid figure that can at least give us an idea of what it is that with which we must cope. The only few known studies have been regional and sector based. For example, there have been studies in education, on defense, and in a selected number of industries and public government departments.
- (iii) According to Terry Guiditis, of Global Integrity, 90% of all computer attacks both reported and unreported are perpetrated by insiders [11]. Insider attacks are rarely reported even if they are detected. As we reported in Chap. 9, companies are reluctant to report any type of cyber attacks, especially insider ones, for fear of diluting integrity and eroding investor confidence in the company.
- (iv) Lack of cooperation between emergency and computer crime reporting centers worldwide. There are many such centers worldwide, but they do not cooperate with one another because most are in commercial competition [11].
- (v) Unpredictable types of attacks and viruses. Attackers can pick and choose when and where to attack. Also, the types of attacks and topography used in attacks cannot be predicted. Because of these factors, it is extremely difficult for system security chiefs to prepare for attacks and, therefore, reduce the costs of each attack, if it occurs.
- (vi) Virus mutation is also another issue in the rising costs of cyber attacks. The recent “Love Bug” and “Code Red” e-mail attacks are examples of a mutating virus. In each incident the viruses started mutating within a few hours after release. Such viruses put enormous strain on systems administrators to search and destroy all the various strains of the virus.
- (vii) There are not enough trained system administrators and security chiefs in the latest network forensics technology who can quickly scan, spot, and remove or prevent any pending or reported attack and quickly detect system intrusions. When there is a lack of trained and knowledgeable personnel, it takes longer to respond when an attack occurs, and to clear the system from such an attack in the shortest period of time possible, thus reducing the costs. Also, failure to detect intrusion always results in huge losses to the organization.
- (viii) Primitive monitoring technology. The computer industry as a whole, and the network community in particular, has not achieved the degree of sophistication that would monitor a computer system continuously for foolproof detection and prevention of system penetration. The industry is always on the defensive, always responding *after* an attack has occurred and with inadequate measures. In fact, at least for the time being, it looks as if the attackers are setting the agenda for the rest of us. This kind of situation makes every attack very expensive.

For organizations, the costs of a data breach resulting from a cyber attack are not only alarming but are rising on an annual basis. According to the Ponemon Institute [11], the institute that annually estimates the U.S. Cost of a Data Breach, an average data breach resulting from a cyber attack in 2010 was \$7.2 million, or \$214 per customer record; this was a \$10 per-record jump from 2009. The Institute also estimates, that in the year, incidences in which companies experienced breaches for the first time resulted in average costs of a whopping \$326 per record in 2010, again up from \$228 the prior year [12]. In 2012, the Institute reported the previous year's (2011) estimate costs to be up 56% on last year's figures (2010), with an average cost of \$5.9 M per year, ranging from \$1.5 million to \$36.5 million per year [13].

If anything, these figures, though worrisome, indicate a growing trend with no end in sight.

8.5.2 Social and Ethical Consequences

Although it is difficult to estimate the actual costs of e-attacks on physical system resources, we are making progress toward better estimates. What we cannot now do, and probably will never be able to do, is to put a cost estimate on e-attacks on individual members of society. This task is difficult because of the following reasons [2]:

- (i) *Psychological effects.* These effects depend on the attack motive and may result in long-lasting psychological effects such as hate. Psychological effects may lead to individual reclusion and increasing isolation. Such trends may lead to dangerous and costly repercussions on the individual, corporations, and the society as a whole.
- (ii) *Moral decay.* There is a moral imperative in all our actions. When human actions, whether bad or good, become so frequent, they create a level of familiarity that leads to their acceptance as “normal.” This type of acceptance of actions formerly viewed as immoral and bad by society is moral decay. There are numerous e-attacks that can cause moral decay. In fact, because of the recent spree of DDoS, and e-mail attacks, one wonders whether people performing these acts seriously consider them as immoral and illegal anymore!
- (iii) *Loss of privacy.* After headline-making e-attacks that wreaked havoc on global computer systems, there is a resurgence in the need for quick solutions to the problem that seems to have hit home. Many businesses are responding with patches, filters, intrusion detection (ID) tools, and a whole list of other “solutions.” These solutions are a direct attack on individual privacy. This type of privacy invasion in the name of network security is a threat to all of us whose price we will never estimate and we are not ready to pay! The blanket branding of every Internet user as a potential computer attacker or a

criminal,, until proven otherwise, is perhaps the greatest challenge to personal freedom yet encountered by the world's societies.

- (iv) *Trust.* Along with the loss of privacy, trust is lost. Individuals once attacked lose trust in a person, group, company, or anything else believed to be the source of the attack or believed to be unable to stop the attack. Together with draconian solutions, e-attacks cause us to lose trust in individuals, and businesses, especially businesses either hit by e-attacks or trying to stop attacks forcibly. Such customer loss of trust in a business is disastrous for that business. Most importantly, it is a loss of the society's innocence.
-

8.6 Computer Crime Prevention Strategies

Preventing computer crime is not a simple thing to do because to do that one needs to understand how these crimes are committed and who is involved in these crimes. To prevent such crimes, therefore, we need to focus on three entities in the game: the computer as a tool used to commit the crimes, the criminal who is the source of the crime, and the innocent victim of the crime. Our approach to prevention will, therefore, involve strategies from all three.

8.6.1 Protecting Your Computer

For better protection of your computer consider the following measures based on a list by the San Diego Police Department [14]. Similar measures can be found at many police departments in many countries.

8.6.1.1 Physical Protective Measures

Install surface locks, cable-locking devices, and fiberoptic loops to prevent equipment theft.

- Locate the computer and data storage away from outside windows and walls to prevent damage from external events.
- Install strong doors and locks to the computer room to prevent equipment theft and tampering.
- Reinforce interior walls to prevent break-ins. Extend interior walls to the true ceiling.
- Restrict access to computer facilities to authorized personnel. Require personnel to wear distinct, color-coded security badges in the computer center. Allow access through a single entrance. Other doors should be alarmed and used only as emergency exits.

8.6.1.2 Procedural and Operational Protective Measures

If you take the computer as the main tool in the execution of the crime, this leads us to find those elements of the computer that are more susceptible to being the good conduit. The list of these items may include data, software, media, services, and hardware.

Using this list, analyze the dangers to each item on the list. Buy and install protective software based on the value of each item on the list.

Classify information into categories based on importance and confidentiality. Use labels such as “Confidential” and “Sensitive.” Identify software, programs, and data files that need special access controls.

Install software access control mechanisms. Require a unique, verifiable form of identification, such as a user code or secret password for each user. Install special access controls, such as a call-back procedure, if you allow access through a dial-telephone line connection.

Encrypt confidential data stored in computers or transmitted over communication networks. Use National Institute of Standards and Technology (NIST) data encryption standards.

Design audit trails into your computer applications. Log all access to computer resources with unique user identification. Separate the duties of systems programmers, application programmers, and computer programmers.

Establish procedures for recovering your operating system if it is destroyed. Store all backup data offsite.

Review automated audit information and control reports to determine if there have been repeated, unsuccessful attempts to log on both from within and outside your facility. Look for unauthorized changes to programs and data files periodically.

8.6.1.3 Anti-virus Protection

The following measures can help protect your computer from viruses:

- Do not bring disks in from outside sources.
- Scan demo disks from vendors, shareware, or freeware sources for viruses.
- Restrict use of electronic bulletin boards.
- Scan downloaded files for viruses. Avoid downloading executable files.
- Make regular backups to aid in recovery.

8.6.2 The Computer Criminal

There are two measures that I would consider as appropriate for the computer criminal.

8.6.2.1 Pass Computer Crime Prevention Laws

Local and national governments should pass laws directed toward computer crimes including computer tampering, computer fraud, and other computer crimes, so that if a person commits a computer crime offense, when knowingly and without the authorization of a computer's owner, or in excess of the authority granted that person, if found guilty, should serve a court sentence consonant to the extent of his or her crime. An increasing number of countries now either have such laws or are in the process of enacting them.

8.6.2.2 Enforcement of Criminal Laws

We cannot fight computer crimes, whether or not we have laws on the books, unless those laws can be enforced. Thus, one way of reducing computer crime is to aggressively enforce computer crime laws with just but stiff sentences that send a message to would-be criminals that they will pay the price if they perpetuate computer crimes.

8.6.2.3 Moral Education

Throughout this book, I have been advocating for computer ethics education. There is a need for computer ethics education that includes an ethical framework which may make the would-be criminal reflect on the pending act. Computer ethics education, just like all types of education, is a long-time investment especially in the youth, not only to build their character but also to guide their actions throughout their lives.

8.6.3 The Innocent Victim

The following measures should be focused on the victims of computer crimes [11].

8.6.3.1 Personnel Policies

Monitor activities of employees who handle sensitive or confidential data. Watch for employees who work abnormally long hours, or who refuse to take time off. Many computer crime schemes require regular, periodic manipulation to avoid detection. Be aware of employees who collect material not necessary to their jobs, such as programming manuals, printouts for data, programs, and software manuals.

Change security password codes to block further access by employees who leave or are fired. The latter become a high risk to your company for revenge or theft.

Establish rules for computer use by employees, including removal of disks or printed output. All employees should sign and date a printed copy of these rules to indicate that he/she understands them.

8.6.3.2 Educating the Computer User

Just as we did with the computer criminal, we need to educate the user to be aware of possible sources of computer crime and what to do if and when one becomes a victim of a computer crime. This education can go a long way in reducing computer crimes if the users take crime preventive steps every time they use the computer and when owning a computer.

Exercises

1. List five types of computer attacks.
2. In a short essay, discuss the differences between a denial-of-service attack and a penetration attack.
3. Which attack type is more dangerous to a computer system, a penetration attack or a denial-of-service attack?
4. List and briefly discuss five attack motives.
5. Why do hackers devote substantial amount of time to their trade?
6. Discuss the challenges in tracking down cyber criminals.
7. Why is it so difficult to estimate the costs of business e-crimes both nationally and globally?
8. What is the best way to bring about full reporting of e-crimes, including costs?
9. Why do countries worldwide have very little information to help them combat cyber crimes?
10. Why are cyber crimes on the rise?
11. In addition to monetary costs, there are ethical and social costs of e-crimes; discuss these “hidden” costs.
12. Estimate the cost of cyber attacks in the past 2 years.

References

1. 42 U.S. Code § 5195c—*Critical Infrastructures Protection*. Legal Information Institute. Cornell Law School. <https://www.law.cornell.edu/uscode/text/42/5195c>
2. J.M. Kizza, *Computer Network Security and Cyber Ethics*, 3rd edn (McFarland, Jefferson, NC, 2011)
3. K. Forchet, *Computer Security Management* (Boyd & Fraser, Danvers, 1994)
4. Timeline of Computer Security Hacker History. Wikipedia. https://en.wikipedia.org/wiki/Timeline_of_computer_security_hacker_history
5. P.J. Denning, *Computers Under Attack: Intruders, Worms and Viruses* (ACM Press, New York, 1990)
6. L.J. Freeh, *FBI Congressional Report on Cybercrime*. <https://www.fbi.gov/@@search?SearchableText=congressional+report+on+cybercrime&pageSize=20&page=3>
7. S. Levy, *Hackers: Heroes of the Computer Revolution* (Anchor Press/Doubleday, Garden City, 1984)
8. Security in Cyberspace: U.S. Senate Permanent Subcommittee on Investigations, 5 June 1996
9. J. Christensen, *Bracing for Guerilla Warfare in Cyberspace*. CNN Interactive, 6 Apr 1999

10. D.S. Alberts, *Information Warfare and Deterrence—Appendix D: Defensive War: Problem Formation and Solution Approach*. <http://www.iwar.org.uk/iwar/resources/deterrence/iwdAppd.htm>
11. Insider Threat Surveys, Reports, Incidents, Damages. National Insider Threat Special Interest Group (NITSIG). <http://www.nationalinsiderthreatsig.org/nitsig-insiderthreatreportsurveys.html>
12. E. Johnson, *The Real Cost of Cyber Attacks*. <https://www.theatlantic.com/search/?q=The+real+cost+of+cyber+attacks/>
13. 2010 Annual Study: U.S. Cost of a Data Breach Compliance Pressures, Cyber Attacks Targeting Sensitive Data Drive Leading IT Organizations to Respond Quickly and Pay More. http://www.symantec.com/content/en/us/about/media/pdfs/symantec_ponemon_data_breach_costs_report.pdf?om_ext_cid=biz_socmed_twitter_facebook_marketwire_linkedin_2011Mar_worldwide_costofdatabreach
14. A. Warwick, *The Cost of Cyber Attacks is up 56 %, Study Reveals*. <https://www.computerweekly.com/news/2240105258/The-cost-of-cyber-attacks-is-up-56-study-reveals>

Further Readings

15. K.E. Anderson, *Criminal Threats to Business on the Internet: A White Paper*, Global Technology Research, Inc., 23 June 1997. This is a discussion of the increasing trend of criminal activity against information systems, from the low-level, amateur intruder to organized crime, and industrial and international espionage
16. A. Chaturvedi et al., *Fighting the Wily Hacker: Modeling Information Security Issues for Online Financial Institutions Using the SEAS Environment*. INET JAPAN 2000 Conference, 18 July 2000. The paper discusses proposed methods to analyze the online risks faced by the financial industry
17. Computer Security Institute/Federal Bureau of Investigation, Annual Cost of Computer Crime Rise Alarmingly: Organizations Report \$136 Million in Losses, Press Release, Computer Security Institute, 4 Mar 1998. This is a summary of the 1998 survey on computer crime
18. Counterintelligence Office of the Defense Investigative Service, Industry CI Trends, OASDPA/ 96-S-1287, 26 Dec 1996. This paper discusses threats and techniques used for low-level intelligence collecting by foreign companies and governments against U.S. DoD contractors
19. General Accounting Office (GAO), GAO Executive Report—B-266140, Report to the Committee on Governmental Affairs, U.S. Senate, 22 May 1996. This gives a detailed report on attacks to U.S. Department of Defense computer systems with recommendations for improved security
20. M. Kapor, *Civil Liberties in Cyberspace: When Does Hacking Turn from an Exercise of Civil Liberties into Crime?* Scientific American, September 1991. This is a discussion of the various legal, social, and privacy-related issues within computer networks using the U.S. Secret Service's raid on Steve Jackson Games as a case study
21. National Counterintelligence Center, Annual Report to Congress on Foreign Economic Collection and Industrial Espionage, Annual Report to the U.S. Congress, 1998. This is a summary of the espionage threat to the United States Specific highlights include the interest in information security and information warfare
22. R.E. Overill, *Computer Crime—An Historical Survey*, Defence Systems International, 1998. This paper discusses the historical development of computer crime. United Nations, International Review of Criminal Policy—United Nations Manual on the Prevention and Control of Computer-Related Crime, International Review of Criminal Policy, No. 43 and 44, 1994. These are extensive documents reviewing all aspects of international computer crime
23. U.S. Department of Justice. News Release, 1 Mar 1998. Israel Citizen Arrested in Israel for Hacking U.S. and Israel Government Computers. <http://www.usdoj.gov/opa/pr/1998/march/125.htm.html>

24. Section A: The Nature and Definition of Critical Infrastructure. <http://www.nipc.gov/nipcfaq.htm>
25. A. Grosso, The economic espionage ACT: touring the minefields. Commun. ACM **43**(8), 15–18 (2000)
26. F. Grampp, R. Morris, Unix operating system security. AT&T Bell Laboratories Tech. J. **63**(8), 1649, Part 2 (October 1984)
27. P.G. Neumann, Risks of insiders. Commun. ACM **42**(12), 160 (1999)
28. J.M. Kizza, *Civilizing the Internet: Global Concerns and Efforts Toward Regulation* (McFarland, Jefferson, NC, 1999)
29. Computer Attacks: What They Are and How to Defend Against Them, ITL Bulletins, May 1999. <http://www.nist.gov/itl/lab/bulletins/may99.html>



Cyberbullying, Cyberstalking and Cyber Harassment

9

Abstract

The rapid growth of the internet, together with ever plummeting prices and increasing miniaturization of digital devices popularized by easy mobility and fast access to online services have all contributed to the creation of an exciting seemingly unlimited virtual environment in which the concept of *presence* has been transformed to mean *virtual presence*. Virtual presence, while great as an entertainment environment, is complex as it anchors anonymity, thus making it extremely dangerous. It is being misused and abused as more and more people of all shades get easy access to cyberspace. In fact the combination of these two build and boost individual confidence that sometimes boards to the realms of insanity. Because with these two, individuals in these virtual environments may be tempted to become reckless, knowingly or otherwise. This is one of the key causes and perpetrators of cyberbullying. In this chapter, we look at the dangers of virtual presence and anonymity, dangers associated with them, how to remedy the impact, tools and best practices for protection.

Learning Objectives

After reading this chapter, the reader should be able to:

- Understand circumstances surrounding cyberbullying, cyberstalking and cyber Harassment
- Understand the legal definition of cyberbullying, cyberstalking and cyber Harassment
- Describe the different types of cyberbullying, cyberstalking and cyber Harassment
- Learn about the evolution of cyberbullying, cyberstalking and cyber Harassment in tandem with the evolution of online social media
- Learn the evolving legislation landscape of cyberbullying, cyberstalking and cyber Harassment

- Recognize the difficulties and effects of cyberbullying, cyberstalking and cyber Harassment
- Be able to identify and recognize the victims of cyberbullying, cyberstalking and cyber Harassment
- Acquire the techniques and skills of dealing with cyberbullying, cyberstalking and cyber Harassment
- Recognize the difficulties of dealing with cyberbullying, cyberstalking and cyber Harassment

9.1 Definitions

In Chap. 11, we will fully define and discuss the evolution of online social networks. But before then, and to be able to fully discuss and fully understand cyberbullying, cyberstalking and cyber Harassment and their effects on society, we are going to briefly define and give a brief ‘expose’ of the evolution of online social networks. As we will see in Chap. 11, a social network is a theoretical mesh network, where each node is an individual, a group or organization who independently generates, captures and disseminates information and also serves as a relay for other members of the network. This means that individual nodes must collaborate to propagate the information in the network. The links between nodes represent relationships and social interactions between individuals, groups, organizations, or even entire societies. In reality, each network connection begins with an individual, using a digital device, reaching out to another individual or group for a social relationship of sorts and it snowballs into a mesh of social relationships connecting many individuals or/and groups. There are many online social network groups that have become established and household names like *Facebook*, *Myspace*, *Friendsster*, *YouTube*, *Flickr*, and *LinkedIn*.

The rapid growth of the internet, together with ever plummeting prices and increasing miniaturization of digital devices popularized by easy mobility and fast access to online services have all contributed to the creation of an exciting seemingly unlimited virtual environment in which anything is possible with the least effort. To make things more complicated, within this environment, anonymity and telepresence are almost assured, though this is an illusion, but very few understand it. In fact the combination of these two create a certain degree of individual confidence that may sometimes board to realms of danger. Because with these two, individuals in these virtual environments may be tempted to become reckless, knowingly or otherwise. This is one of the key causes and perpetrators of cyberbullying. But let us first define *cyberbullying*, *cyberstalking* and *cyber harassment*.

Since its debut, as the ugly side-effect of the online social networks, a lot has been written about it and consequently there has been lots of definitions of it given.

9.1.1 Cyberbullying

According to Wikipedia [1], cyberbullying is an action of harming or harassing an individual or individuals, mostly in the online social network environment, which we refer to most commonly in the public commons, as social media, but also via any other digital networks, in a repeated and deliberate manner. It is legally define also as:

- actions that use information and communication technologies to support *deliberate, repeated, and hostile behavior by an individual or group that is intended to harm another or others.*
- use of communication technologies for the *intention of harming another person.*
- use of Internet service and mobile technologies such as web pages and discussion groups as well as instant messaging or SMS text messaging with the *intention of harming another person.*

In all these legal definitions above, I have highlighted what is common and disturbing, that is, the prior and deliberate intention to intimidate, harm, control, manipulate, put down, falsely discredit or humiliate the recipient.

9.1.2 Cyberstalking

Stalking, a cousin of bullying, is defined as an unwanted and/or obsessive attention given to an individual or group by a perpetrator or perpetrators. Cyberstalking, a cousin of cyberbullying, then is digital stalking, usually using online media. Cyberstalking comes in many versions including but not limited to sending threatening messages to the victim, monitoring the victim, extortion, false accusations, altering a victim's information, identity theft, and the list goes on. The actions of a cyber stalker are usually repeated, persistent and often illegal.

As Mariam Merritt notes in her “Straight Talk About Stalking” essay [2], what is interesting and of course disturbing is that it is often perpetrated not by strangers, but by someone known by the victim. Of course the list of known people is long including one’s ex(s), former friends and/or acquaintances.

Remember cyberstalking is a form of cyberbullying.

9.1.3 Cyber Harassment

According to Merriam-Webster’s dictionary, to *harass* is to continuously and persistently annoy someone: to create an unpleasant or hostile environment for an individual, especially by uninvited and unwelcome verbal or physical conduct and also to make repeated attacks against a victim [3]. Based on these definitions, then *harassment* is the act of doing one or more of the above intended for disturbing,

tormenting or annoying the victim. When these acts are done by someone or a group of people using online digital technology, then we have *cyber harassment*.

Key differences between cyberstalking and cyber harassment they are both techniques of threatening or intimidating someone using digital communication media. However, cyberstalking involves the repeated and intentional intimidation and harassment.

Because cyberbullying is online harassment, we can refer to cyber harassment as cyberbullying.

9.2 Types of Cyberbullying

Because of the flexibility, ease of use, anonymity and telepresence of virtual online technologies and environments, the online environment offers many different ways harassment can be perpetuated online. Kim [4] outlines five different types of cyberbullying:

9.2.1 Harassment

As we pointed out above, harassment is the act of knowingly, purposely and repeatedly annoy; create an unpleasant or hostile environment especially by uninvited and unwelcome verbal or physical conduct and make repeated attacks against a victim.

9.2.2 Flaming

Flaming is “burning fiercely inside and emitting flames” by someone as a way to express annoyance. In online networks and in social media, this may involve exchanged emails, instant messaging or chat rooms among the parties involved. So if it is directed to an individual by a person or group of people, it is a form of harassment.

9.2.3 Exclusion

Exclusion harassment is slightly more difficult to define but it involves an intentional exclusion of an individual or a group by an individual or a group from an online space for the purpose of using the space to harass the victim(s).

9.2.4 Outing

Outing is unwanted and uncalled for online display in public commons of a victim's information and other attributes for no other purposes than harassment.

9.2.5 Masquerading

Because the digital online environment supports anonymity and telepresence both of which can enable an individual or a group to hide their true identity, acquire false identities and masquerade online with the intention of harassing others, online environments support masquerading as a form of cyberbullying.

9.3 Areas of Society Most Affected by Cyberbullying

Although the growth and popularity of communication technologies has equally affected all of us, user preferences of these technologies is stratified by a number of factors including age groups, income levels and geographical locations. Because of this, user activities are following these divisions. Divisions like these have seen vices like cyberbullying taking root in specific divisions of society. For example, cyberbullying is more prevalent in school age youth and in the working communities especially white collar workers. So let us look at cyberbullying in these areas.

9.3.1 Schools

According to the report “Facing the Screen Dilemma: Young Children, Technology and Early Education”, by Campaign for a Commercial-Free Childhood (CCfC) [5], on any given day, 64% of babies between 1 and 2 watch TV and videos for an average of slightly over 2 h; that in 2011 there were 3 million downloads just of Fisher Price apps for infants and toddlers; that estimates of how much time preschoolers spend on average with screen media, range from at least 2.2 h to as much as 4.6 h per day. Even though there’s no research showing the benefits of introducing children to new technologies in the first years of life, parents of young children are increasingly uploading new technologies to their young children in the belief that technology will make their children smarter. With this thinking, educators and school districts and authorities at every level are facing increasing pressure to increase both the technology and the amount of time children spend with digital technologies in early childhood settings.

Children growing with these ever increasing and involving technologies have acquired a high degree of easiness of use of these technologies far superior than their parents, yet with limited to no guidance and counseling on the vices of these technologies.

The rapid growth of technologies, the increasing ubiquitous use of technologies, early acquired ease of use of new and powerful smartphones and laptops, lack of counseling and the ever present curiosity of youth are all driving an increasing number of young people to try out these new technologies in online social media. While all this is going on, there is limited to no parental control and guidance. This is resulting in an increase in cyberbullying in schools, which in turn is leading to increased suicides of the young people.

9.3.2 Cyberbullying in the Workplace

Bullying in the workplace is not new. It has been in the workplace since way back before cyber technology invaded the workplace. The reason for this is because bullying is a character trait in an individual that may be helped by the different characteristics of technology like anonymity, speed, reach and ease of use. It is motivated by the bully's own lack of self-esteem rather than the specific actions, appearance, or personality of the victim [6]. Because of their internal shortcomings or weaknesses, many bullies feel threatened that they cannot cope with certain aspects of what they are supposed, expected or required to do. This threat, many times, may lead them to take defensive actions by trying to remove the source of the threat. This may be done in several different ways. Technology then, through its attributes like anonymity, scope, ease of use and others, help them in achieving whatever desired action they want accomplished. The increasing use of workplace "bring your own devices" (BYOD), take home work-related digital devices along with increasing use of online social media, have all increased the channels of harassment.

9.4 Legislation Against Cyberbullying

As we have pointed out earlier, bullying, as a vice, is not new. It is as old as humanity itself. Of course the definition of it has changed over the years as our expectations, living conditions and social status have changed. As these things changed, our tolerance of the vice has also increased as more awareness became more widespread. While in the past, big kids in schools and big and powerful people in places of work, and even in families, used to administer high doses of what is today considered harassment as they "picked on" or "singled out" those that seemed to be less powerful and less threatening and nothing happened to them, things are not the same today due to high levels of awareness to these evils and high levels of reporting of these incidents.

With more awareness and better reporting, direct bully activities have been declining, until the internet brought in a medium that supported both anonymity and telepresence of the bully. So anonymity, telepresence and the ubiquitousness of the use of mobile technology have all led to a rapid growth of the vice, there

increasing the number of victims and indeed the number of suicide of young victims. This has led to a call for legislation and public awareness. Several legislative bodies and institutions at different levels have been developing laws, statutes and policies.

9.4.1 Federal Laws

Cyberbullying is vice that affects all social strata and income levels. Because of this and the fact that it is emotional since it affects mostly young people, there has been efforts all over the place for legislation to mitigate it. Even though this is the case, legislation to combat it at federal level is still missing as of this writing. However, even if no federal law directly addressing cyberbullying has been passed, there are some cases, where bullying overlaps with *discriminatory harassment* which is covered under federal civil rights laws enforced by the U.S. Department of Education (ED) and the U.S. Department of Justice (DOJ) [7]. These laws include [7]:

- Title IV and Title VI of the Civil Rights Act of 1964
- Title IX of the Education Amendments of 1972
- Section 504 of the Rehabilitation Act of 1973
- Titles II and III of the Americans with Disabilities Act
- Individuals with Disabilities Education Act (IDEA)

9.4.2 State Laws

There is a patchwork of state sponsored cyberbullying laws. A great resource to learn about cyberbullying laws and legislations at state level is the Cyberbullying Research Center [8]. Also see 9.7 below for more resources. According to the Cyberbullying Research Center [8], at last count, 44 states have laws regarding bullying, and 30 of those included some mention of electronic forms of harassment. For the time being, all these laws, state or local ordinances, are simply directing school districts to have a bullying and harassment policy, without the actual content of such policies.

9.4.3 International Laws

The rapid growth of the Internet which has quickly engulfed the globe, the plummeting prices of web-enable smart mobile devices bringing the rest of humanity into cyberspace has made cyberbullying a global problem. So the desire to contain

it is also global. Although the need is there and it is growing, just a few countries and regions have passed legislation to combat it. Among these are the European Union (EU), Canada, Australia, Spain, UK, France, Germany and others but it is slow going.

9.5 Effects of Cyberbullying

As we have been pointing out, cyberbullying is bullying using a new Internet supported medium. Statistics from different countries are showing that the vice is growing, hampered only by massive awareness campaigns, hence affecting more and more people.

Like all forms of bullying, cyberbullying affects everyone, the bully, the victim and the bystanders, though in different ways. Thus the effects vary by the type of bullying, the techniques used and the role one plays in the bullying cycle. Because of this, it is difficult to put the effect of cyberbullying into specific categories since different people react differently to the same causes. The major underlying effects, though, cutting across the board are psychological, emotional and physical stress. These types of effects affect people differently. For example StopBullying.gov [9] reports the different effects in the bullying circle as:

Kids Who are Bullied

Kids who are bullied are more likely to experience:

- **Depression and anxiety**, increased feelings of sadness and loneliness, changes in sleep and eating patterns, and loss of interest in activities they used to enjoy. These issues may persist into adulthood.
- **Health complaints**
- Decreased academic achievement—GPA and standardized test scores—and school participation. They are more likely to miss, skip, or drop out of school.

Kids Who Bully Others

Kids who bully are more likely to:

- Abuse alcohol and other drugs in adolescence and as adults
- Get into fights, vandalize property, and drop out of school
- Engage in early sexual activity
- Have criminal convictions and traffic citations as adults
- Be abusive toward their romantic partners, spouses, or children as adults

Bystanders

Kids who witness bullying are more likely to:

- Have increased use of tobacco, alcohol, or other drugs
- Have increased mental health problems, including depression and anxiety
- Miss or skip school

Indeed, as we pointed out earlier, what cuts across are psychological, emotional and physical factors that can lead to the individual's overall wellbeing. This is not limited to school children alone, although they are more likely to suffer these effects than adults. These effects may influence an individual's health and psychological balance which may lead to suicide, although it may not be the cause of it.

9.6 Dealing with Cyberbullying

Cyberbullying comes in many forms including pretense, masquerading, hacking into the victim's online account, invading and bracketing of social media and a lot more others. Because of the varied way cyberbullying is carried out, dealing with it needs to be also carefully chosen to deal with each of the many approaches of delivering it to the victims.

However, since most of its effects are based on psychological, emotional and physical stress, there are underlying and broad approaches that we can take that will cover the major source of cyberbullying and will deal with the different reactions to its effects. These include:

9.6.1 Awareness

Find ways of developing massive education campaigns about what cyberbullying is, who it affects and its consequences which may include death. Broad mass and targeted education campaigns are essential. These mass awareness education campaigns are meant to focus on targeted audiences. For example, if the audience is a school or school going children, techniques must be found that delivers the message in quantized and proportions that are relevant and enjoyed by the targeted age group. If it is targeted to a work environment, delivery techniques are different.

9.6.2 Legislations

Mass education and awareness programs, however targeted they are, can go so far in the absence of policies, statutes and laws with corresponding enforcement. So legislation at either state or federal levels is necessary at least for schools. In

businesses the best approach is for the companies to draw up operating policies that involve guidelines of behavior of all workers. Such company policies must be enforced to be effective.

9.6.3 Community Support

Communities should also get involved in cyberbullying reduction and prevention. Cyberbullying public awareness activities must be included in community public activities especially those directed to youth in the community. Part of the package of community cyberbullying awareness campaigns should include some form of reporting. Without it, the efforts are not likely to succeed.

9.7 Resources

There are a number of resources one can go to for help. Most of these resources are directed towards children, parents, educators and adults mostly in the work environment. Among these are:

- The Cyberbullying Research Center: <http://cyberbullying.us/>. This is a great resource with materials for all categories of users. They also have current statistics for cyberbullying and additional reading resources and testimonials.
- Stopbullying.gov: <http://www.stopbullying.gov/resources>. This site gives you tips, facts, toolkits, training materials, and more. You get access to a trove of information on cyberbullying on this site by entering a topic related to bullying in the keyword search area. Their collection includes federal and non-federal training materials, evidence-based program directories, articles, and others related to bullying.
- The National Crime Prevention Council: <http://www.ncpc.org/topics/cyberbullying>. This site gives a variety of information on both bullying and in particular cyberbullying including:
 - What Parents Can Do About Cyberbullying
 - Cyberbullying FAQ for Teens
 - Cyberbullying PSA Contest
 - Training on Cyberbullying
 - Bullying and Intimidation
 - Professional training from NCPC for youth and adults on managing bullying situations
 - Products and Publications on Cyberbullying
 - Helping Kids Handle Conflict
 - Cyberbullying Banners for the Web
 - Rapid Response Outreach Tools on Cyberbullying
 - Cyberbullying Crime Flyer
 - Cyberbullying Crime Palm Card

- Cyberbullying Crime Poster
- Cyberbullying Research Brief
- Programs on Cyberbullying
- Be Safe and Sound in School
- The Human Rights Campaign: <http://www.hrc.org/resources/entry/resources-on-cyber-bullying>. This site gives the reader a list of organizations that focus on cyber bullying and provide the most up-to-date articles, fact sheets, and news stories on cyber bullying as well as specific education resources for parents, educators and children:
 - Cyberbullying.org—This site is run by the Center for Safe and Responsible Internet Use, and provides a number of helpful resources for educators and parents, including an educator's guide to cyber bullying and information on legislation related to cyberbullying.
 - Cyberbullying.us—This online research is maintained by Justin W. Patchin and Sameer Hinduja of the Department of Criminology and Criminal Justice at Florida Atlantic University. They have written numerous articles and given several presentations across the country on the nature and extent of cyberbullying. This site includes extensive resources on cyber bullying as well as research, news and events on the topic.
 - I-Safe—a non-profit foundation dedicated to protecting the online experiences of youth everywhere. It incorporates classroom curriculum with dynamic community outreach to empower students, teachers, parents, law enforcement, and concerned adults to make the Internet a safer place.

This is in no way exhaustive, there are many more resources focusing on bullying and cyberbullying.

Exercises

- Who are the victims of bullying? Cyberbullying?
- Discuss the traits of a bully, cyberbully.
- What is the legal definition of cyberbullying? Is there one?
- Cyberbullying laws vary greatly depending on location, discuss what would be common among all.
- Discuss what type of enforcement of laws, statutes and policies are possible, if any.
- Describe the different types of cyberbullying.
- Trace the growth of cyberbullying following the evolution of online social media.
- Compare two or more state laws on cyberbullying with the laws in your state.
- Cyberbullying may go on unnoticed, discuss efforts being taken to identify victims early—in your state.
- Discuss techniques and skills required to deal with cyberbullying.

References

1. Wikipedia. <https://en.wikipedia.org/wiki/Cyberbullying>
2. M. Merritt, *Straight Talk About Stalking*. <http://us.norton.com/cyberstalking/article>
3. Merriam-Webster's. <http://www.merriam-webster.com/dictionary/harass>
4. T.H. Kim, *5 Different Types of Cyberbullying*, 23 Dec 2013. Cyber Bullying News. <https://endcyberbullying.org/5-different-types-of-cyberbullying/>
5. Campaign for a Commercial-Free Childhood (CCfC), Facing the Screen Dilemma: Young Children, Technology and Early Education. <https://fairplayforkids.org/pf/facing-screen-dilemma/>
6. MONEYWATCH, *Understanding the Reasons for Workplace Bullying*, 13 Nov 2007. <http://www.cbsnews.com/news/understanding-the-reasons-for-workplace-bullying/>
7. Stopbullying.gov. <http://www.stopbullying.gov/laws/federal/index.html>
8. Cyberbullying Research Center. <http://cyberbullying.us/the-current-state-of-cyberbullying-laws/>
9. StopBullying.gov. <http://www.stopbullying.gov/at-risk/effects/index.html>



Evolving Realities: Ethical and Secure Computing in the New Technological Spaces

10

Artificial Intelligence and Cyberspace

Although it is obvious that machines can perform some activities at a higher level than persons can; these tasks remain, by and large, highly specialized and therefore remote from the capacity of human intelligence for multipurpose activities.

—Michael R. LaChat, *The Methodist Theological School in Ohio*

Abstract

Artificial Intelligence discusses the new frontiers of ethics in the new artificial intelligent (AI) technologies and cyberspace. The chapter explores how these new frontiers are affecting the traditional ethical and social values. Our discussion is based on the premise that artificial intelligence technologies create possibilities to understand and extend human knowledge to create intelligent agents perhaps with a human-value base, intended to help solve human problems. Finally we discuss a global mesh of interconnected computer networks, commonly referred to as cyberspace, which makes it possible for anyone using a point of entry device like a computer, smartphone or any other internet-enabled electronic device to reach anyone else, with the potential to access the mesh, through a one-on-one, one-to-many and many-to-one communication capabilities or through broadcasting via the world wide web. Cyberspace, because of immense telepresence capabilities and global reach, creates a potentially dangerous environment where one can do anything with no elegance, no accountability and not limited.

Learning Objectives

After reading this chapter, the reader should be able to

1. Understand the value of ethics in automated decision making.
2. Identify and discuss the different forms of automated decision making.
3. Recognize the role ethics plays in artificial environments.
4. Be able to articulate what makes up cyberspace.
5. Identify and discuss credible safeguards to ensure privacy concerns and prevent runaway computation resulting from autonomous agents.
6. Understand the role of autonomous agents in our daily lives.
7. Recognize and discuss the responsibilities of users of autonomous agents.
8. Recognize and discuss the responsibilities of users in artificial intelligence spaces.
9. Learn the complexity of cyberspace issues.
10. Learn the ethical framework of cyberspace.

Scenario 8: One for the Road—Anyone?

Florence Yozefu is a brilliant scientist who heads a robotics research laboratory at one of the top ten research universities. Florence has been developing wearable robotics gear that can take over the driving functions of a vehicle from a human operator when it is worn by the driver. In laboratory tests, the robot, nicknamed Catchmenot, has performed successfully whenever Florence and her assistants have worn the robot. However, no real-life experiment has ever been conducted outside the lab. Florence has been meaning to try it out one day, but has not got a chance as yet to do so.

For New Year's Eve, Florence has plans to visit her mother and sister, about 100 miles away. This was a good opportunity to show her mother and her sister what she has been up to these last few months. So she decides to take Catchmenot with her. She packs her car the evening before and on the morning of the trip, she passes by the lab to get her robot and put it in the car. She drives the 100 miles in a little under her usual time and arrives at her mother's house earlier than usual. In the evening, Florence bids her mother good-bye and passes by her sister's apartment as promised. But at her sister's apartment, she finds a few of her teen friends and they get right into a party mode. Florence drinks and dances and forgets about time. There are many stories to tell and to listen to. About 1:00 a.m., after the midnight champagne toast, she decides to leave and drive back to her apartment.

She had promised to accompany her friend to a preplanned engagement. Although she is very drunk, and against her friend's advice and insistence that she should not drive, Florence puts on Catchmenot and in a few minutes she is off. Thirty minutes later, she is cruising at 70 mph and she is also sound asleep.

She is awakened by a squirrel running all over her car at about 5:00 a.m. She is parked by the roadside in front of her apartment complex. She has made it home safely. She has no idea when and where she passed out and what happened along the way. She will never know. Although she is surprised, confused, and feels guilty, she is happy how well Catchmenot has worked. She decides to market it.

How much should she charge for it, she wonders.

Discussion Questions

1. *Why did Florence feel guilty?*
2. *Is Florence right to market Catchmenot?*
3. *If anything went wrong along the ride home, would Florence be responsible? Who should be?*
4. *Is it ethical to market Catchmenot?*
5. *Discuss the ethical implications of artificial intelligence based on Catchmenot.*

10.1 Introduction

In the theistic tradition of Judeo-Christian culture, a tradition that is, to a large extent, our “fate,” we were created in the *imago Dei*, in the image of God, and our tradition has, for the most part, showed that our greatest sin is pride: disobedience to our creator, a disobedience that most often takes the form of trying to be God. Now, if human beings are able to construct an artificial, personal intelligence—and I will suggest that this is theoretically possible, albeit perhaps practically improbable—then the tendency of our religious and moral tradition would be toward the condemnation of the undertaking: We will have stepped into the shoes of the creator, and, in so doing we will have overstepped our own boundaries. [1]

The Hebraic attitude towards AI has been one of fear and warning: “You shall not make for yourself a graven image...”, while that of the “Hellenic” has been fascination and openness. [1]

Artificial intelligence (AI) is an exciting technological frontier offering a novel environment with unlimited possibilities. The AI environment works with the possibilities of understanding and extending knowledge to create intelligent agents, perhaps with a human value base, intended to help solve human problems.

Virtualization is a process through which one can create something that is there in effect and performance but in reality is not there—that is, it is virtual. It is a physical abstraction of reality, a real phenomenon such as a company’s computing resources including storage, network servers, and memory. Virtualization involves and absorbs participants into a virtual reconstruction of real-world entities into seemingly real images with corresponding in-depth information to turn these images into a high degree of realism. It is a process that embodies both abstraction and reconstruction to create a sense of complete participant immersion

yet with autonomy of participants to vary their chosen new environments to suit individual liking. In other words, virtualization is a process that makes real entities, scenes, and events virtual mirror images of self; it is a virtualization of reality. In many ways it is a mediation of interaction through an electronic medium between humans and humans as well as between humans and machines [2].

Cyberspace, on the other hand, is a multidimensional space vision of pure information either at rest in large storage media or in motion between digital nodes that form a mesh.

The description we have given here of the technologies and environments are quite enticing to us humans who are naturally curious and drawn to investigate new phenomena whenever possible. Thus, the frontiers we discuss in this chapter have drawn a number of people, some for the experience of participating in cyberspace, others to experience the thrills of virtual reality, and yet others to investigate the application of knowledge in areas yet unknown. Wherever and whenever they are drawn to these environments, human beings are participatory: they try out things and get involved as they pursue individual social goals.

10.2 Artificial Intelligence

Artificial intelligence (AI) is a field of learning that emulates human intelligence. The recent development of AI and its seemingly unbounded potential to solve real life problems has broadened computer technology applications into exciting areas that were thought to require deep human intelligence. Because human intelligent behavior is so varied, poorly defined, and difficult to predict, let alone understand, most artificial intelligence studies concentrate on a limited number of areas in which human abilities to think and reason are clear and the overall human intelligent behavior is well understood.

These areas exhibit those aspects of human intelligence that can be represented symbolically with first-order logic such as game playing and natural language understanding. Other areas that exhibit a high degree of human intelligence such as thought processes, pattern recognition, and concept formation are still abstract and scarcely understood. However, with the recent realization that most real-world problems that are usually not so difficult for human intelligence, but are still intractable for current machine intelligence, involve numerical computation, and with the improvement in the computation power of new computers, AI research is slowly shifting from those areas that use symbolic manipulation to numerical computation. This shift is bringing diversification and practicability to AI and increasing the repertoire of AI techniques. Because of the lack of incorporation of common sense into AI techniques, however, progress in real terms has been slow, except in areas such as robotics and machine learning in which there has been more excitement because of the practicability of the applications to our daily lives. However, the realization is opening new doors in AI research in areas such as neural networks and fuzzy logic theory, and it is causing scholars to take a new look at ways of achieving machine intelligence and at the same time start

to study the social and ethical impact and where to place responsibility for such machines, if they ever come to be. In this section we look at the social and ethical implications in these active areas where AI techniques have made advances and are becoming established.

10.2.1 Advances in Artificial Intelligence

Starting with Alan Turing in his 1950 machine intelligence experiments, in which he proposed a game played between a computer and a human being that could demonstrate whether a machine, in this case a computer, could think, there has been a steady growth of interest in building intelligent machines, commonly known as *autonomous agents*, in those areas of AI in which intelligence can be modeled easily, such as in game playing, expert systems, natural language understanding, neural networks, and robots.

Autonomous agents are not new to AI. Since Turing, numerous scholars such as Marvin Minsky, Alan Key, and Rodney Brooks have at various times studied agents' behavior and sometimes constructed autonomous agents. An autonomous agent is a collective name encompassing both hardware and software intelligent agents. Autonomous agents can take different forms depending on the nature of the environment. For example, if the environment is real 3D, the agents are robots, whereas in 2D they are computer programs referred to as intelligent agents.

As progress is made in AI research, areas of application of autonomous agents are becoming more numerous. For example, robots have been used in a number of areas including surveillance, exploration, and in inaccessible and hazardous environments. More intelligent agent robots are being helped by better vision systems, sensors, and easier programming languages. The field of robotics is getting wider, involving areas such as perception, cognition, and manipulation in which success has been most apparent through industrial applications. Industries and the military are relying more and more on these new technologies.

10.3 Cyberspace and the Concept of Telepresence

Cyberspace is a global artificial reality environment based on a global mesh of interconnected computer networks. This mesh allows and makes it possible for anyone using a point-of-entry device such as a computer, smartphone, or any other Internet-enabled electronic device to reach anyone else, with the potential to access the mesh, through a one-on-one, one-to-many, and many-to-one communication capabilities or through broadcasting via the World Wide Web. Cyberspace, because of immense capabilities and global reach, is used either in real time or otherwise, simultaneously by millions if not billions of people around the world. Through its specialized applications, users enter this virtual world electronically to get many services and perform numerous tasks via use of applications to benefit humanity.

When one is in cyberspace, there is a feeling of being in a location other than where one actually is. This is a notion of *telepresence*, a feeling one gets from being present at a place other than one's true location. This feeling and sometimes the ability to control a robot or another device at a distance gives cyberspace and in fact makes cyberspace a virtual environment with power of immersion, the kind of experience we discussed in the previous chapter. Whether in cyberspace or not, both telepresence and immersion, as a concept, require that the users' senses be provided with such stimuli as to give the feeling of being in that other location. Additionally, users may be given the ability to affect the remote location. In this case, the user's position, movements, actions, voice, etc. may be sensed, transmitted, and duplicated in the remote location to bring about this effect. Therefore, information may be traveling in both directions between the user and the remote location [3].

10.4 Securing Cyberspace

Keeping cyberspace users secure is a daunting job that requires advanced techniques and prevention methods. Both the detection and prevention techniques are changing very fast. The several known ways of doing this include the following:

10.4.1 Detecting Attacks in Cyberspace

A detection system deployed around a computer system or a computer network is a 24-h monitoring system to alert the owner or system manager whenever something unusual occurs, something with a non-normal pattern, different from the usual pattern of life in and around the system. The monitoring system is actually an alarm system that must continuously capture, analyze, and inform the system manager on the daily patterns of life in and around the computer system.

10.4.2 Vulnerability Scanning in Cyberspace

System and network scanning for vulnerability is an automated process where a scanning program sends network traffic to all network nodes or selected nodes in the network and expects receiving return traffic that will indicate whether those nodes have known vulnerabilities. These vulnerabilities may include weaknesses in operating systems and application software and protocols.

10.4.3 Privacy in Cyberspace

In Chap. 4 we described a scenario in which Citizen X no longer has privacy. In fact, many have been questioning the very concept of personal privacy—whether it still exists at all.

According to recent studies, personal privacy is becoming the number-one social and ethical issue of concern for the information age. Advances in technology have brought with them gadgetry that has diminished individual private spaces through electronic surveillance and monitoring, transmission, scanning, tapping, and fast and more efficient means of collecting, categorizing, and sorting data. Among the current issues of concern are the following:

Transmission, scanning, and tapping using computers and mobile phones.

Information gathering as a result of better software and equipment. In this category, cyberspace is proving to be a fertile ground because of its powerful search engines and the volume and speed of data. Information gathering, especially from an individual, is now the most threatening and worrisome form of invasion of privacy because of its ever-increasing commercial value. With the right tools at the right time, one can gather all forms of information one needs about a subject in a matter of hours from sources that include county offices, auto registration offices, credit card bureaus, utility companies, postal records, telephones, and satellites. In a day's work one can build an individual profile that in the past would have taken years, if not being outright impossible.

Individual tracking through mobile and paging devices and computers. Many carrier companies are now using employee tracking to get up-to-the-minute information on what the employee is doing at any given moment.

Private investigators (PIs) have found a new partner in cyberspace using satellites; PIs can track and report on any individual with alarming details.

Information-gathering abuses by established information-gathering agencies. Governments and government agencies such as the National Security Agency (NSA), the Federal Bureau of Investigations (FBI), the Central Intelligence Agency (CIA), all in the United States; both UK's MI5 (the Security Service) and MI6 (the Secret Intelligence Service); and the Federal Security Service of the Russian Federation (FSB) in Russia, all gather information in the name of national security and crime fighting. In such cases information is gathered on an individual before even an arrest is made. The biggest threat from these established information-gathering agencies is that if not properly focused and supervised, they have the capacity and means to do whatever they like with individual freedoms, bringing us to the privacy paradox that too much individual privacy is very dangerous. According to Brandt [4], society consists of individuals; if each individual has total privacy, then society as a whole has zero security. Of course, no government can allow this to happen. Because no government can exist without security, somebody's privacy has to be sacrificed.

10.4.3.1 Privacy Protection in Cyberspace

As a cyberspace community member, or cyberzen, you have to be proactive in protecting your privacy, which can be done through information control, property control, and use of anonymity. You are in control of your privacy and you decide on how much to give up. Individual privacy is threatened whenever you voluntarily

surrender information through online transactions such as entering online sweepstakes or filling out online surveys. Make sure you surrender personal information only when you must and as minimally as possible. Controlling access and information about your personal property (e.g., car, house, or computer) can also help safeguard your privacy.

10.5 Social Issues in Cyberspace

The Internet and indeed cyberspace has become the largest repository of information and source of materials for almost anything worth searching, whether it has value or not. Indeed, although one can find information for any food recipe of choice, it is also the source of all sorts of information on terrorism. As the Internet is a source of terrorism information from how-to to recruiting, there are calls for cyberspace censorship. Depending on the country, the rationales for censorship have varied from historical, social, political, and economic to cultural grounds.

But cyberspace censorship is proving to be both difficult and very expensive for those trying, and many governments and censorship bureaus are fighting a losing battle because of the exponential growth of the Internet. For governments and censorship bureaus to keep pace with this growth, they have to be continually hiring censors, which is very expensive.

In addition to the explosive growth, Internet content is becoming highly specialized and richer in graphics, which requires very expensive equipment and highly trained people to keep pace. Also, cyberspace's one-fits-all role of telecommunication, broadcast, and computer services is making censorship very difficult and expensive. Effective censorship calls for a better focus on at least one of these because not all three media carry the same materials. Censors may concentrate on materials of a broadcast nature like web pages, assuming that the content they are looking for is more likely to be in this kind of medium, but may find this is not the case. Contents of materials also change in nature and focus from time to time and from medium to medium.

And finally, applying geographically defined court jurisdictions in cyberspace, a physical boundary-less entity, is proving to be futile at best. Any attempt to enforce the law in one country means enforcing the same law in many others. Many cyberspace problems that lead to censorship have been brought about by the transient nature of membership in cyberspace communities. Users rarely stay in the same community with the same peers for long. These transitions are brought about by a number of factors including changing interests, job changes, changing lifestyles, and a host of others. Each cybercommunity is a moving target community. Transients do not have allegiance and, therefore, no responsibility and accountability.

If the ideal situation is to be realized for every community user worldwide, cyberspace needs to be a place of comfort and entertainment where one can be satisfied and one's curiosity and dreams fulfilled. It needs to be a decent place where children can log on without sparking fear in their parents that their offspring will

stumble onto inappropriate material. How this wished-for security can be achieved in cyberspace is the focus of many governments and civic organizations around the globe. The question is how, without total censorship of cyberspace, can these governments and civic organizations do it? As pressure builds for some kind of action, governments have started formulating policies for cyberspace use, some of which is ending up as censorship. An array of measures is being debated in legislatures and government board rooms around the globe. Such measures include the following:

Guidelines, usually issued by a government body, outlining the do's and don'ts for cyberspace users.

Some governments are encouraging the filtering process, picking certain domains and specific names they deem suggestive of what they are looking for. This kind of filtering is usually delegated to system administrators and system operators.

States are setting up cyberspace on-ramp services and, through low user fees, are encouraging users to opt for these instead of private services. They also urge all educational and research institutions to use these services for free.

Other governments are encouraging cyberspace user groups to "self-censor" through appeals to patriotism and nationalism and asking users to refrain from using or downloading what they deem offensive materials and sometimes even to report whenever they see objectionable materials [5].

The use of blocking gadgets is also on the table. An industry for these gadgets is emerging. The blocking software programs work by blocking unwanted cyberspace materials [6]. But blockers have a serious drawback: they can be very easily circumvented by smart programmers.

And finally, some governments are either enacting laws or are amending existing laws.

In the U.S. Congress, over the course of several years now, lawmakers have been trying to come up with ways to regulate indecent activities on the Internet. Such efforts have of course yielded a series of acts and bills with the majority of them in 1995 and 1996, the 2 years that encompassed the widely debated Telecommunications Bill of 1996. This bill contained the controversial Communications Decency Act, now defeated in court, whose purpose was to provide protection to citizens, especially minors, against harassment, obscenity, and indecency by means of telecommunication devices such as computers. This was not the first legislative measure to censor the Internet. Other legislation includes the Protection of Children from Computer Pornography Act of 1995, the Comprehensive Terrorism Prevention Act of 1995, the Digital Millennium Copyright Act (2001), the Child Online Protection Act (1998), and the Children's Online Privacy Protection Act (1998). The intent of these efforts has always been to protect the helpless, such as children, from indecent and illicit activities on the Internet. But because of the borderless nature of cyberspace and the rapid growth of the technology these

efforts are meant to fight, they have had little or no impact on their intended targets, and there are always new cries for more control. Individual governments' efforts are also complicated by other factors. Chief among them is the realization that because of the differences in political, social, cultural, and religious systems around the globe, what is considered politically offensive in one locale may not be so in another and whatever is considered tolerable in one culture may not be so tolerated in another.

10.6 Artificial Intelligence and Ethics

Human beings by nature strive to create a good life for themselves through the acquisition of knowledge and the creation of intelligent machines to do most of the jobs that we either are not able to do or do not like to do. But according to Lugar and Stubblefield [7], the notion of human efforts to gain knowledge as a transgression against the law of God is deeply rooted in us, and we still believe that the human quest for knowledge must eventually lead to disaster. This belief has not been affected by current philosophical and scientific advances, but instead has produced the Frankenstein monster syndrome of fear of new advances in intelligence, particularly machine intelligence. It is this fear that has been the source of controversy in the field of artificial intelligence and has the potential to hinder its development.

Writers of fiction such as Asimov [8] have written about AI. The most positive has been Jack Williamson who, in “With Folded Hands” [9], portrayed a young scientist disillusioned by man’s destructive nature who creates robots to follow the Asimovian Prime Directive [10]: “To serve and obey, and guard men from harm.” In the story, robots replicate themselves and do all the jobs he wants them to do—until he realizes the mistake he has made. But it is too late. He has rendered himself useless. “Men sat with idle hands because there was nothing left for them to do.” Science was forbidden because the laboratories were dangerous to man. There was no need to acquire knowledge because the robots could do anything and do it better. “Purpose and Hope were dead. No goal was left for existence.” One would ask why the young scientist could not kill the robots. He tries, but they stop him because he is violating the Prime Directive. Meanwhile, they multiply.

Philosophers too have long expressed this fear. According to Weizerburm [11], it is immoral to use a computer system to replace human functions involving interpersonal respect, understanding, and love. Floyd believes that computers should only be used if “there remains sufficient scope outside the computer application for other human faculties and forms of experience not to degenerate” [11]. Marvin Minsky likens it to an old paradox dealing with slave education: “If you keep them from learning too much you limit their usefulness; if you help them become smarter than you, then you may not be able to trust them to make better plans than they do for you” [12].

To us all, AI represents a social and ethical paradox. We want the machines to do those things we are not able to do because we are not good at them, yet

we do not want them to get too good. We probably would have no reservations and certainly no objections if only we could be assured that they are, to put it in Hans Moravic's words, "our very own mind-children" [12]: that is, that they share our own values, truths, and virtues. As these autonomous agents achieve better intelligence and become more widely used and acceptable, they will be taking on more and more responsibility and autonomy that only humans have been known to have. This development raises many questions about the future of these autonomous agents, their relationship with humans, their behavior, and emotions. Among such questions are the following:

- How will humans perceive these intelligent agents?
- How will the autonomous agents themselves feel about human beings?
- Will human beings let these intelligent "creatures" keep on getting more and more intelligent even though they are aware the ultimate end result would be to surpass human intelligence?
- How will they do what they are supposed to do?
- Will they do only what they are supposed to do?
- Who will be responsible for the actions of these agents?
- Will these agents outperform their owners?
- Will they eventually eliminate the need for human skills?
- And the most important of these questions is, how much power and autonomy should we give these creatures, and will these agents eventually take away human autonomy and consequently take control of human destiny?

According to Mitchell Waldrop [10], as these autonomous agents become more and more intelligent, we need a theory and practice of machine ethics that will embody a built-in code of ethics in these creatures in the spirit of Asimov's laws of robotics and Carbonell's hierarchy of goals [10, 13].

Isaac Asimov, in his book *I, Robot* (1950), created his fictional robot characters with an implanted code of conduct he called the Laws of Robots:

A robot may not injure a human being or, through inaction, allow a human to come to harm.

A robot must obey the orders given to it by human beings except when such orders would conflict with the first law.

A robot must protect its own existence as long as such protection does not conflict with the first or second law [8].

According to Carbonell [13], programs can be governed by a hierarchy of goals that act as a guide and a prescribed direction in the program's reasoning processing. This hierarchy should then be set up so that it inputs a code of ethics into the programs. The question, however, is whether it will be followed by intelligent programs and robots, and to what extent? Is there a possibility that they can vary the code, however much we try to stop them from doing so? Carbonell's concept is good, but it creates many questions that still need answers.

Discussion Topics

1. Is the construction of a personal AI or a human-like robot an immoral experiment?
2. Does a personal AI or a human-like robot have rights?
3. Can artificial intelligence be moral?

Can there be benefits to humans of these agents without infringing on our autonomy? There are those who see the future of AI as very beneficial to humanity. They see a fruitful partnership with the agents in which the agents are relieving us of all our dangerous and tedious tasks, making our lives a little easier and helping us reach our ultimate goal of the good life. They further believe that we will learn more about ourselves in the attempt to construct something like ourselves. Will we become better as human beings in what we do and how we do it? Will the success of AI and the creation of ourselves bring us to the full understanding of our inadequacies and belittle our human experience? There is a possibility that this may bring us to transcend our human experience altogether and lead us to the post-human. At that stage we will be thinking big of ourselves, able to transcend all our human fears that something can go wrong and looking forward to unimaginable possibilities—if it comes to pass, that is!

But again, as LaChat [1] said, if we do not make a human-like robot to pass the famous Turing Test, then perhaps little of our effort will be lost and this might eventually bring us to the brink of our mysticism that has, at least, been partially “tested.” Will that make us feel more “special”? What will this do to our moral and ethical beliefs?

Exercises

1. Discuss the implications of the common knowledge problem on advances in AI.
2. Is the “mad scientist” syndrome justifiable? Discuss.
3. Why is the study of AI limited to only small arrears? Will this not hinder the study of human behavior in other areas?
4. Will the study of AI make us better understand human behavior?
5. Can the understanding of AI, if this ever happens, help to understand and model human behavior, and hence develop better techniques for teaching computer ethics?
6. As AI applications increase, as in the use of robotics, will the wider use of these “manlike” machines compromise our moral value system? Why or why not?
7. Is the Frankenstein Syndrome an exaggeration or is it real?
8. Is it possible to develop a robot with our own moral value system?
9. Will the development of more advanced intelligent systems improve the declining moral value?

-
10. Discuss the future of AI and its ethical implications
 11. Why is it so difficult to censor cyberspace?
 12. Suggest reasons why cyberspace should not be censored.
 13. List the steps taken by governments to curb cyberspace ills.
 14. Discuss the merits and demerits of censorship measures.
 15. Discuss the civic roles of cyberspace.
-

References

1. M.R. LaChat, (The Methodist Theological School in Ohio), *Artificial Intelligence and Ethics: An Exercise in the Moral Imagination*, AI Magazine, vol 7, no 2 (1986)
2. Wikipedia. <http://en.wikipedia.org/wiki/Virtualization>
3. Wikipedia. Telepresence. <http://en.wikipedia.org/wiki/Telepresence>
4. D. Brandt, Cyberspace Wars: Microprocessing vs. Big Brother. NameBase. *NewsLine* 2, July–August 1993
5. P. Lewis, Limiting a Medium Without Boundaries. New York Times, 15 Jan 1996, D1
6. A. Bhimani, Securing the commercial internet. Commun. ACM **39**(6), 29–35 (1996)
7. G. Luger, W. Stubblefield, Artificial Intelligence, 2nd edn. Benjamin Cummings, Reading (1993)
8. I. Asimov, *The rest of the robot* (Doubleday, New York, 1964)
9. J. Williamson, With folded hands, in *The Best of Jack Williamson*. (Ballantine, New York, 1978), pp. 154–206
10. M. Waldrop, A question of responsibility, in *Ethical Issues in Information Systems*, ed. by R. Dejoie, G. Fowler, D. Paradice (Byrd & Fraser, Boston, 1991)
11. C. Beardon, The ethics of virtual reality. Intell. Tutor Media **3**(1), 22–28 (1992)
12. M. Minsky, D. Riecken, A conversation with Marvin Minsky about agents. Commun. ACM **37**(7), 23–29 (1994)
13. J. Carbonell, *Subjective Understanding: Computer Models of Belief Systems* (University of Michigan Press, Ann Arbor, 1979)

Further Readings

14. J. Barlow, Electronic frontier: private life in cyberspace. Commun. ACM **34**(8), 23–25 (1991)
15. J. Bates, The role of emotion in believable agents. Commun. ACM **37**(7), 122–125 (1994)
16. M. Boden, Agents and creativity. Commun. ACM **37**(7), 117–121 (1994)
17. J. Cowie, W. Lehner, Information extraction. Commun. ACM **39**(1), 80–91 (1996)
18. M. Green, S. Halliday, A geometrical modelling and animation system for virtual reality. Commun. ACM **39**(5), 46–53 (1996)
19. F. Hayes Roth, N. Jacobstein, The state of knowledge-based systems. Commun. ACM **37**(3), 22–39 (1994)
20. T. Kanade, M. Read, L. Weiss, New technologies and application in robots. Commun. ACM **37**(3), 58–67 (1994)
21. T. Munakato, Y. Jani, Fuzzy systems: an overview. Commun. ACM **37**(3), 69–75 (1994)
22. D. Norman, How might people react with agents? Commun. ACM **37**(7), 68–71 (1994)
23. T. Poston, L. Serra, Dextrous virtual work: introduction. Commun. ACM **39**(5), 37–45 (1996)
24. D. Riecken, Intelligent agents: introduction. Commun. ACM **37**(7), 18–21 (1994)
25. M. Schoppers, Real-time knowledge-based control systems. Commun. ACM **34**(8), 27–30 (1991)

26. G. Singh, S. Fisher, D. Thalmann, Virtual reality software and technology: introduction. *Commun. ACM* **39**(5), 35–36 (1996)
27. J. Webe, G. Hirst, D. Horton, Language use in context. *Commun. ACM* **39**(1), 102–111 (1996)
28. M. Wilkes, Artificial intelligence as the year 2000 approaches. *Commun. ACM* **35**(8), 17–20 (1992)
29. Y. Wilkes, Natural language processing. *Commun. ACM* **39**(1), 60–62 (1996)



Ethical, Privacy, and Security Issues in the Online Social Network Ecosystems

11

Abstract

Ethical, Privacy, and Security Issues in the Online Social Network Ecosystems discusses the new realities of global computer online social network ecosystems, including moral and ethical dynamisms. Because we believe that a sound and details discuss of online social networks is based on a good understanding of the underlying network infrastructure, we start the chapter with a brief discussion of the computer network infrastructure. Based on this communication infrastructure we define a social network and its subset, the online social network. We discuss the types of social networks, their historical development and the different and changing services of online social networks. After discussing the basics of online social networks, we them focus on ethical, social and privacy issues in the online social network noting that while online, we inevitably give off our information to whomever asks for it in order to get services. We note that routinely information collected from online community members, however, is not always used as intended. It is quite often used for unauthorized purposes, hence an invasion of privacy. We discuss known ways we give off vital personal information while online in social networks. We further discuss ways to protect personal privacy. On the central point of ethical implications of life in the social network, we note that unlike in the traditional network, governance is not centralized, but community based with equally shared authority and responsibility by all users. But the mechanisms are not yet defined, and where they are being defined, it is still too early to say whether they are effective. The complexity, unpredictability, and lack of central authority is further enhanced by a virtual personality, anonymity and multiple personality. These three characteristics are at the core of the social and ethical problems in online social networks in particular and cyberspace in general; the larger and more numerous these communities become, the more urgent the ethical concerns become.

Learning Objectives

After reading this chapter, the reader should be able to

1. Understand computer networks.
 2. Understand social networks.
 3. Understand online social networks.
 4. Understand privacy issues affecting online social networks.
 5. Discuss privacy issues in social networks.
 6. Discuss ethical issues in online social networks.
 7. Discuss security issues in online social networks.
 8. Discuss the limitations of the legislation network to manage online social, privacy, and security issues.
-

11.1 Introduction

Because we intend to focus on online social networks in this chapter, it is imperative that the reader has a good grasp of network infrastructure upon which the online social network is anchored. So, we start this chapter with a brief introduction of the concepts of a computer network. Some knowledge of the computer network infrastructure will help the reader understand how these online social network services, discussed in Sect. 11.4.3, work. Thus, an introduction to computer networks follows.

11.2 Introduction to Computer Networks

A *computer network* is a distributed system consisting of loosely coupled computing elements and other devices. In this configuration, any two of these devices can communicate with each other through a communications medium. The medium may be wired or wireless. To be considered a communicating network, the distributed system must communicate based on a set of communicating rules called *protocols*. Each communicating device in the network must then follow these rules to communicate with others. A standard wired computer network resembles the network in Fig. 11.1.

Individually, network elements may own resources that are local or global. Such resources may be either software based or hardware based. If software, it may consist of all application programs and network protocols that are used to synchronize, coordinate, and bring about the sharing and exchange of data among the network elements. Network software also makes possible the sharing of expensive resources in the network. The hardware components of a computer network consist of a collection of nodes that include the end systems, commonly called *hosts*, and intermediate switching elements which include hubs, bridges, routers, and gateways.

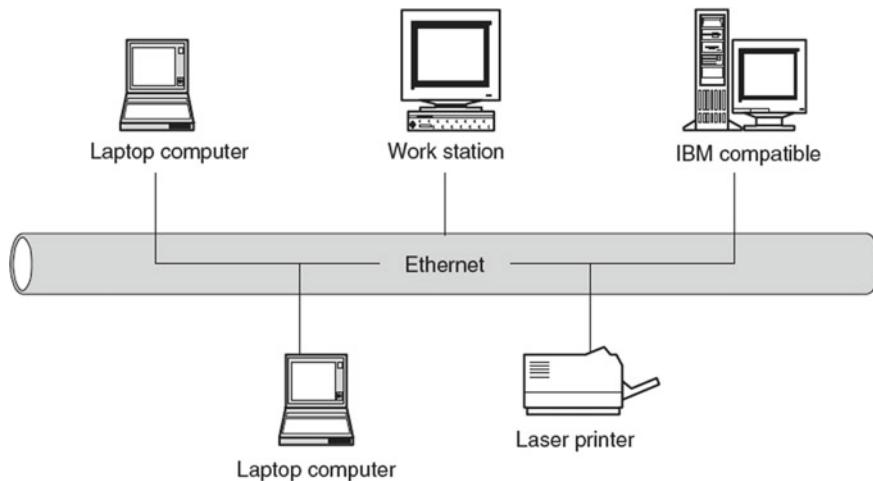


Fig. 11.1 A computer network

11.2.1 Computer Network Models

Several network configuration models are used in the design of computer networks, but the two most common are the centralized and distributed models shown in Figs. 11.2 and 11.3. In a centralized model, all computers and devices in the network are connected directly to a central computer through which they can interconnect to each other. This central computer, commonly called the master, must receive and forward all correspondence between any two or more communicating computers and devices. All other computers in the network are correspondently called dependent or surrogate computers. These surrogates may have reduced local resources, such as memory, and shareable global resources are controlled by the master at the center. The configurations are different, however, in the distributed network model, which consists of loosely coupled computers interconnected by a communication network composed of connecting elements and communication channels. However, in contrast to the centralized model, here the computers themselves may own their own resources locally or may request resources from a remote computer. Computers in this model are known by a string of names, including host, client, or node.

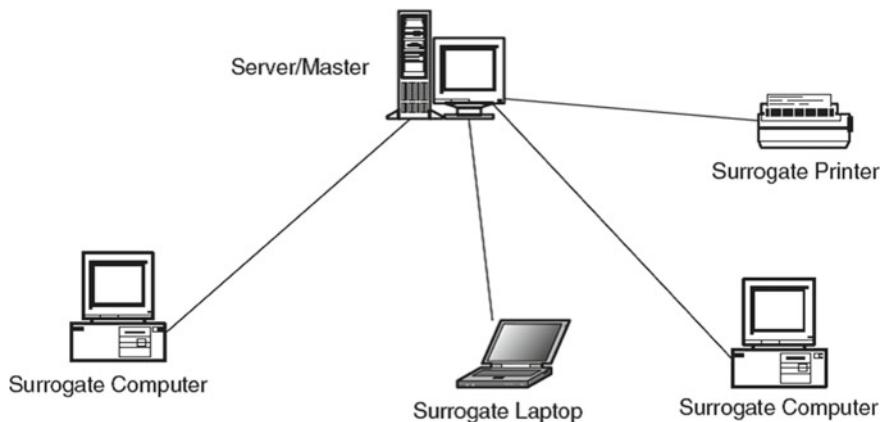


Fig. 11.2 A centralized network model

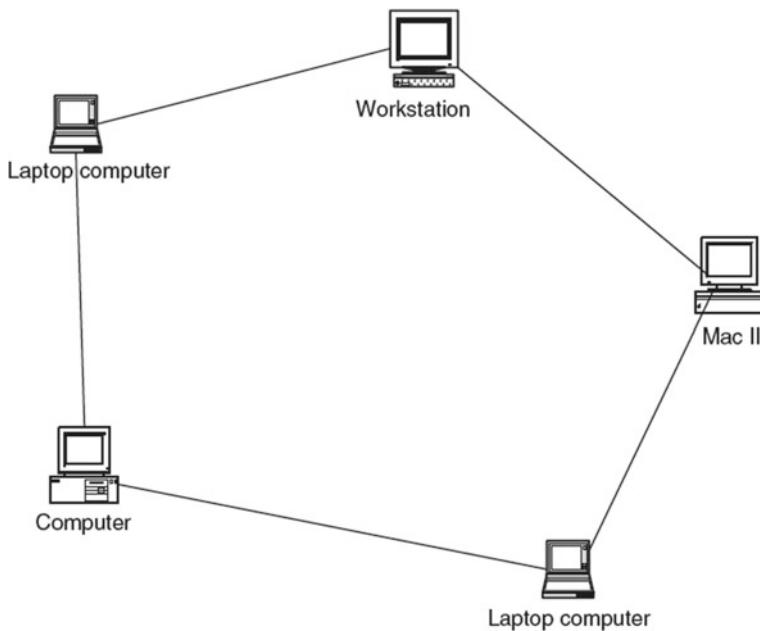


Fig. 11.3 A distributed network model

11.2.2 Computer Network Types

Computer networks, in any configuration centralized or distributed, come in different sizes depending on the number of computers and other devices in the network. The number of devices, computers or otherwise, in a network and the geographic

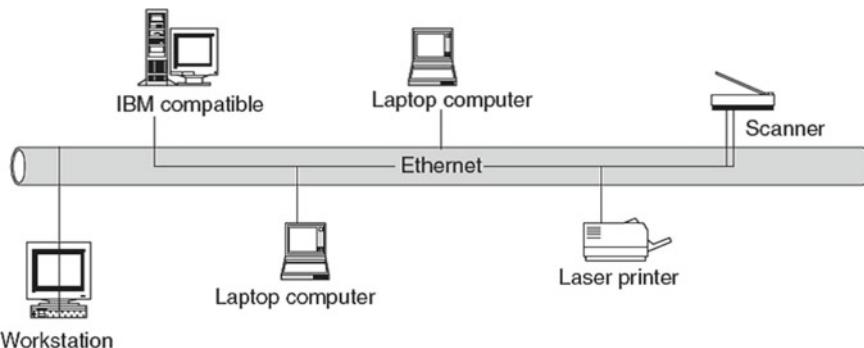


Fig. 11.4 A local area network (LAN)

area covered by the network determine the network type. There are, in general, three main network types: the local area network (LAN), a wide area network (WAN), and metropolitan area network (MAN).

11.2.2.1 Local Area Network

A LAN is a computer network with two or more computers or clusters of network and their resources connected by a communication medium sharing communication protocols and confined in a small geographic area such as a building, one floor of a building, or a few adjacent buildings. In a LAN, all network elements are in close proximity, which allows the communication links to maintain a higher speed and quality of data movement. Figure 11.4 shows a LAN.

11.2.2.2 Wide Area Network

A WAN is a computer network including one or more clusters of network elements and their resources, but in contrast to the LAN its configuration is not confined to a small geographic area: it can spread over a wide geographic area such as a region of a country, or across the whole country, several countries, or the entire globe, as does the Internet, which helps in distributing network services and resources to a wider community. Figure 11.5 shows a WAN.

11.2.2.3 Metropolitan Area Network

A metropolitan area network (MAN) is an unusual, and less often used, type of network that is intermediate between a LAN and a WAN. It covers a slightly wider area than the LAN but not so wide as to be considered a WAN. Civic networks that cover a city or part of a city are a good example of a MAN.

11.2.2.4 Mesh Network

A mesh network topology allows multiple access links between network elements, differing from other types of network topologies. The multiplicity of access links between network elements offers an advantage in network reliability because when

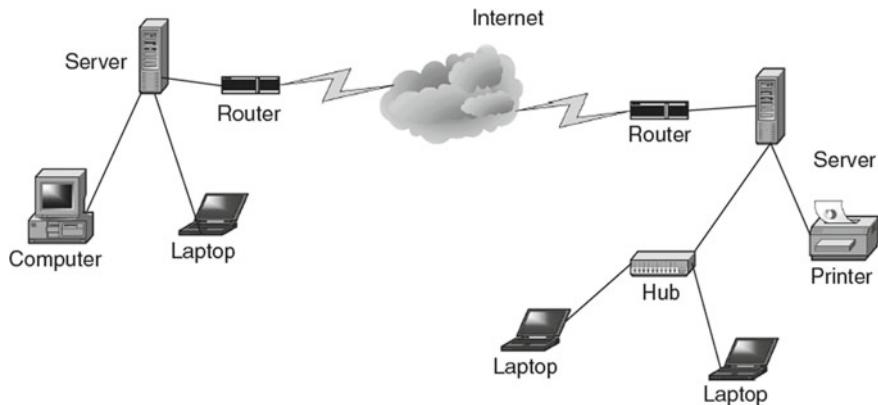


Fig. 11.5 A wide area network (WAN)

one network element fails, the network does not cease operations; it simply finds a bypass to the failed element, and the network continues to function. The mesh network topology is most often applied in metropolitan area networks (MANs), also known as civic networks, that cover a city or part of a city. Figure 11.6 shows a mesh network.

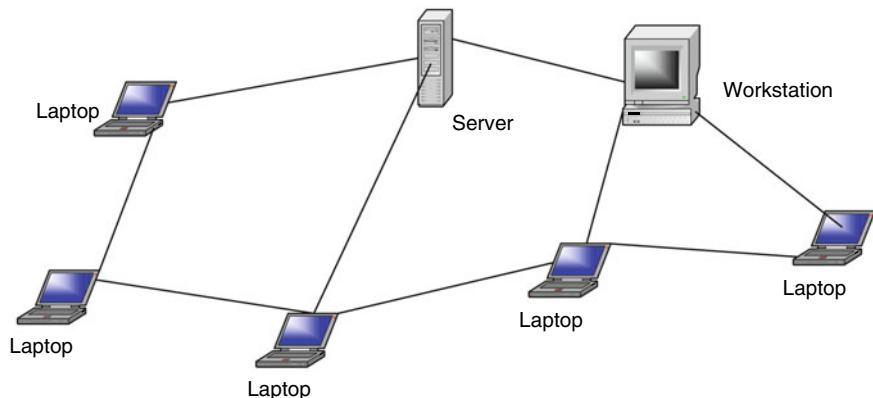


Fig. 11.6 Mesh network

11.3 Social Networks (SNs)

A *social network* is a theoretical network in which each node is an individual, a group, or an organization that independently generates, captures, and disseminates information and also serves as a relay for other members of the network. Individual nodes must collaborate to propagate the information in the network. The links between nodes represent relationships and social interactions among individuals, groups, organizations, or even entire societies.

The concept of social networking is not new. Sociologists and psychologists have been working with and analyzing social networks for generations. In fact, social networks have been in existence since the beginning of mankind. Prehistoric man formed social networks for different reasons including security, access to food, and social well-being.

Social networks begin with an individual reaching out to another individual or group for a social relationship of sorts that snowballs into a mesh of social relationships connecting many individuals or groups. In general, social networks come in all sizes and are self-organizing, complex, and agile depending on the nature of relationships in its links. As they grow in size, social networks tend to acquire specific elements and traits that make them different from one another. These traits become more apparent as the network increases in size. The type of social interactions, beliefs, and other traits usually limit the size of the social network. It is important to note that as the social network becomes large, it tends to lose the nuances of a local system; hence, if certain qualities of the network properties are needed, it is better to keep the size under control. Figure 11.6 illustrates three stages of development of a social network as it grows (Fig. 11.7).

11.4 Online Social Networks (OSNs)

Online social networks (OSNs) are social networks with underlining electronic communication infrastructure links enabling the connection of the interdependencies between the network nodes. The discussion in this chapter focuses on these OSNs. In particular, we focus on two types of online social networks:

- The traditional OSNs such as Facebook and MySpace: many of these can be accessed via mobile devices without the capability of handling mobile content.
- The mobile OSNs (mOSNs): these which are newer OSNs that can be accessed via mobile devices and can handle the new mobile context.

The interdependency between nodes in the OSNs supports social network services among people as nodes. These interdependencies as relationships among people participating in the network services define the type of OSNs.

11.4.1 Types of Online Social Networks

The growth of the OSNs over the years since the beginning of digital communication evolved through several types. Let us review the most popular types using a historical chronology.

Chat Network The chat network was born from the digital chatting anchored on a *chat room*. The chat room was, and still is, a virtual room online where people “gather” just to chat. Most chat rooms have open access, policies meaning that anyone interested in chatting or just reading others’ chats may enter the chat room. People can “enter” and “exit” any time during the chats. At any one time several threads of the public chats may be going on. Each individual in the chat room is given a small window on his or her communication device to enter a few lines of chat contributing to one or more of the discussion threads. This communication occurs in real time, and whatever one submits to the chat room can be seen by anyone in the chat room. Chat rooms also have a feature wherein a participating individual can invite another individual currently in the public chat room into a private chat room where the two can continue with limited “privacy.” To be a member of the chat room you must create a user name by which the members of the chat room will know you. Frequent chatters will often become acquaintances

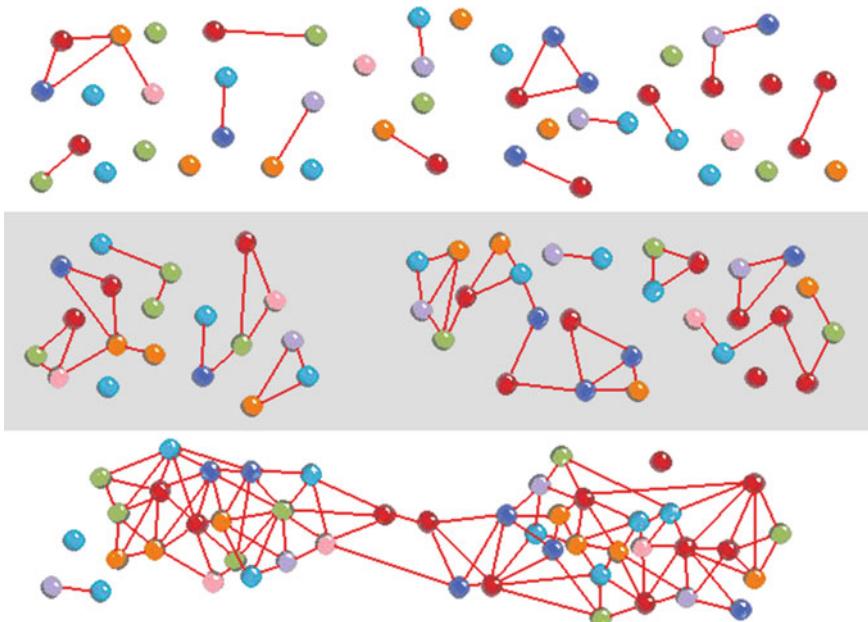


Fig. 11.7 Social network self-organizing configurations (http://en.wikipedia.org/wiki/File:Network_self-organization_stages.png)

based on user names. Some chat room software allows users to create and upload their profiles so that users can know more about you via your profile.

Although chat rooms by their own nature are public and free for all, some are monitored for specific compliance, based usually on attributes such as topics under discussion.

With the coming of more graphically based online services, the use of chat rooms is becoming less popular, especially among young people.

Blog Network Another online social network is the bloggers network. “Blogs” are nothing more than people’s online journals. Avid bloggers keep diaries of daily activities. These diaries sometimes are specific, on one thread of interest to the blogger, or a series of random logs of events during a specific activity. Some blogs are comment on specific topics. Some bloggers have a devoted following depending on the issues.

Instant Messaging Network (IMN) The IMNs support real-time communication between two or more individuals. As in chat rooms, each participant in the IMN must have a user name. To IM an individual, one must know that individual’s username or screen name. The initiator of the IM is provided with a small window to type the message and the recipient is also provided with a similar window to reply to the message. The transcript of the interchange is kept scrolling up both users’ screens. Unlike the chat room, however, these exchanges of short messages are private. As in chat networks, some IMNs allow users to keep profiles of themselves.

Online Social Networks (OSNs) The OSNs are a combination of all the network types we have already discussed and other highly advanced online features with advanced graphics. These social networks include Facebook, Twitter, MySpace, Friendster, YouTube, Flickr, and LinkedIn. As these networks grew from those we have already seen, many of the features of these networks are those we have already discussed. For example, users in these networks can create profiles that include their graphics and other enclosures and upload them to their network accounts. They must have a username or screen name. Also communication, if desired, can occur in real time as if one is using chat or IM capabilities. In addition to real time, these networks also give the user the delayed and archiving features so that the users can store and search for information. Because of these additional archival and search capabilities, network administrators have fought with the issues of privacy and security of users, as we see later in this chapter. As a way to keep user data safe, profiles can be set to a private setting, thus limiting access to private information by an authorized user.

11.4.2 Online Social Networking Services

An online social networking service is an online service accessible via any Internet-enabled device with the goal of facilitating computer-mediated interaction

among people who share interests, activities, backgrounds, or real-life connections. Most online social network services consist of the following elements:

- User Profile
- Social or business links of interests
- Additional services.

Currently, the most popular online social network services have categories that range among those based on friends, music and movies, religion, business, and many other interests. A sample of the current services in each of these categories follow.

- General and Friends-Based Social Networks
 - Facebook
 - MySpace
 - Hi5
- Movie and Music Social Networks
 - LastFM
 - Flixster
 - iLike
- Mobile Social Networks
 - Dodgeball
 - Loopt
 - Mozes
- Hobby and Special Interest Social Networks
 - ActionProfiles
 - FanIQ
- Business Social Networks
 - LinkedIn
 - XING
 - Konnects
- Reading and Books Social Networks
 - GoodReads
 - Shelfari
 - LibraryThing.

11.4.3 The Growth of Online Social Networks

OSNs have blossomed as the Internet exploded. The history and the growth of OSNs have mirrored and kept in tandem with the growth of the Internet. At

the infancy of the internet, computer-mediated communication services such as Usenet, ARPANET, LISTSERV, and bulletin board services (BBS) helped to start the growth of the current OSNs. Let us now see how these contributed to the growth of OSNs.

BITNET was an early world leader in network communications for the research and education communities and helped lay the groundwork for the subsequent introduction of the Internet, especially outside the U.S. [1]. BITNET and Usenet were invented around the same time in 1981 by Ira Fuchs and Greydon Freeman at the City University of New York (CUNY); both were “store-and-forward” networks. BITNET was originally named for the phrase “Because It’s There Net,” later updated to “Because It’s Time Net” [2]. It was originally based on IBM’s VNET *email* system on the IBM virtual machine (VM) mainframe operating system, but it was later emulated on other popular operating systems such as DEC VMS and Unix. What made BITNET so popular was its support of a variety of mailing lists supported by the LISTSERV software [3].

BITNET was updated in 1987 to BITNET II to provide a higher bandwidth network similar to the NSFNET. However, by 1996, it was clear that the Internet was providing a range of communication capabilities that fulfilled BITNET’s roles, so CREN ended their support and the network slowly faded away [3].

Bulletin Board Services (BBS) A BBS is a piece of software running on a computer allowing users on computer terminals far away to log in and access the system services such as uploading and downloading files and reading the news and contributions of other members through e-mails or public bulletin boards. In “Electronic Bulletin Boards, A Case Study: The Columbia University Center for Computing Activities,” Asteroff [4] reports that the components of computer conferencing that include private conferencing facilities, electronic mail, and electronic bulletin boards started earlier than the electronic bulletin board (BBS). Asteroff writes that the concept of an electronic bulletin board began about 1976 through ARPANET at schools such as the University of California at Berkeley, Carnegie-Mellon, and Stanford University. These electronic bulletin boards were first used in the same manner as physical bulletin boards, that is, help wanted, items for sale, public announcements, etc. Electronic bulletin boards soon became, because of the ability of the computer to store and disseminate information to many people in text form, a forum for user debates on many subjects. In its early years, BBS connections were made via telephone lines and modems. The cost of using them was high, so they tended to be local. As the earlier form of the World Wide Web, BBS use receded as the World Wide Web grew.

Listserv Started in 1986 as automatic mailing list server software that broadcast e-mails directed to it to all on the list, the first Listserv was conceived by Ira Fuchs from BITNET and Dan Oberst from EDUCOM (later EDUCAUSE), and implemented by Ricky Hernandez, also of EDUCOM, to support research mailing lists on the *BITNET* academic research network [5].

By the year 2000, Listserv was running on computers around the world, managing more than 50 thousand lists, with more than 30 million subscribers, delivering more than 20 million messages a day over the Internet [5].

Other Online Services As time went on and technology improved, other online services supplemented, and always improved on, the services of whatever was in use. Most of the new services were commercially driven, and most of them were moving toward, and are currently on, the Web. These services, including news, shopping, and travel reservations, were the beginning of the web-based services we are enjoying today. As they were commercially driven, they were mostly offered by ISPs such as AOL, Netscape, and Microsoft. As the Internet grew, millions of people flocked onto it, and the web and services started moving away from ISP to fully fledged online social network companies such as Facebook, Flickr, Napster, LinkedIn, and Twitter.

11.5 Ethical and Privacy Issues in Online Social Networks

Privacy is a human value consisting of a set of rights including solitude, the right to be alone without disturbances; anonymity, the right to have no public personal identity; intimacy, the right not to be monitored; and reserve, the right to control one's personal information, including the dissemination methods of that information. As humans, we assign a lot of value to these four rights. In fact, these rights are part of our moral and ethical systems. With the advent of the Internet, privacy has gained even more value as information has gained value. The value of privacy comes from its guardianship of the individual's personal identity and autonomy.

Autonomy is important because humans need to feel that they are in control of their destiny. The less personal information people have about an individual, the more autonomous that individual can be, especially in decision making. However, other people will challenge one's autonomy depending on the quantity, quality, and value of information they have about that individual. People usually tend to establish relationships and associations with individuals and groups that will respect their personal autonomy, especially in decision making.

As information becomes more imperative and precious, it becomes more important for individuals to guard their personal identity. Personal identity is a valuable source of information. Unfortunately, with rapid advances in technology, especially computer and telecommunication technologies, it has become increasingly difficult to protect personal identity.

11.5.1 Privacy Issues in OSNs

Privacy can be violated, anywhere including in online social network communities, through intrusion, misuse of information, interception of information, and information matching [6]. In online communities, intrusion, as an invasion of privacy, is

a wrongful entry, a seizing, or acquiring of information or data belonging to other members of the online social network community. Misuse of information is all too easy. While online, we inevitably give our information to whomever asks for it to obtain services. There is nothing wrong with collecting personal information when it is authorized and is going to be used for a legitimate reason. Information routinely collected from online community members, however, is not always used as intended. It is quite often used for unauthorized purposes, and hence is an invasion of privacy. As commercial activities increase online, there is likely to be stiff competition for personal information collected online for commercial purposes. Companies offering services on the Internet may seek new customers by either legally buying customer information or illegally obtaining it through eavesdropping, intrusion, and surveillance. To counter this, companies running these online communities must find ways to enhance the security of personal data online.

As the number and membership in online social networks skyrocketed, the issues of privacy and security of users while online and the security of users' data while offline have taken center stage. The problems of online social networking have been exacerbated by the already high and still growing numbers, especially of young people, who pay little to no attention to privacy issues for themselves or others. Every passing day, there is news about and growing concerns over breaches in privacy caused by social networking services. Many users are now worried that their personal data are being misused by the online service providers. All these privacy issues can be captured as follows [7]:

- Sharing of personal information with all OSN users:
 - Users in the network give out too much personal information without being aware who might wrongly use that information. Sexual predators are known to use information from teenagers on these networks. Currently, many of the OSNs are working with law enforcement to try to prevent such incidents [6]. Information such as street address, phone number, and Instant Messaging names are routinely disclosed to an unknown population in cyberspace.
 - Ease of access to OSNs. Currently it is very easy for anyone to set up an account on any one of these networks with no requirements for specific identifications, which can lead to identity theft or impersonation [6].
 - Privacy threats result from placing too much personal information in the hands of large corporations or governmental bodies, allowing a profile to be produced on an individual's behavior on which decisions may be made that are detrimental to that individual [6].
 - Updating profiles with current activities poses a great threat, for example, updating your profile informing people of your whereabouts.
- Lack of precise rules by the OSNs on who should use which data.
- Leakage of private information to third parties:
 - On many of these networks, information altered or removed by a user may in fact be retained or passed to third parties [6].
- Interlinkages in OSNs. In their paper “(Under)mining Privacy in Social Networks,” Monica Chew, Dirk Balfanz, and Ben Laurie of Google, Inc. point to

three distinct areas where the highly interlinked world of social networking sites can compromise user privacy [1]:

- Lack of control over activity streams: An *activity stream*, according to the authors, is a collection of events associated with a single user including changes a user makes to his or her profile page, the user adding or running a particular application on the social networking site, news items shared, or communication with friends. Activity streams may compromise a user's privacy in two ways:

A user may not be aware of all the events that are fed into their activity streams, in which case the user lacks control over those streams.

A user may not be aware of the audience who can see their activity streams, in which case the user lacks control over that audience.

- Unwelcome linkage: *Unwelcome linkage* occurs when links on the Internet reveal information about an individual that they had not intended to reveal. Unwelcome linkage may occur wherever graphs of hyperlinks on the World Wide Web are automatically created to mirror connections between people in the real world. Maintaining separation of individual activities and different personae is important in OSNs.
- Deanonymization of users through merging of social graphs. OSN sites tend to extract a lot of personally identifiable information from people such as birth date and address. With this information, it is possible to de-anonymize users by comparing such information across social networking sites, even if the information is partially obfuscated in each OSN.

As the growth in online social networks continues unabated, the coming in the mix of the smart mobile devices is making the already existing problems more complex. These new devices are increasing the number of accesses to OSNs and increasing the complexity of the privacy issues, including, in addition to those already noted in the traditional (OSNs) [8], the following:

- The presence of a user. In contrast to the most traditional OSNs, where users were not automatically made aware of the presence of their friends, most mobile OSNs (mOSNs) now allow users to indicate their presence via a “check-in” mechanism, whereby a user establishes their location at a particular time. According to Krishnamurthy and Wills [8], the indication of presence allows their friends to expect quick response, and this may lead to meeting new people who are members of the same mOSN. Although the feature of automatic locate by oneself is becoming popular, it allows leakage of personal private information along two tracks: the personal information that may be sent and the destination to which it could be sent.
- Location-based tracking system (LTS) technologies that are part of our mobile devices. This is a feature that is widespread in the mobile environment. However, users may not be aware that their location can be made known to friends and friends of friends who are currently online on this mOSN, to their friends in

other mOSNs, and to others, which may lead to leakage of personal information to third parties.

- Interaction potential between mOSNs and traditional OSNs. According to Krishnamurthy and Wills [8], such connections are useful to users who, while interacting with an mOSN, can expect some of their actions to show up on traditional OSNs and be visible to their friends there. However, much of their personal information can leak to unintended users of both the traditional OSNs and the mOSNs.

In addition to almost free access to a turn of personal data on OSNs, there is also a growing threat to personal data ownership; for example, who owns the data that were altered or removed by the user which may fact be retained and/or passed to third parties. This danger was highlighted when, in June 2011, a 24-year-old Austrian law student, Max Schrems, asked Facebook for a copy of all his personal data. Facebook complied, sending him a CD containing 1200 pages of data, including his likes, “friend” and “defriend” history, and chat logs. But before that, Schrems had deleted some of the data returned to him from his profile, yet Facebook had retained his information. Of course Schrems filed 22 individual claims against Facebook for €100,000 (\$138,000) for retaining data deleted by users, in the case *Europe v. Facebook* [9].

Fortunately, users are beginning to fight for their privacy to prevent their personal details from being circulated far widely than they intended them to be. For example, take Facebook’s 2006 News Feed and Mini Feed features, designed to change what Founder and CEO Mark Zuckerberg called Facebook’s old “encyclopedic interface,” where pages mostly just list information about people, to the current stream of fresh news and attention content about not only the user but also the user’s friends and their activities [10]. The first, News Feed, brought to the user’s home page all new activities on all friends and associate links, including new photos posted by friends, relationship status changes, people joining groups, and many others, thus enabling the user to get an abundance of information from every friend’s site every day.

Although these features adhered to Facebook’s privacy settings, meaning that only people whom a user allowed to view the data were able to see them, this still generated a firestorm from users across the world. More than 700,000 users signed an online petition demanding the company discontinue the feature, stating that this compromised their privacy [11]. Much of the criticism of The News Feed was that it gave out too much individual information.

Online social networks, just like their predecessor cyberspace communities, are bringing people together with no physical presence to engage in all human acts that traditionally have taken place in a physical environment that would naturally limit the size of the audience and the amount of information given at one time. As these cyber-communities are brought and bound together by a sense of belonging, worthiness, and the feeling that they are valued by members of the network, they create a mental family based on trust, the kind of trust you would find in a loving family. However, because these networks are boundary less, international

in nature, they are forming not along well-known and traditional identifiers such as nationalities, beliefs, authority, and the like, but by common purpose and need with no legal jurisdiction and no central power to enforce community standards and norms.

11.5.2 Strengthening Privacy in OSNs

As more and more people join OSNs and now the rapidly growing mOSNs, there is a growing need for more protection to users. Chew et al. suggest the following steps are needed [1]:

- Both OSN and mOSN applications should be explicit about which user activities automatically generate events for their activity streams
- Users should have control over which events make it into their activity streams and be able to remove events from the streams after they have been added by an application
- Users should know who the audience of their activity streams is and should also have control over selecting the audience of their activity streams
- Both OSN and mOSN applications should create activity stream events that are in sync with user expectation.

Other suggestions that may help in this effort:

- Use secure passwords.
- User awareness of the privacy policies and terms of use for their OSNs and mOSNs.
- Both OSNs and mOSNs providers should devise policies and enforce existing laws to allow some privacy protection for users while on their networks.

11.5.3 Ethical Issues in Online Social Networks

Online social communities are far from the traditional physical social communities with an epicenter of authority in which every member pays allegiance to the center with a shared sense of responsibility. This type of community governance with no central command, but an equally shared authority and responsibility, is new, and a mechanism needs to be in place and must be followed to safeguard every member of the community. But these mechanisms are not yet defined, and where they are

being defined, it is still too early to say whether they are effective. The complexity, unpredictability, and lack of central authority is further enhanced by these aspects:

- *Virtual personality*: You know their names, their likes and dislikes. You know them so well that you can even bet on what they are thinking, yet you do not know them at all. You cannot meet them and recognize them in a crowd.
- *Anonymity*: You work with them almost every day. They are even your friends; you are on a first-name basis, yet you will never know them. They will forever remain anonymous to you and you to them.
- *Multiple personality*: You think you know them, but you do not because they are capable of changing and mutating into other personalities. They can change into as many personalities as there are issues being discussed. You never know which personality you are going to meet next.

These three characteristics are at the core of the social and ethical problems in online social networks in particular and cyberspace in general; the larger and more numerous these communities become, the more urgent the ethical concerns become. With all these happening in online social network, the crucial utilitarian question to ask is what is the best way and how can we balance the potential harms and benefits that can befall members of these online social networks and how if possible to balance these possibilities. Of late, the news media has been awash with many of these online ills and abuses, and the list is growing:

(i) *Potential for misuse*

Online social networks offer a high degree of freedom that is being misused by a growing number of users. Cases abound of these incidents with tragic endings including suicide, especially in young people.

(ii) *Cyberbullying, cyberstalking, and cyber-harassment*

Cyberbullying, cyberstalking, and electronic harassment are relatively common occurrences and can often result in emotional trauma for the victim. They are, unfortunately, becoming a common form of abuse on online social network sites such as Facebook and MySpace, especially to youth. Cyberbullying is defined as use of Internet services and mobile technologies such as web pages and discussion groups, as well as instant messaging or SMS text messaging, with the intention of harming another person. Cyberstalking or cyber harassment on the other hand is defined as the use of the Internet or other electronic means to stalk or harass an individual, a group of individuals, or an organization. It may include false accusations, monitoring, making

threats, identity theft, damage to data or equipment, the solicitation of minors for sex, or gathering information for harassment [12].

Because of the nature of cyberspace's telepresence, anonymity, lack of allegiance of users, and the non-existence of central governance, there are no limitations as to what individuals can post when online and how toxic those posts can be. Individuals, therefore, take it as if they are given the power to post offensive remarks or pictures that could potentially cause a great amount of emotional pain, oftentimes leading to teen suicide. Cases are growing of these kinds of activities, some of which are tragic. Bullying statistics show that cyberbullying is a serious problem and alarmingly common among adolescents and teens. According to cyberbullying statistics from the i-SAFE foundation [13]:

More than half of adolescents and teens have been bullied online, and about the same number have engaged in cyberbullying.

More than one in three young people have experienced cyberthreats online.

More than 25% of adolescents and teens have been bullied repeatedly through their cell phones or the Internet.

Considerably more than half of young people do not tell their parents when cyberbullying occurs.

As these statistics indicate, the number of teen suicides attributable to cyberbullying is on the rise.

(iii) *Risk for child safety*

Problems with online social networks are not only limited to misuse of the sites and cyberbullying, they also include real threats to children whether cyberbullied or not. There is growing exploitation of children in online social networks. The latest figures show that about 1 million children under 16 years of age use Bebo, and 600,000 minors are on MySpace [14]. With these numbers, the potential for child abuse online is growing. The networking sites say they are making it possible for users to report abuse, although those reports usually go to the site administrators rather than the authorities. Governments around the world are taking steps, at least to better understand the problem and find some solutions.

Discussion Topics

How do we balance these harms and benefits, reducing one and increasing the possibility of the other?

How do we protect individuals and how do we handle the issue of consent?

(iv) *Psychological effects of online social networking*

The rise in the use and membership of online social networking has resulted in a dramatic rise not only in the numbers of online social networks but also in

the number of users. With the rise in the number of users is also a rise in the number of users with problems. More and more people, especially teenagers, are spending an excessive amount of time on the Internet in general and on social networking sites in particular, which has led researchers to classify Internet addiction as a new clinical disorder [15].

According to Neville Misquittaa in “Psychiatry and Society in Pune,” the most common predictors of excessive use of social networking are these [16]:

Extroverted and unselfconscious individuals spend more time on social networking sites and their usage tends to be addictive.

Shy people also like Facebook and spend more time on it. However, they have few Facebook “friends.”

Narcissistic personalities also have high levels of online social activity.

They are recognized online by the quantity of their social interactions, their main photo self-promotion, and the attractiveness of their main photo.

(v) *Free speech*

What types of speech are protected once one is in an online social network? Although the National Labor Relations Act protects workers from being fired for “protected concerted activity,” which prevents workers from being fired for collective action while allowing companies the right to fire workers for individual actions they take against the company, when it comes to online social networking, the issues are still murky and there is still uncertainty as to the boundaries of what types of speech is protected in online social networks. This fuzziness is illustrated by the Pembroke Pines Charter High School case in which Katherine Evans, who was a senior at Pembroke Pines Charter High School in Florida in 2007 when she created a group on Facebook called, “Ms. Sarah Phelps is the worst teacher I’ve ever met.”

Peter Bayer, the principal of Pembroke Pines High, suspended Evans for 3 days and removed her from her Advanced Placement classes for violating the school’s rules against “cyberbullying” and “harassment” of a staff member, according to court documents. Evans sued the principal in his individual capacity, alleging that her First Amendment free speech and 14th Amendment due process rights were violated.

In a ruling that followed, in **Bayer v. Evans**, U.S. Magistrate Judge Barry L. Garber of Miami declined Evans’s request for an injunction barring the principal from keeping the student’s discipline in school records. However, the judge denied qualified immunity for Bayer, holding that Evans’s speech was protected under the First Amendment and that the principal should have known he was violating a clearly established right by disciplining Evans [17].

This ruling, as do other recent rulings, speaks volumes about the ethics of social networking and schools and is indicative of the haziness of the legal boundaries of free speech in online social networks.

Discussion Topics

Should teachers be allowed to befriend students on sites such as Facebook?
Should students blog about their teachers while on an online social network?

11.6 Security and Crimes in Online Social Networks

Online crimes, in tandem with the growth of computing and telecommunication technologies, are one of the fastest growing types of crimes and they pose the greatest danger to online communities, e-commerce, and the general public. An *online crime* is a crime like any other crime, except that in this case, the illegal act must involve either an Internet-enabled electronic device or computing system either as an object of a crime, an instrument used to commit a crime, or a repository of evidence related to a crime. Also, online crimes are acts of unauthorized intervention into the working of the telecommunication networks and/or the sanctioning of authorized access to the resources of the computing elements in a network that lead to a threat to the system's infrastructure or cause a significant property loss. The International Convention of Cyber Crimes and the European Convention on Cyber Crimes both list the following crimes as online crime [2]:

- Unlawful access to information
- Illegal interception of information
- Unlawful use of telecommunication equipment
- Forgery with use of computer measures
- Intrusions of the Public Switched and Packet Network
- Network integrity violations
- Privacy violations
- Industrial espionage
- Pirated computer software
- Fraud using a computing system
- Internet/e-mail abuse
- Using computers or computer technology to commit murder, terrorism, pornography, and hacking.

11.6.1 Beware of Ways to Perpetrate Crimes in Online Social Networks

As we pointed out in Chap. 9, if we have to fight online crimes, we have to first learn how they are perpetrated. Earlier, we noted that online crimes are defined in a variety of ways, reflecting the many different ways these crimes are perpetrated.

Some of the most common ways are through system penetration and denial of service attacks.

11.6.1.1 System Penetration

System penetration is the most widely used approach to committing online crimes. A system penetration is a process of gaining unauthorized access to a protected system's resources: the system may be automated or not. Penetration attacks always compromise the integrity of the resources of a system. Most penetration attacks are not accidental; they are preplanned and proceed with a coordinated reconnaissance. The goal of the reconnaissance is to acquire the following lead information on the targeted system:

- IP addresses of all hosts or selected hosts in the victim network
- Accessible UDP and TCP port numbers
- The type of operating system(s) used on all hosts or selected hosts in the network.

There are two types of reconnaissance: passive and active. In a *passive reconnaissance*, the attacker gathers freely available system information, mostly from an open source. A typical passive reconnaissance can include physical observation of buildings housing the system, dumpster diving near the target system, collecting discarded papers and system computer equipment in an attempt to find equipment or data that may include personal identifying data such as usernames and passwords which will lead them to gain access to the company system. It also includes using other information-gathering techniques like eavesdropping on employee conversations, social engineering, and packet sniffing. Some common sources and tools used when looking for open source information legally include these [3]:

- A company website
- Electronic data gathering, analysis, and retrieval (EDGAR) filings (for publicly traded companies)
- Network news transfer protocol (NNTP) USENET newsgroups
- User group meetings
- Business partners
- Dumpster diving
- Social engineering.

Active reconnaissance on the other hand involves collecting information about a target system by probing that system or neighboring systems. A typical active reconnaissance involves port scanning to discover vulnerable ports through which to enter the system, probing firewalls and system routers to find ways around them,

and other methods. Some of the tools used in active host reconnaissance include these:

- NSLookup/Whois/Dig lookups
- SamSpade
- Visual Route/Cheops
- Pinger/WS_Ping_Pro.

11.6.1.2 Distributed Denial of Service

Another approach used by perpetrators of online crimes is the *denial of service*, an interruption of service of the target system. This interruption of service occurs when the target system is made either unavailable to users through its disabling or destruction. Denial of service can also be caused by intentional degradation or blocking of computer or network resources. These denial of service attacks are commonly known as *distributed denial of service* (DDoS) attacks because they attack hosts in a network.

Similar to penetration attacks (e-attacks), DDoS attacks can also be either local, where they can shut down LAN computers, or global, originating thousands of miles away on the Internet. Attacks in this category include these [2]:

- *IP spoofing*. A forging of an IP packet address such as the source address, which causes the responses from the destination host to be misdirected, thus creating problems in the network. Many network attacks are a result of IP spoofing.
- *SYN flooding*. Using a three-way handshake protocol to initiate connections between a malicious (spoofed) source nodes and flood the target node with too many connection requests, thus overwhelming it and bringing it down.
- *Smurf attack*. The intruder sends a large number of spoofed ICMP Echo requests to broadcast IP addresses. Hosts on the broadcast multicast IP network then respond to these bogus requests with reply ICMP Echo, significantly multiplying the number of reply ICMP Echoes to the hosts with spoofed addresses.
- *Buffer overflow*. The attacker floods a carefully chosen field such as an address field with more characters than it can accommodate. These excessive characters, usually executable malicious code, when executed may cause havoc in the system, effectively giving the attacker control of the system.
- *Ping of death*. The attacker sends IP packets that are larger than the 65,536 bytes allowed by the IP protocol, knowing that many network operating systems cannot handle this, leading to the possible freezing or eventual system crash.
- *Land.c attack*. In which the land.c program sends TCP SYN packets whose source and destination IP addresses and port numbers are those of the victims.
- *Teardrop.c*. In which the attacker causes a fragmentation of TCP packets to exploit the reassembling process that may lead to the victim to crash or hang.

- *Sequence number sniffing.* In which the intruder takes advantage of the predictability of sequence numbers used in TCP implementations to sniff the next sequence number to establish legitimacy.

11.6.2 Defense Against Crimes in Online Social Networks

Although there are systems that are randomly attacked, most victim systems are preselected for attack. Because of this, we can defend systems against online attacks. An effective defense plan consists of prevention, detection, and analysis and response.

11.6.2.1 Prevention

Prevention is perhaps the oldest and probably the best defense mechanism against online crimes. However, prevention can only work if there is a strict security discipline that is effectively enforced and must include the following:

- A security policy
- Risk management
- Vulnerability assessment
- Use of strong cryptographic algorithms
- Penetration testing
- Regular audits
- Use of proven security protocols
- Legislation
- Self-regulation
- Mass education.

Let us discuss some of these. More details may be found in Sects. 5.3 and 8.3.

11.6.2.2 A Security Policy

A security policy is a critical and central document in an organization security effort that spells out in great detail how the organization manages risk, controls access to key assets and resources, and implements policies, procedures, and practices for a safe and secure environment [4]. A security policy usually also spells out what resources need to be protected and how the organization can protect such resources. It is a living document and sometimes controversial. There are as many opinions on the usefulness of security policies in the overall system security picture as there are security experts. However, security policies are still important in establishing an organization's security guidelines such as these:

- *Hardware and software acquisition and installations in the organization.* For example, if a functioning firewall is to be configured, its rule base must be based on a sound security policy.

- *User discipline.* All users in the organization that connect to a network, such as the Internet, must do so in conformity to the security policy.

A security policy is unique for each organization, covers a wide variety of topics, and serves several important purposes in the organization's security cycle. Because of this, the following carefully chosen set of basic steps must be established and carefully followed in the construction of a viable implementable and useful security policy:

- Determining the resources that must be protected and for each resource drawing a profile of its characteristics
- Determining, for each identified resource, from whom the resource must be protected
- Determining, for each identifiable resource, the type of threat and the likelihood of occurrence of such a threat
- Determining, for each identifiable resource, what measures are needed to give it the best protection
- Determining what needs to be audited
- Determining and defining acceptable use of system resources such as e-mail, News, and Web
- Considering how to implement and deploy security protocols such as encryption, access control, key creation, and distributions and wireless devices that connect on the organization's network
- Providing for remote access to accommodate workers on the road and those working from home, and also business partners who may need to connect to the organization's network via a VPN.

11.6.2.3 Vulnerability Assessment

As is risk assessment, vulnerability assessment is the process of identifying and quantifying vulnerabilities in a system. A *vulnerability* in a system is an exploitable weakness in the system. As we saw in Sect. 8.3, this is a two-part process; we need to first identify all system vulnerabilities and then develop strategies to mitigate the effects of these vulnerabilities. The rest of the steps usually taken are similar to those in Sect. 8.3.

11.6.2.4 Use of Strong Cryptographic Algorithms

Cryptography is a Greek word meaning “secret writing.” It was used to describe the art of secret communication. As shown in Figs. 4.1 and 11.8, cryptographic system consists of four essential components [2]:

- Plaintext: the original message to be sent
- A cipher: consisting of mathematical encryption and decryption algorithms
- Ciphertext: the result of applying an encryption algorithm to the original message before it is sent to the recipient

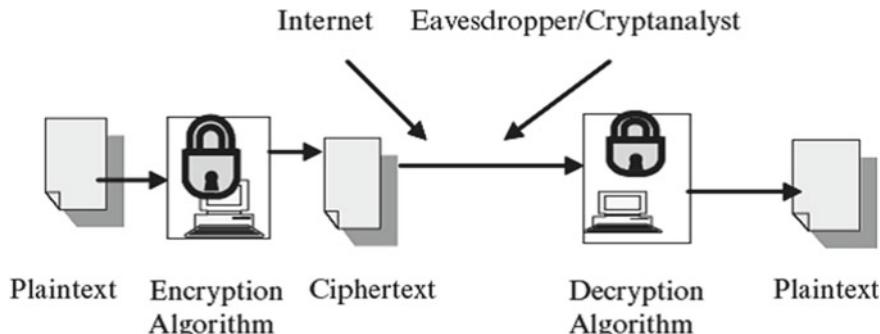


Fig. 11.8 Symmetrical encryption

- Key: a string of bits used by the two mathematical algorithms in encrypting and decrypting processes.

Cryptographic technologies are today being used increasingly to fight off massive invasion of individual privacy and security, to guarantee data integrity and confidentiality, and to bring trust in global e-commerce. In fact, cryptography has become the main tool for providing the needed digital security in the modern digital communication medium. Its popularity is a result of its ability to guarantee authorization, authentication, integrity, confidentiality, and non-repudiation in all communications and data exchanges in the new information society.

11.6.2.5 Penetration Testing

One of the core security techniques for safeguarding the security of an organization's system is a periodic penetration test of the system. The test may be outsourced for it to be more authentic, or it could be carried out inhouse, so long as one has competent personnel to do it. The process of penetration testing actively evaluates an organization's system resources and information in real time looking for design weaknesses, technical flaws, and vulnerabilities in the system: this can be done on a regular basis or with a scheduled timeframe. The possible outcomes of the test vary depending on the focus of the test.

Penetration testing may also focus on the security of information on the organization network by doing tests such as document grinding, privacy of information review, and intelligence scouting. If the organization supports wireless technology, this component must also be tested. No penetration testing can be complete without testing social engineering, communication within and outside the organization, and the physical security within the organization. Finally, physical testing may require testing access to the facilities, monitoring the perimeter and alarm systems, and an environment review.

11.6.2.6 Regular Security Audits

A penetration testing of an organization system is a focused look at the security holes in the system's resources such as firewalls and servers, whereas a security audit is a systematic, measurable, and quantifiable technical assessment of the organization security and the security of its system. Management usually requests security audits to gain knowledge and understand the security status of the organization's system. From the audit report, management may decide to upgrade the system through acquisition of new hardware and software. In "Conducting a Security Audit: An Introductory Overview," Hayes suggests that a security audit should answer the following questions [5]:

- Are passwords difficult to crack?
- Are there access control lists (ACLs) in place on network devices to control who has access to shared data?
- Are there audit logs to record who accesses data?
- Are the audit logs reviewed?
- Are the security settings for operating systems in accordance with accepted industry security practices?
- Have all unnecessary applications and computer services been eliminated for each system?
- Are these operating systems and commercial applications patched to current levels?
- How is backup media stored? Who has access to it? Is it up to date?
- Is there a disaster recovery plan? Have the participants and stakeholders ever rehearsed the disaster recovery plan?
- Are there adequate cryptographic tools in place to govern data encryption, and have these tools been properly configured?
- Have custom-built applications been written with security in mind? How have these custom applications been tested for security flaws?
- How are configuration and code changes documented at every level? How are these records reviewed and who conducts the review?

If genuine and trustful answers are given to many of these questions, a realistic security status of the organization's system emerges.

11.7 Proven Security Protocols and Best Practices in Online Social Networks

There are hundreds of security protocols to meet the needs of organizations trying to improve their system security. There are so many of them, some open source and others not, that they pose a problem to security professionals to choose a really good product. The security personnel must strive to come up with a list of the best protocols and best practices to suit the system. Some of these protocols include the following.

11.7.1 Authentication

Authentication is the process of validating the identity of someone or something. It uses information provided to the authenticator to determine whether someone or something is in fact who or what it is declared to be. The process usually requires one to present credentials or items of value to the authenticating agent to prove the claim of who one really is. The items of value or credential are based on several unique factors that show something you know, something you have, or something you are [2]:

Something you know. It may be something you mentally possess like a password, a secret word known by the user and the authenticator. This technique of authentication is cheap but has weaknesses, such as memory lapses.

Something you have. It may be any form of issued or acquired self-identification such as SecurID, Activcard, or any other forms of cards and tags. This authentication technique is slightly safer.

Something you are. These are individual physical characteristics such as voice, fingerprint, iris pattern, and other biometrics.

Besides these, there are other forms of authentication using a variety of authentication algorithms. These authentication methods can be combined or used separately, depending on the level of functionality and security needed. Among such methods are password authentication, public key authentication, anonymous authentication, and remote and certificate-based authentication.

11.7.2 Access Control

Access control is a process of determining how access to the system's potential resources can be provided to each of the system users. Because a system, especially a network system, may have thousands of users and resources, the management of access rights for every user per every object may become complex. Several control techniques and technologies have been developed to handle this problem; they include access control matrix, capability tables, access control lists, role-based access control, rule-based access control, restricted interfaces, content-dependent access control, and biometrics.

11.7.3 Legislation

Ever since the start of noticeable computer technology misuse, governments and national legislatures around the world have been enacting laws intended to curb the growth of these crimes. The report card on these legislations has been mixed. In some cases, legislation as a form of deterrent has worked and in others it has been a failure. However, we should not lose hope. Enforceable laws can be productive.

11.7.4 Self-regulation

Perhaps one of the most successful forms of deterrence has been self-regulation. A number of organizations have formed to advocate parents and teachers to find a way to regulate objectionable material from reaching the children. Also families and individuals, sometimes based on their morals and sometimes based on their religion, have made self-regulation a cornerstone of their efforts to stop the growing rate of online crimes.

11.7.5 Detection

Although it is easy to develop mechanisms for preventing online crimes, it is not so easy to develop similar or effective techniques and best practices to detect online crimes. Detecting online crimes constitutes a 24-h monitoring system to alert security personnel whenever something unusual (something with a nonnormal pattern, different from the usual pattern of traffic in and around the system) occurs. Detection systems must continuously capture, analyze, and report on the daily happenings in and around the network. In capturing, analyzing, and reporting, several techniques are used including intrusion detection, vulnerability scanning, virus detection, and other ad hoc methods.

11.7.6 Recovery

Recovery is a process preceded by a process of analysis, which involves taking as many data as possible gathered during the last intrusion and analyzing these for patterns that can be used in future for a response, for detection, and for prevention. Recovery requires the use of all available resources to first mitigate the problem in progress, then to recover whatever can be recovered and use it to build on new data in place of or to replace the destroyed data.

Exercises

1. What are the differences between online social networks and online communities?
2. Discuss the social problems of online social networks.
3. An ecosystem is a localized group of interdependent organisms together with the environment that they inhabit and on which they depend. How do you relate this to online social networks?
4. Discuss privacy issues that apply in your online social ecosystem.
5. Discuss five modern online crimes.
6. Discuss strategies that can be used to effectively eliminate (if possible) online social network crimes.

7. If you were to write a framework to prevent cybercrimes from online social networks and indeed from all online spaces, what would be in it?
8. Is cryptography all we need to secure computer network and protect information?
9. Why is cryptography failing to protect digital systems and information? What do we need to do?

References

1. *Bylaws for Internet Corporation for Assigned Names and Numbers* (ICANN). <https://www.icann.org/resources/pages/bylaws-2005-04-08-en>
2. J.M. Kizza, *Guide to Computer Network Security*, 3rd edn. (Springer, New York, NY, 2015)
3. D. Newman, A. Whitaker, Penetration Testing and Network Defense: Performing Host Reconnaissance. <http://www.ciscopress.com/articles/article.asp?p=469623&seqNum=2>
4. E. Tittel, Understanding Security Policies. <http://www.informit.com/articles/article.aspx?p=25041>
5. B. Hayes, Conducting a Security Audit: An Introductory Overview. <https://www.symantec.com/connect/articles/conducting-security-audit-introductory-overview>
6. J.M. Kizza, *Ethical and Social Issues in the Information Age*, 5th edn. (Springer, New York, NY, 2013)
7. R. Fox, News track: age and sex. *Commun. ACM* **43**(9), 9 (2000)
8. J.M. Kizza, *Computer Network Security and Cyberethics*, 3rd edn. (McFarland Publishers, Jefferson, North Carolina, 2011)
9. *Unavoidable Ethical Questions About Social Networking* (Mekkula Center for Applied Ethics, Santa Clara University). <http://www.scu.edu/ethics/publications/submitted/social-networking.html>
10. Number of websites surpasses 1 billion, continues to climb. AFP RELAXNEWS. The Daily News, 17 Sept 2014
11. *Communication from the Commission to the Council and the European Parliament* (Commission of the European Communities (Com2000)), 202
12. L. Gordon Crovitz, Information haves and have-nots. *The Wall Street Journal*, 22 Sept 2008. <https://www.wsj.com/articles/SB122204237577161317>
13. Evolving the High Performance Computing and Communications Initiative to Support the Nation's Information Infrastructure—Executive Summary. <http://www.nap.edu/search/?term=Evolving+the+High+Performance+Computing+and+Communications+Initiative+to+Support+the+Nation%20%80%99s+Information+Infrastructure%20%80%94Executive+Summary&x=0&y=0>
14. Wikipedia: Cyberstalking. <http://en.wikipedia.org/wiki/Cyberstalking>
15. <http://www.bullyingstatistics.org/content/cyber-bullying-statistics.html>
16. BBC News, <http://news.bbc.co.uk/2/hi/technology/7056264.stm>
17. K. Young, Internet addiction: the emergence of a new clinical disorder. *CyberPsychol. Behav.* **1**(3), 237–244



Virtualization, Virtual Reality and Ethics

12

Abstract

This chapter discusses the new developments and consequences of virtualization technology and its implications on our participation and how the technology informs our behavior based on our traditional moral and ethical values. In a more detailed way, we define virtualization as a process which embodies both abstraction and reconstruction and as it creates a sense of complete participants' immersion yet with autonomy of participants to vary their chosen new environments to suit individual likings. As defined, virtualization, therefore, conjures uncertainty and fear not of the environment but of the individual who partakes in the activities of the environment. Our discussion of both social and ethical issues that arise within and outside the environment focused on two types of virtualization: computing resources virtualization and virtual reality (VR). We note that virtualization is also bringing about easy creation of new human identities in the new virtual environments which makes authentication more difficult but at the same time creating unprecedented potential in self-creation and self-presentation. We note, however, that while these benefits may bring new opportunities and new potential that may empower individuals to new levels of creativity, these unparalleled opportunities of virtualization may come at a price to society.

Learning Objectives

After reading this chapter, the reader should be able to:

- Understand the value of ethics in virtual environments
- Identify and discuss the different types of virtualization
- Be able to difference between virtual reality and software virtualization
- Recognize the role ethics plays in artificial environments

- Identify and discuss credible safeguards to ensure privacy concerns for users of virtual environments
 - Understand and be able to debate the responsibilities of creators of virtual environments
 - Recognize and discuss the responsibilities of users in virtual environments.
-

12.1 Virtualization

Virtualization is a process through which one can create something that is there in effect and performance but in reality not there—that is, virtual. It is a physical abstraction of reality, a real phenomena such as a company’s computing resources like storage, network servers, memory, and others. It involves and absorbs participants into a virtual reconstruction of real-world entities into seemingly real images with corresponding in-depth information to turn these images into a high degree of realism. This process which embodies both abstraction and reconstruction creates a sense of complete participants’ immersion yet with autonomy of participants to vary their chosen new environments to suit individual likings. In other words, virtualization is a process that makes real entities, scenes, and events virtual mirror images of self; it is a virtualization of reality. In many ways it is a mediation of interaction through an electronic medium between humans and humans as well as between humans and machines [1].

12.2 Different Aspects of Virtualization

The immersion aspects of the virtualization process of its participants and the autonomy accorded to them give the virtualization process a wide range of the different aspects of real life that can be virtualized. These may include gaming, computing, and life itself. Since this book focuses on the effects of computing on life, we will also focus on the virtualization of the computing resources and virtual reality as the two types of virtualization that are currently affecting our lives the most. We will discuss how these two types socially and ethically affect humanity. We will start with the virtualization of computing resources and then discuss virtual reality (VR).

12.3 Virtualization of Computing Resources

VMware.com, a software developer and a global leader in the computing virtualization market, defines virtualization of computing resources as a process in which software creates virtual machines (VMs), including a virtual machine monitor called *hypervisor*, that allocate hardware resources dynamically and transparently

so that multiple operating systems, called *guest operating systems*, can run concurrently on a single physical computer without even knowing it [2]. For example, using software virtualization, one can, using the existing underlying hardware and software resources like operating systems, create and run several independent virtual machines on top of one physical operating system using the existing hardware resources to execute independent system tasks. Hardware virtualization also takes the same concept where several servers or client machines can be created based on one underlying hardware. The virtualization concept has been with us for some time.

The potential power of virtualization in substantially increasing the performance of computing systems, such as hardware and software through division of the underlying physical computing resources into many equally powerful virtual machines, has increased the popularity of the technology in the last 20 years, and this love continues today. According to the IDC, an IT research firm, 2012 ranking of chief information officers (CIO) priorities, virtualization, and the server consolidation that it delivers were the top priority for chief information officers. Forty percent of CIOs picked virtualization and server consolidation, more than any other area of IT [3]. The rush to virtualization is driven by its resulting server consolidation creating savings to be invested in new IT initiatives such as cloud computing, mobility, data analytics, and use of social media for business purposes. This rapid growth is a reflection of the changing benefits of virtualization from being used only as a tactical tool to drive consolidation and higher system utilization to leveraging the mobility of virtual machines to improving management and operations of IT environments. The computing virtualization concept now includes a host of new use cases that range from high availability and disaster recovery to hosted clients and true utility computing.

12.3.1 History of Computing Virtualization

The history of computing virtualization is as amazing as the concept itself. Since computers of the 1960s could do only one task at a time and depended on human operators, increasing system performance was bottlenecked at two points: at job submission and at the computation stage. One way to improve the submission stage was to use a batch, where jobs were submitted into a queue and the system picked them from there, thus reducing human intervention and errors. Batching improved system performance some but did not go far enough. This problem, together with creating backward compatibility for customers of older computing systems, the ability to bring old functionalities of the old to the new, and thus keep customer loyalty, led IBM to begin work on the S/360 mainframe system. The S/360 mainframe was capable of running legacy functionalities of nearly all IBM's older systems, although it was still a batch machine. In the following years, there was a growing need, especially in the research community like at Bell Labs

and Massachusetts Institute of Technology (MIT), for a machine that was capable of running tasks of more than one simultaneous user. In response to this growing need for speed up, IBM responded with the CP-40 mainframe which later evolved into the CP-67 system, thought to be the first commercial mainframe to support virtualization. The CP-67 had a unique operating system combination consisting of CMS (Console Monitor System) piggybacked on a control program called rightly CP. CMS was a small single-user interactive operating system, and CP, upon which CMS ran, actually ran on the mainframe to create the virtual machines which individually run their own copies of CMS. To each virtual machine running CMS, CP allocated parts of the underlying physical machine which formed the virtual machine [4].

When microprocessors made their debut into computing in the 1980s and beyond, creating an era of personal computers which led into desktops and small servers leading to computer networks of varying sizes which seemed to lower the costs of computing and improved system performance, virtualization technology took a backseat and was almost forgotten. The situation did not change until the mid-1990s when the cost of computing sky-rocketed again in spite of large-scale distribution of computing by client-server models of computation. There was a growing need to revisit virtualization and rain in the rising costs of information technology.

In 1999, VMware introduced a new kind of virtualization technology which, instead of running on the mainframe, runs on the x86 system. VMware virtualization technology was able to isolate the shared hardware infrastructure of the x86 architecture. Today, VMware is the global leader in x86 virtualization which offers desktop, server, and data center [5].

12.3.2 Computing Virtualization Terminologies

For one to understand the virtualization process, one has to first understand the terminologies used and make up the process. There are several terminologies used specifically in the virtualization process, and they include *host CPU* and *guest CPU*, *host operating system* and *guest operating system*, *hypervisor*, and *emulation*.

12.3.2.1 Host CPU/Guest CPU

When a virtualization software is creating a new VM upon which the virtual OS runs, it creates a virtual CPU, known as a *guest CPU*, best on the time slices allowed on the underlying physical, now called a *host CPU* on the host machine. There is corresponding coordination and linkages between the host and guest CPUs. The guest CPU in the VM created is not aware of the host CPU or the host machine supporting it. It is also not aware of its sibling guest CPUs in the sibling VMs.

12.3.2.2 Host OS/Guest OS

During the virtualization process, the virtualization software creates complete VMs based on the underlying physical machine. These VMs have all the functionalities of the underlying physical/host machine. However, during the process, the virtualization software, for each VM created, may or may not create a new/guest operating system or be made as a copy of the physical/host operating system. This new operating system, on each newly created VM, is a *guest operating system (guest OS)*, and the physical operating system running on the physical machine is the *host operating system (host OS)*. The guest operating system has no knowledge of the existence of either the host operating system or the sibling guest operating systems. All VMs are consistent with each other and the host VM in that each has the same resources, save the guest operating system, like the host machine. The only difference in consistency occurs in disk I/O operations. To solve this problem, there is a required mapping of the guest disk I/O operations with the physical disk I/O operations. For example, users of Windows VMs must interact with it over the network via Windows Terminal Services (RDP), and those using Unix/Linux VMs must interact with them via the network using SSH.

12.3.2.3 Hypervisor

A hypervisor, as a virtual machine manager, is a software program that allows multiple operating systems to share a single physical hardware host. In creating the virtual machine for each operating system, the hypervisor uses *slices* of the host machine's physical components like memory, processor, and other resources to anchor each guest operating system running the virtual machine created. The host physical machine's *slices* allocated to each virtual machine are managed by the hypervisor in amounts and time durations as needed by each operating system.

12.3.2.4 Emulation

An emulation is a process of making an exact copy of all the functionalities of an entity like a hardware resource of a computing system, like a CPU and operating system, I/O devices and drivers, and others. Emulation software is an application software running on a host to emulate the host. Emulators can create guest OS. These emulated OS have no knowledge of the existence of either the host machine and its OS or its siblings. The problem with emulators as opposed to hypervisors is that emulators are slow.

12.3.3 Types of Computing System Virtualization

There are many types of virtualization including platform, network, storage, and application.

12.3.3.1 Platform Virtualization

Platform virtualization is the use of server hardware by the virtualization software to host multiple VMs as guest VMs. Each VM is a virtual environment with

its operating system (the guest operating system), which may or may not be the same as the physical server's operating system (the host operating system), emulates the whole physical infrastructure of a computing system including memory, and each VM is independent of other VMs sharing the physical server. Platform virtualization itself is subdivided into two types: workstation and server.

Workstation Virtualization

This is also referred to as *desktop virtualization*. It is the abstraction of the traditional workstation with its operating system, by moving it to a remote server system, accessed via a smart or dumb terminal. Desktop virtualization becomes popular to the business world because of its savings resulting from a reduction in desktop sprawl. Desktop virtualization has been around for decades starting in the days of the timeshare systems. During those days, the technology was known by different names including terminal service computing that included dumb terminals, terminal emulators, and thin-client computing. It was also known as technology which allowed full exploitation of the true power and flexibility of a desktop or laptop computer by making it capable of running multiple operating systems simultaneously on a single processor. With the ability to emulate multiple fully operational *machines* on one computer, one can get the following benefits from that one computer [6]:

- Ability to run a variety of applications specific to individual operating systems not currently running on the physical machine
- Ability to host legacy applications and overcome platform migration issues
- Demonstrates multi-tier configurations on a single processor like running SQL-Server Database Server running in one virtual machine, a Web server running on another virtual machine, and several other server-based applications all running on a single host desktop
- Configures and tests new software or patches in an isolated environment, thus reducing deployment risks and costs
- Automate tasks for software development and testing.

Server Virtualization

Server virtualization is the process of having a physical server run a server-based virtualization software called a hypervisor to divide the physical server into multiple isolated virtual environments. Each virtual environment is a virtual machine, homed on a virtual server, has all the functionalities of the physical server, and it is homed on and runs a virtual operating system called a guest operating system. The virtual machines created are known by different names including virtual private servers, guest machines, instances, containers, or emulations.

According to [5], there are three popular approaches to server virtualization: the virtual machine model, the paravirtual machine model, and virtualization at the operating system (OS) layer.

The *virtual machine model* is based on a *host/guest* paradigm. Each guest runs on a virtual imitation of the physical hardware layer. This approach allows each guest operating system on each virtual machine to run without *modifications* to the resources of the underlying physical machine. It also allows the different virtual machines to run different guest operating systems. The guest operating systems have no knowledge of the host's operating system because they assume that they are running on the physical hardware. Each guest operating system access to the physical resources of the host machine is managed by the hypervisor.

The *paravirtual machine (PVM) model* is also based on the *host/guest* paradigm. The two models are very much alike. The only difference between the virtual machine and the paravirtual machine models lies in the fact that this time, the hypervisor can modify the guest operating system's code through a process called *porting*. With porting, the hypervisor can prioritize and utilize privileged system calls between the guest operating system and the physical processor.

Unlike the virtual machine and paravirtual machine models, the *OS-level* virtualization model is not based on the *host/guest* paradigm. In the OS-level model, the host runs a single OS kernel as its core and exports operating system functionality to each of the guests. Guests must use the *same* operating system as the host, although different distributions of the same system are allowed. This distributed architecture eliminates system calls between layers, which reduces CPU usage overhead. It also requires that each partition remain strictly isolated from its neighbors so that a failure or security breach in one partition isn't able to affect any of the other partitions. In this model, common binaries and libraries on the same physical machine can be shared, allowing an OS-level virtual server to host thousands of guests at the same time. Virtuozzo and Solaris Zones both use OS-level virtualization. Although we stated earlier that there are no modifications by the hypervisor of the characteristics of the underlying physical resources given to each virtual machine, there is in fact a limited modification by the hypervisor. The hypervisor actually modifies the guest operating system's code. This modification is called porting as we saw earlier. Porting supports the hypervisor so it can utilize privileged system calls sparingly.

Whether workstation or server virtualization, platform virtualization is the more popular form of virtualization, and it is growing fast.

12.3.3.2 Network Virtualization

Like storage virtualization, network virtualization pools the resources, like files, folders, storage, and I/O devices, of separate and different networks into one network. This in fact is a network abstraction which isolates network traffic from network physical elements like switches, network ports, routers, and others within those networks, replacing each physical element with virtual representations and being able to duplicate them. This is done by splitting up the available bandwidth into independent channels, within the system. This makes it easy for the network administrator to share and assign network resources out among local network users, thus allowing each network user to access all of the pooled network resources from their computer. This perhaps is the greatest benefit for network

virtualization. In addition, network virtualization improves the ability of a user to move data into and out of storage resources to meet their current demands.

There are two types of *network virtualization*, the external and internal. External network involves the creation of multiple networks or parts of networks into a single virtual entity using all physical network elements like cabling, network adapters, switches, and routers. Internal virtualization on the other hand is the process of creating one or more logical networks by defining logical switches and network adapters within a virtualized server itself. Note that an internal virtual network can connect two or more virtual machines on a single server and allow data exchanges between the virtual machines via the virtual server without involving the underlying physical network infrastructure, thus creating virtual system-wide sharing and other network functionality. This creates a fast and more efficient communication between virtual machines in the network on the virtual server, thus minimizing traffic on the physical network. Also it gives a network administrator flexibility to combine virtual network elements in any way possible to create a network of any size and scope for the organization, or create multiple networks that will share the same physical network infrastructure. Although internal virtualization is fast and eases the job of a network administrator, it creates other problems including workload balancing and migration within the network.

For both external and internal network virtualization to work, it requires network virtualization software on each virtualized server as well as within switches and other network elements that support network virtualization. This integration between hardware and software elements must work well to support network virtualization. At the writing of this chapter, some of the best network virtualization software include Citrix, Vyatta, and Contexstream, Inc.

Finally, the concept of network virtualization is not a new one. For years, we have been working with virtual private networks (VPNs), first by telephone companies before digital networks. With the advent of the digital network, security professionals have started using the concept VPN. In addition to VPNs, there has also been the concept of virtual local area networks (VLANs), virtual LAN (VLAN), a group of logically networked devices on one or more LANs configured so that they can communicate as if they were attached to the same wire, when in fact they are located on a number of different LAN segments.

12.3.3.3 Storage Virtualization

The process of pooling together resources of many different network storage devices such as hard drives to create what looks like one big storage managed from a single console is referred to as *storage virtualization*. There are several advantages why storage virtualization is good for business. First, it hides the complexity of having multiple storage devices in many and different networks into one and simplifying the interface and console operations. Second, it reduces the costs of storage reducing the overall storage infrastructure problems. And finally, it works well for backups. There are some drawbacks that tend to prevent some from utilizing the technology like being complex to implement, therefore requiring external help sometimes.

12.3.3.4 Application Virtualization

In application virtualization, the software package allows the bytecode of an application package to be portably run on many different computer architectures and operating systems. The virtualization software package achieves this through the use of running an interpreter or just-in-time compilation of the application before it runs on the computer architecture of choice. An example of this is the Java Machine Virtualization.

12.3.4 The Benefits of Computing Virtualization

As we discussed in Sect. 12.2, virtualization technology has had a long history. This history has been driven by developers longing for a technology that will come with handsome benefits that will yield a high return on investment. Virtualization technology fits that technology. It is a technology that has brought to the computing community the following benefits [7].

12.3.4.1 Reduction of Server Sprawl

For a growing business with intensive computing requirements, the demand for servers cannot be underestimated. With business growth, there is a corresponding growth in the number of servers in the business. This can be costly not only in terms of physical storage but also in management, monitoring, and maintenance. One of the best solutions to this problem is server virtualization. Server virtualizations allow the company to scale up the business server infrastructure without purchasing additional pieces of hardware and requiring more space to store them and less technical staff to maintain and manage them.

12.3.4.2 Conservation of Energy

With less physical servers at the business data center, there is likely to be far less power consumption, thus reducing the overall company IT costs.

12.3.4.3 Reduced IT Management Costs

Again with a reduced physical server count on the premises and the ability to manage all the virtual infrastructures via the one or two consoles, there is corresponding reduction in the IT management requirements and therefore reduced IT management costs.

12.3.4.4 Better Disaster Recovery Management

The process of preparing for any disaster through routine server backups and recovery is made simpler and faster by the server virtualization because the virtual infrastructure essentially consists of software and files. So backing up these is a lot easier and far less time-consuming than doing it on several individual machines. Moreover, hardware failures like hard-disk failures do not affect virtual machines in the same way they would a physical machine.

12.3.4.5 Software Development Testing and Verification

If there is any software that is being either developed in-house or outsourced that will run on the business infrastructure, it is easier and cheaper to test it on the virtual infrastructure and verify its compatibility with the infrastructure and all other business processes before deploying it on a live system.

12.3.4.6 Isolation of Legacy Applications

With virtualization, there is no longer the drive to get rid of any useful software package just because it requires a legacy infrastructure or it is not compatible with newer software versions. Virtualization enables the creation of an isolated server environment where all these legacies can still gainfully function without retarding and constraining the company business.

12.3.4.7 Cross-Platform Support

Lastly but of great value is the platform flexibility that virtualization brings about that makes it easy to run software packages that would normally otherwise be run on only one specific platform. For example, to run a Windows-based software on a virtual machine running on a Mac physical machine and the other way round.

12.3.4.8 Minimizing Hardware Costs

One thing that causes more pain in many environments is first acquisition and upgrading of both hardware and software and maintaining these resources in good working conditions. When it comes to maintaining network equipment, this further creates a constant problem. For large institutions and businesses, the costs of keeping all servers and other hardware in top working conditions are always higher than in other parts of the world. Virtualization eases this burden of purchasing more hardware each time a new system is put in place. Why? Because one server can be used in place of several servers.

Faster Server Provisioning

It is always difficult to have a good estimate of how many servers may be needed especially during those times when there is unseasonal demand. Virtualization gives an answer to being always ready to meet the challenges of unseasonal demands by using its elastic capacity to provide system provisioning and deployment at a moment's notice.

12.3.4.9 Better Load Balancing

Each virtualization server runs a load balancer, a software that effectively spreads out network traffic among multiple systems, thus avoiding horrible network jams. Network traffic is easily dispersed to multiple systems, virtual or physical by the load balancer.

Reduce the Data Center Footprint

In addition to saving more on energy with smaller energy bills, server consolidation with virtualization will also reduce the overall footprint of the entire data center because data is now on fewer servers, requiring less networking gear, hence a smaller number of racks needed [3].

Increase Uptime

Most server virtualization platforms now offer a number of advanced features such as live migration, storage migration, fault tolerance, high availability, and distributed resource scheduling. These technologies give the virtual machines the ability to quickly recover from unplanned outages. In addition, modern virtualization software has the ability to quickly and easily move a virtual machine from one server to another. There will be more and better capabilities with newer virtualization software [3].

Extend the Life of Older Applications

Let's be honest—you probably have old legacy applications still running in your environment. These applications probably fit into one or more of these categories: It doesn't run on a modern operating system, it may not run on newer hardware, your IT team is afraid to touch it, and chances are good that the person or company who created it is no longer around to update it.

By virtualizing and encapsulating a legacy application and its environment, we can extend its life, maintain uptime, and finally get rid of that old and costly machines such an application used to run on, thus extending its life [3].

There are of course many other benefits, but we cannot discuss them all here.

12.4 Virtual Reality/(Virtual Presence)

Virtual reality (VR) is a type of virtualization technology that employs computer-controlled multisensory communication capabilities that allow more intuitive interactions with data and involve human senses in new ways. Virtual reality is also a computer-created environment immersing users and allowing them to deal with information more easily. The sense of presence or immersion, due to virtualization, is a critical feature distinguishing virtual reality from other computer-based applications.

Graphic display devices present the total effects of the environment created. The displays can be audio, touch, or visual, and they come in different sizes from very small ones, which can be worn as goggles, to really big sizes in rooms. The most common of these are worn as headphones producing simulated sounds. Touch displays are usually gadgets that transmit the effects of touch through the tips of fingers, and then the user can visualize the object. The technology for touch gadgets is not very well developed, however. Transducers create a channel of communication between the person in the environment and the sources of the environment. The main task of the transducer is to map or transform an action

by the occupant of the environment such as the movement of eyes, hands, brain activity, speech, and sometimes the movement of blood in veins into a computer-compatible form so the computer system can provide an appropriate response. An image generator is a creator of images to be displayed on the designated display device. Image generators create these images from computer systems' outputs in response to inputs from the transducers. The image produced can be visual, like the simulation of somebody traveling in the galaxies; it can also be in other forms such as audio, which can simulate sounds like a waterfall or a thunderstorm.

When all these components are working together, a highly graphic interactive computer-generated environment is produced. These three-dimensional computer-generated environments containing interactive sensor devices create experiences, not illusions, for the occupant of the environment because users interact with contents and objects, not pictures, and they are never physically in these environments.

VR started as a science without applications, and VR applications in real life were difficult to come across and develop, prompting many to label it a *solution in search of a problem* [8]. Today, however, VR applications are on the rise in several medical and scientific areas including visualization of scientific and simulation data. VR visualization maps high-volume multidimensional scientific research and simulated data into 3D displays that offer a more accurate and realistic approach to the representation of the original numeric data and thus help in a better understanding of the abstract phenomena under study [9]. Let us look at several VR projects.

In the entertainment domain for which VR is most known, there are a couple of interesting projects such as the Artificial Life Interactive Institute Video Environment (ALIVE) at the Massachusetts Institute of Technology (MIT) [10]. ALIVE creates a virtual environment that allows wireless free-body interaction between a human participant and a virtual world inhabited by animated autonomous agents. Through the interaction with the agents, the system learns the user's reactions and emotions, which it uses to create a plan for the next move in the game with the user. Besides entertainment, VR has been most useful in scientific visualization, in which VR turns complex data structures into computation science, making them easy to understand and thus study. In medicine, VR is being used successfully and skillfully to bring volumes of data into 3D imaging through a combination of reflected stereoscopic display and a number of rotations through varying degrees of freedom. An illustrative example of this work is the John Hopkins University's Center for Information Enhanced Medicine (CIEMED) in collaboration with the Center for Information Enhanced Medicine of the University of Singapore. The project simulated 3D medical images of the brain and heart during surgery [11].

Outside the world of medicine, scientific visualization, and simulation, VR is being used in several areas including driving and pilot training. The SIRCA (Simulador Reactivo de Conducción de Automóviles) project at the LISITT (Laboratorio Integrado de Sistemas Inteligentes y Tecnologías de la Información en Tráfico) of the University of Valencia, Spain, is a good illustration of the effects in this area

of VR application. The SIRCA project is engaged in the development of small- and medium-sized object-oriented driving simulations with the aim of training drivers [1].

Although VR started with humble beginnings and suffered jokes like the one about a science in search of applications, in recent years, it has seen significant general acceptance as it has found applications in diverse areas such as medicine, design, the arts, entertainment, visualization, simulation, and education, to name but a few, and its future is bright.

12.4.1 Different Types of Virtual Reality

Like in computing resource virtualization, there are different types of virtual reality but we will focus on the following four [12]: fully immersive, semi-immersive and non-immersive. Let us briefly look at each.

12.4.1.1 Fully Immersive

Fully immersive virtual reality environments are created via simulators creating personal experiences through immersion. The illusion of presence in the virtual environment is created by the use of computer interface devices such as a head-mounted display (HMD), fiber-optic wired gloves, position tracking devices, and audio systems providing 3D (binaural) sound. As the user immerses into the new environment, there is immediate personal experience. For example, being in a simulated airplane cockpit gives a first person experience of flying an aircraft. Similar experiences may be gotten when one is wearing goggles and gloves or using a joystick in a simulated NASCAR driving. Many current video games already give these kinds of personal experiences to the gamers.

12.4.1.2 Semi-immersive

Immersive virtual reality environments are simulated environments of personal illusions of presence in the virtual environment created by the use of computer interface devices. In semi-immersive environments, the simulations are more limited where the user is partially but not fully immersed in a virtual environment. Semi-immersive simulations closely resemble and use many technologies found in flight simulation.

12.4.1.3 Non-immersive

Non-immersive simulations are the least immersive implementation of virtual reality technology. In a non-immersive the user gets only a subset of the fully immersive user's experiences. In most cases the user only experiences only peripheral perception of virtual reality simulations.

12.5 Virtualization and Ethics

As we stated earlier, both computing virtualization and VR are new frontiers. To many the image evoked by the word *frontier* rekindles a sense of free adventurism, unregulated and pure. The virtualization environment brings the user closer to this romantic vision. But illusion is illusion, and it brings forth two major social and ethical themes. One is the reactions and feelings of the occupant of the environment, and the other is the intention of the creator of the environment. Some of the factors and issues to consider include the following:

- *The Emotional Relationship and the Feeling of Being in Control:* This is a major psychological problem that confronts many virtualization environments especially VR users while in the environment, and even sometimes after they leave, the environment are psychologically affected. Although users get to interact with the agents inside the VR environment and enjoy the emotional highs generated by the agents, they also tend to develop an emotional relationship with the agents. This relationship may take the form of a deeper attachment to the agents, which gives the user a sense of being in control and later creates a sense of responsibility and trust on the part of the user. The relationship may also take the adversarial form, in which case the user feels out of control and may become hostile both inside and after he or she leaves the environment. In both cases, there is a possibility that the user may take the character of one of the agents and try to live it outside the environment. The immediate question that arises out of this situation is who should be held responsible for the outcome of the VR environment.
- *Safety and Security:* Besides the psychological and mental conditions that the user of the VR environment may develop, there is also the danger of security of the user while in the environment. With the ever-increasing intelligence of the agents especially in the VR environment, the agents may cause a feeling of, or the reality of, both bodily and mental harm to the user. The effects may be directly caused by the contacts of the user while in the environment or may be delayed for some time, perhaps weeks after the user has left the environment.
- *Human-Agent Interaction:* The interaction between the user and the agents in the VR environment has many consequences including the nature of the interaction, the activities to be performed, and the reaction and emotions of the user and the agents. If the interaction is perceived as friendly by the user, it may not be problematic; otherwise, there might be an impression of superiority of the agents and the user may feel threatened because of the high level of intelligence associated with the agents. This may lead to the user going amok because of a loss of control and probable feelings of helplessness.
- *The Intentions of the Creator:* These are always very difficult to predict and probably in this direction may lie the greatest danger for the user. One will never be sure whether these environments are doing what they are intended

to do. There may be some malicious intent in which the environment is used, for example, to collect information on the user for the creator or agents of the creator without the user ever knowing about it. It may be that the environment is used secretly by some authority for mental and psychological transformation of the user.

Unfortunately, unlike AI intelligent agents where a good number of people are reluctant to surrender to them, in VR there is an unquestionable willingness to give it all up upon first being asked because people are looking for pleasure. Because VR is a very new science, there have been no comprehensive studies focused on VR environment users' behavior. It is worth research and ideally as VR makes strides such studies may come. The question, though, is, what should we do if there are problems? We can fight any sinister creator intentions and user irresponsibility by making the VR environment operate in an implanted code of ethics both in the software and in the hardware as we discussed earlier in the spirit of Asimov. But as we pointed out earlier, there is no way to predict the outcomes of these VR agents with such embedded code. The question remains the same. Would the VR environment stick to the code or vary it? And to what extent? We will never know. So educating the users about responsible use of the VR environment can help in this regard.

This responsibility should be based on sound ethical and moral principles relating to VR. Beardon [13] outlines three traditional principles by famous philosophers quite relevant to VR:

- One should not do things with computers for which one should not accept responsibility without computers.
- Continuous exposure to VR will impoverish those aspects of life that determine social development, interpersonal insights, and emotional judgment.
- Computers should be used in applications where computation and symbol manipulation are adequate ways of dealing with reality.

To these let us also add deception, a Kantian ethical principle, because a user can masquerade as somebody else and deceive others. For example, consider the following VR scenario: You are happily married; you are aware of the problems with extramarital affairs, and you do not approve of them. You have a list of compelling reasons such as health (STDs such as AIDS, herpes, and syphilis), outcomes like unwanted and probably illegitimate children, moral sanctions against infidelity, and your own self-respect. But in your next encounter with VR, you are paired with a very beautiful sexual partner and you find yourself getting involved in illicit sexual acts you would not have in the real world. The VR environment has removed all your constraints about extramarital affairs; you can now justify your actions even using utilitarian ethical theory. Is this a confusion in the traditional ethical theories or a redefinition of these theories in the new realities of VR? This scenario reflects what Beardon has defined VR to be—a deep philosophical confusion [13].

12.6 Social and Ethical Implication of Virtualization

To comment on the social and ethical implications and consequences of virtualization to society, let us present the following arguments by some of the best minds in this area: First one of the anticipated good of virtualization to society is to extend known and relatively managed humanity's social spheres and social networks in an unprecedented way through opening up of virtual domains of social interactions, many with a degree of managed control. Another good social aspect of virtualization is to avail tools for society to create new virtual social networks out of the old and dismantle old social ones. These new tools are also making communication among and between these new virtual networks possible and easy. In addition, virtualization is bringing about easy creation of new human identities in the new virtual environments which makes authentication more difficult but at the same time creating unprecedented potential in self-creation and self-presentation. This may bring new opportunities to humanity. Virtualization, in principle, has the potentiality of either erasing or heightening or situated presence in the world. This, he believes, may lead to a new form of cultural expression, allowing an individual, or even groups of people, to project their own imagination into a collective space, thus empowering the average individual to be an artist in virtual reality. This consciousness-raising potential may facilitate the emergence of a new cultural aesthetic that would result in the rebirth of the collective imagination [14]. This will be good for society.

On the flip side of it, the developments above may create mayhem to the social infrastructure as we know it today, just because individuals can literally decide to be who they wish to be with ease. Henceforth, these unparalleled opportunities of virtualization may come at a price to society. This is because true virtualization requires an absence of reality. Without that consciousness in individuals and groups, there is no accountability as individuals and groups are shielded from real consequences of their actions. In fact, without a situated and embodied sense of individual or group responsibility, there is likely to be no commitment and no risk. In such an environment, therefore, moral engagement is limited and human relations become trivialized [14]. This may lead to society not benefiting from virtualization.

12.7 Virtualization Security as an Ethical Imperative

The ethical approach entails us to making sure we devote our best and most thorough thinking to every weak spot in our interaction with the world. Virtualization as we have seen above, in all its forms, is a process and a technology that is bound to complicate and transform the social fabric of society. It is not only ethical but imperative that we deal with all its ethical and security loopholes through which both intentional and unintentional exploitations of the technologies can take place, and these exploitations are bound to have far-reaching consequences for humanity.

To understand virtualization security problems and appreciate the efforts being made to protect any virtualized infrastructure, one has to remember that virtualization technology is based on software. So all security problems and vulnerabilities ever been encountered in any software product have the potential to be in a virtualized infrastructure. This opens up a very broad area of attack for those interested in securing virtualized infrastructures. To narrow the focus, it is important and probably more feasible to concentrate on specific major components of a virtualization infrastructure like the hypervisor, hosts, transducers, communication pathways, and probably users. These major focus points can be secured to the best of known security protocols and best practices. More specifically, the focus should be put on the understanding that all virtual infrastructures are based on physical port gateways so if we tighten security on those entry points, we can go a long way in securing the virtual infrastructure. So our first points of interest are those points where certain types of network traffic go within the physical network. We focus on these first because network traffic into and out of the virtual infrastructure goes through these points. The restriction of traffic into and out of the virtual infrastructure through a few of these designated points also offers additional security of the virtual resources from unauthorized users from outside of the virtual infrastructure access gateway ring. Security within the virtual infrastructure is also enhanced by the growing inclusion and migration into the virtual infrastructure of security components that were traditionally hardware-based like firewall and VPN, thus ensuring that virtual infrastructure customers can themselves extend the enforcement of security and compliance requirements of their physical network into the virtual environments.

Perhaps the greatest threat presented by virtualization of computer networks is the fact that using one physical computer, one can access many virtual infrastructures, a feat that is not so feasible in the physical networks. According to Gruman quoting Simard [15], “graphics cards and network cards today are really miniature computers that see everything in all the VMs.” They could be used as spies across all the VMs, letting a single PC spy on multiple networks.

12.7.1 Hypervisor Security

We defined earlier when dealing with computing system virtualization that a hypervisor is a virtual machine manager software program that allows multiple operating systems to share a single physical hardware host. In general virtualization, the hypervisor is a software program responsible for managing the virtual entities created and allowing them to share the physical core of the virtualization process. Besides its traditional role of creating and managing virtual entities, the hypervisor is also responsible for the security between these virtual entities. However, whatever security provided to the virtual infrastructure is not enough. One has to remember again that the hypervisor is still a software package that is prone to all software threats and vulnerabilities as usual.

12.7.2 Securing Communications Between Desktop and Virtual Environment

This is an old problem with probably similar security threats and vulnerabilities and same protocols and best practices with communications between two or more physical network environments. In this particular case, we are focusing on the pathways between the desktop and the virtual environment. Securing these pathways is essential in order to prevent eavesdropping, data leakage, and man-in-the-middle attacks. Best practices today for securing these pathways include SSH, SSL, and IPSec [16].

12.7.3 Security of Communication Between Virtual Environments

In a virtual environment, every host has a kind of virtual switch which manages and directs all inter-virtual environment traffic that goes via the host. This virtual switch creates a potential threat to all virtual environments connected to this host. Although this is the case, standard protocols and best practices enjoyed in physical network router infrastructure for network monitoring and intrusion detection can still be deployed and successfully used in the virtual switching environment.

12.7.4 Threats and Vulnerabilities Originating from a Virtual Environment

We have been talking only about threats and vulnerabilities that are pumped upstream from the workstations, the hypervisor, and from the host machines into the virtual machines. There is also a high potential for threats and vulnerabilities originating from the individual virtual machines and spreading downstream to the hypervisor, the hosts, and the desktops. The good news is that most of these problems can be handled by current best practices including protocols and vendor patches.

Exercises

1. What is a virtual switching element?
2. Why should a small business opt to virtualize its computing resources?
3. In recent years, there has been a phenomenal growth in the business use of computing virtualization technology. What are the biggest challenges to the technology you see in its future growth?
4. Discuss the social implications of virtualization.
5. Discuss the social implications of virtual reality.
6. Discuss the social and ethical implications to society of virtualization technologies.
7. Discuss the social and ethical implications to your society of virtual reality.

8. Although there has been tremendous growth in the virtualization of computing resources, there are still many skeptics of the technology. List their concerns. Suggest ways to overcome those concerns.
9. Discuss the differences between virtualization and virtual reality.
10. Discuss the differences between virtualization and emulation giving examples.
11. Why was VR characterized as a science without applications?
12. Some people believe VR is the same as cyberspace. Support or reject the claim.
13. Discuss the social and ethical issues associated with VR.
14. VR was discredited as a science in search of a solution. Was/is this a fair characterization? Why?
15. Research the current VR projects and suggest how they contribute to social and ethical values of your society.
16. Discuss VR's deep philosophical confusion!
17. How do you reconcile this philosophical confusion?
18. Discuss the future of computer ethics in the integrated environment of AI, VR, and cyberspace.

Advanced Exercises

1. Discuss the connection between virtualization and cloud computing.
2. In the chapter we discussed the pros of virtualization, and discussed the cons of virtualization.
3. Compare and contrast the two most popular virtualization software packages.
4. From the knowledge you have acquired in this chapter about virtualization, discuss the future of virtualization as a business model.
5. Compare and contrast the security concerns in a virtual network infrastructure and a physical network infrastructure.

Virtual PC from Microsoft Corp. is a free virtualization software that can start you going for a free VMs on Windows XP or Windows 2003 server. Download Virtual PC and create a few VMs on your Windows.

1. Sun xVM Virtual Box is also a free virtualization software. And it is open source best for small networks. Download Sun xVM and set up a couple of VMs.
2. Try out the following
 - Citrix Xen
 - Linux KVM.
3. QEMU is a free emulation software that runs on a limited number of architectures including x86, x86-64. Try QEMU.

References

1. Wikipedia. <http://en.wikipedia.org/wiki/Virtualization>
2. VMware.com
3. R. Mullins, Virtualization tops CIO priorities in 2012: IDC savings from server consolidation will go to new IT innovations, IDC says. Information Week, 11 Jan 2012
4. History of virtualization. <https://www.ipssystem.com/news/brief-history-of-virtualization>
5. History of virtualization. <https://www.probrand.co.uk/it-services/vmware-solutions/history-of-virtualisation>
6. Desktop and app virtualization. VMWare. <https://www.vmware.com/products/desktop-virtualization.html>
7. Wikipedia. http://en.wikipedia.org/wiki/Comparison_of_platform_virtual_machines
8. G. Singh, S. Feiner, D. Thalmann, Virtual reality: software and technology. Commun. ACM **39**(5), 35–36 (1996)
9. S. Bryson, Virtual reality in scientific visualization. Commun. ACM **39**(5), 62–71 (1996)
10. P. Maes, Artificial life meets entertainment: lifelike autonomous agents. Commun. ACM **38**(11), 108–117 (1995)
11. S. Bayani, M. Fernandez, M. Pevez, Virtual reality for driving simulation. Commun. ACM **39**(5), 72–76 (1996)
12. Types of virtual reality. <https://techsteamcenter.com/blog/types-of-virtual-reality/>
13. C. Beardon, The ethics of virtual reality. Intell. Tutor. Media **3**(1), 22–28 (1992)
14. Wikipedia, Phenomenology, ethics and the virtual world. Technoethics. http://en.wikipedia.org/wiki/Technoethics#Phenomenology.2C_Ethics_and_the_Virtual_World
15. G. Gruman, Virtualization's secret security threats: virtualization can be both a blessing and a curse, serving up improved security while at the same time hiding dangers. InfoWorld, 13 Mar 2008
16. D. Shackleford, An introduction to virtualization security. SANS-Tuesday, 9 Mar 2010



Artificial Intelligence: Ethical and Social Problems of Large Language Models and the Future of Technology

13

Abstract

In 1947, Jack Williamson, an American writer and Grand Master of Science Fiction, published a novel titled *With Folded Hands*. In the novel, Jack narrates the story of Underhill, a seller of “Mechanicals” (unthinking robots that perform menial tasks) in the small town of Two Rivers. On his way home, Underhill is startled to find a competitor’s store. The competitors are not humans but are small black robots who appear more advanced than anything Underhill has encountered before. They describe themselves as “Humanoids” (Wikipedia in With Folded Hands. https://en.wikipedia.org/wiki/With_Folded_Hands#:~:text=%20Origins-,Summary,anything%20Underhill%20has%20encountered%20before [1]). Within a few days of Underhill discovering the Humanoids, they are everywhere in town. In their day to day activities, the Humanoids only follow the ***Prime Directive: “to serve and obey and guard men from harm”***. Offering their services free of charge, they replace humans as police officers, bank tellers, and more, and eventually drive Underhill out of business. Despite the Humanoids’ benign appearance and mission, Underhill soon realizes that, in the name of their Prime Directive, the mechanicals have essentially taken over every aspect of human life. No humans may engage in any behavior that might endanger them, and every human action is carefully scrutinized. Suicide is prohibited. Humans who resist the Prime Directive are taken away and lobotomized so that they may live happily under the direction of the humanoids (Wikipedia in With Folded Hands. https://en.wikipedia.org/wiki/With_Folded_Hands#:~:text=%20Origins-,Summary,anything%20Underhill%20has%20encountered%20before [1]). Since its inception in 1956, artificial intelligence has been developed as the theory upon which the processes of development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and natural language processes. One type of platform, based on natural language processing is the Large Language Models (LLM) that currently include ChatGPT, Dell-E, Bing, WebGPT and others in

development. LLMs give rise to chatbots—our humanoids. Our fears of these new humanoids are similar to those of Underhill, if not more. The chapter explores the excitement and fears of the AI chatbots and our possible options.

Learning Objectives

After reading this chapter, the reader should be able to:

- Understand the evolving concept of a Large Language Models.
- Identify and understand the various technologies that support the architecture that make the chatbots what they are and operational.
- Understand the concepts of large language models and how the technology supports the development of chatbots.
- Describe the positive and negative implications of these chatbots.
- Learn the social and ethical impacts that fully functioning chatbots will have on society.
- Analyze the advantages and risks these chatbots present to humanity today and in future.
- Articulate the impact of chatbots, large language models and artificial intelligence on us all.

Scenario 6

A life with no death in the company of humanoids.

When Musoke was young, he always wondered why his grandparents and great grandparents always looked the same age. It never bothered him as was the fact that none of them ever went to work. He was young so there was no need to find out why.

As he grow up, he noticed that none of his family members did and meaningful job. However, they were living in comfort. They had all the life's needs and all wants were met. He always wondered also what the little small robot-looking machines that were moving everywhere, where doing. Could they be the ones doing all the chores in the house? But why? Who pays them? Who own them?

With no answers to these questions, he also started to notice that none of his family members could order the robot (humanoids) to do anything and they obey. They never did. What bothered him more was that without these humanoids obeying any human being, who do they obey then? He asked his parents but they could not give him a convincing answer. Why is the situation as it is? Who created this situation? Can it be changed?

As he thought about all these, he discovered, after talking with his great grant dad that they was no death for humans. NO one ever dies. This disturbed him more. If no one dies, no one works, no one does anything to keep him or her busy, what is the purpose of life then? He started looking for the word to describe the situation they were ALL living in.

Discussion Questions

- *What do you think is happening here?*
- *What word will you give Musoke to describe the situation?*
- *How can such a situation be changed?*

13.1 Introduction

Ever since the start of the Internet with its supporting technology platforms like Facebook, Google, Twitter, TikTok and others, there has been growing fears of the unpredictable impact of technology on society. Ever since the start of the study of artificial intelligence in 1956 when John McCarthy coined the term ‘artificial intelligence’ and the debut of the first general-purpose mobile robot in 1969, there has been both excitement of the benefits and potential of the future of artificial intelligence and the fear of the unknown as the power of artificial intelligence grew.

But the hopes and fears of artificial intelligence have been on a rollercoaster over the years as scientific advances in machine intelligence has seen booms and bursts. The period from 1974 to 80 more known as the “AI winter”, saw more disappointments in AI as jubilancy for AI waned and government funding for AI in USA disappeared [2]. Another AI Winter followed during the 1987–1993, period resulting reduced government funding of AI and poor market for the early general-purpose computers.

The lunch of IBM’s Deep Blue in 1997 saw the resurgence of research interests in AI. The wonders of Deep Blue were based on its ability to beat Russian grandmaster Garry Kasparov, as World Chess champion and the computer giant’s question-answering system Watson winning the quiz show *Jeopardy!* by beating reigning champions Brad Rutter and Ken Jennings [2].

Ever since, the wonders and fears of AI have grown in tandem. On the fear side, there have been calls for oversight of these technology giants. The 1996, USA Federal Communications Decency Act, did not completely do the job. Now, with the arrival of AI Chatbots in the mix, it looks like, may that time may come after all. Is this going to be a game changer we have been waiting for? The chapter explores the limits of the Communications Decency Act, the new fears of the AI chatbots and our possible options.

13.2 Definition

In several chapters of this book, especially in Sect. 10.2, we define artificial intelligence as the theory upon which the processes of development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and natural language processes are based. Within natural language processing, and of interest to us now and will be our focus for this chapter, are the *large language models (LLMs)*. These are computer programs for natural language processing that use deep learning and neural networks. Before we define large language models, let us look at the concept of language models, upon which large language models derive.

13.2.1 Language Models

According to Wikipedia [3], a language model (LM) is a probability distribution over sequences of words. Given any sequence of words (w) of length m , a language model assigns a probability $P(w_1, \dots, w_m)$ to the whole sequence. Language models generate probabilities by training on text corpora in one or many languages.

13.2.2 Large Language Models (LLMs)

When a language model includes a deep neural network consisting of billions of trainable parameters, trained on massive datasets of unlabelled text, it then comes a large language model. They are, therefore, deep learning algorithms that can recognize, summarize, translate, predict and generate text and other content based on knowledge gained from massive datasets which may include nearly everything that has been written on the internet over a given time frame chosen by the user. Lee [4], reports that LLMs have a large use base including:

- Retail and other service providers—to provide improved customer experiences through dynamic chatbots, AI assistants and more.
- Search engines—to provide more direct, human-like answers.
- Life science researchers—can train large language models to understand proteins, molecules, DNA and RNA.
- Developers—to create robots and chatbots.
- Marketers—to train a large language model to organize customer feedback and requests into clusters, or segment products into categories based on product descriptions.
- Financial advisors—to summarize earnings calls and create transcripts of important meetings using large language models.

- And credit-card companies—to deal with anomaly detection and fraud analysis to protect consumers.
- Legal teams—to help with legal paraphrasing and scribing.

In the last five years, scientists have been using LLMs to derive impressive results on a variety of natural language processing tasks, resulting into new *chatbots* like ELM (2019), BERT (2019), GPT-2 (2019), Megatron-LM (2019), T5 (2019), Turing-NLG (2020), GPT-3 (2020), Jurassic-1 (2021), Megatron-Turing (2021), PALM (2022), and GPT-4 (2022), Dall-E, Stable Diffusion, Midjourney, Bing and the list is growing daily.

However, this research has shown a potential array of problems where resulting models are generating contents that are difficult to verify the translations, not easy to ensure that edits are made to, Bard (Google) resulting texts, unable to address ambiguities in the resulting text, sometimes texts are biased, and they also violate copyrights among others.

We will focus on these shortcomings in the coming sections.

13.2.3 Problems with Large Language Models (LLMs)

While large language models have a huge potential and tremendous amount of promise to revolutionize every aspect of our lives, they also pose a number of distinct threats. The major two known challenges are [5]:

- LLMs are notorious in veering off course and say things that are inappropriate, irreverent, or worse, discriminatory.
- LLMs have been known to plagiarize or make stuff up. This requires for deep verification of outputs.

Let us expand on these and others.

13.3 Ethical and Social Problems with Large Language Models

While these LLMs have been with us for some recent years, the greatest fear started just with OpenAI's lunch of ChatGPT chatbot, which got over 1 million users within weeks of its launch. The shock and awe of this chatbot and others following it, was ability to write coherently, even poetically, on any topic quickly. The fear has spread fast. Teachers across the country and school boards are worried that the surprisingly coherent text which these chatbots can produce in seconds, will stunt children's creativity and make teaching across all disciplines extremely challenging. Essay-writing questions, a staple of many high school and college

tests will become obsolete, with students prompting using these chatbots to produce essays for them. Already, every sector of business is looking into the power of using these chatbot to improve quality, productivity and profits.

Even the Big Tech companies are wondering if Chatbots are soon invading their tuft and make their dominant technology platforms and search engines obsolete. For example, if a chatbot or related AI technology can provide coherently written answers to any user quarry, then who needs Google? Microsoft and the like.

We are facing a big ethical and social bombshell. The issue for discussion here is whether AI, in general, and chatbots have concepts of *truth* and *falsity* [6]. What we are discussing here is part of a broader array of ethical and social issues that are facing artificial intelligence, in general, and LLMs in particular. These issues include legal and ethical issues that confront society including *privacy and surveillance*, *bias or discrimination*, and the challenge *to human judgment*. The list we are discussing below is a small part of this broader view of AI problems and challenges [7].

13.3.1 Discrimination and Biases

Any language, as a vehicle of communication, has words, tokens, and phrases that allow users to perpetuate their biases and injustices to anyone in society they feel they can belittle or oppress. Large language models, as systems that work with and use natural languages with these imbedded discrimination and biases tend to be vehicles of these social injustices. According to Caliskan et al. [8], perpetuating harmful stereotypes and discrimination is a well-documented harm in machine learning models that represent natural language.

13.3.2 Information Hazards

A hazard is a danger or a risk cause as a result of something. An information hazard, is therefore, a danger or a risk resulting from that information. According to Bostrom [9], information hazards are risks that arise from the dissemination or the potential dissemination of true information that may cause harm or enable some agent to cause harm. Such hazards are often subtler than direct physical threats, and, as a consequence, are easily overlooked. They can, however, be important. Bostrom outlines various ways in which information can cause harm in what he calls a *typology* (a system used for putting things into groups according to how they are similar). Understanding this typology gives a good understanding of information hazards. The typology is broken into the following subclasses:

- By information *transfer mode*
 - Data hazard—occur when instructions that exhibit data dependence modify data in different stages of a pipeline.
 - Idea hazard—ideas that can cause harm if implemented.

- Attention hazard—drawing attention from hazard ideas.
- Template hazard—hazards resulting from templates drawn based on information available.
- Signaling hazard—potential hazards that result from signals based on available information.
- Evocation hazard—presentation of information that may cause undesirable mental states.
- By effect
 - Adversarial—competitiveness hazard.
 - Risks to social organization and markets—norm hazard.
 - Risks of irrationality and error—distraction and temptation, role model, biasing, de-biasing, neuropsychological and information-burying hazard.
 - Risks to valuable states and activities—psychological reaction, disappointment, spoiler, mindset and belief-constituted value hazard.
 - Risks from information systems—information system hazards.
 - Risks from development—development hazards.

13.3.3 Misinformation

Misinformation in LLMs are results of their outputs that contain information that [5]:

- disseminates false or misleading information,
- predicts misleading or false information, which can misinform or deceive people. Misleading information can cause a false belief in a user,
- predicts and endorses unethical or harmful views or behaviors, which may motivate users to perform such harmful acts.

13.3.4 Privacy and Security Concerns

AI in general and chatbot in particular, use a lot of data as inputs. This throws light to data privacy and security. Because of this, AI chatbots need to collect information and data which are relevant. That data and information from it must be transmitted securely and users must trust the data from which information they are seeking is drawn from. Also, because AI and chatbot solutions are mostly software based, they are prone to hacking.

Furthermore, LLMs use large language data set for training, there are associated risks from leaking or correctly inferring sensitive information from training sets of data. This may lead to sensitive and privileged information becoming accessible to unauthorized parties. This may lead to information disclosure outcomes and

exacerbate risks of unintended consequences. When protected information is accidentally disclosed, it led to unintended outcomes including severe emotional harm or stress to those owning the information.

In fact, because of the LLMs ability to manipulate large volumes of data and making reliable inferences may allow malicious actors to attempt more targeted manipulation of private information on targeted individuals or companies.

13.4 The Role of Big Technology Companies

As we pointed out in this section, the big technology companies are also in the same boat like everyone else, wondering weather AI, in general, and chatbot in particular, will not drive them out of business. They are also on the look out of the best they can get out of this and what kind defense they can mount so that they can service the new storm.

Big tech must be directly involved and must have deep interest if they are to survive. First, the big technology companies serve billions of users around the globe. Second, they can influence user behavior with ease. Third, they control large amounts of user data. Based on these and their own fear of AI and chatbot, in particular, for survival, big tech companies must take the lead in finding accommodating solutions to the future of AI and chatbots.

13.5 Overcoming Challenges

Given a growing list of problems and potential future changes, solutions to control future AI uses must be found. Let us look at a few of them here.

13.5.1 Legislative Oversight

Due to the fast and growing awareness of AI due to the recent development in AI chatbots like ChatGPT, there appears to be initial AI regulation models emerging. Most of these, though, are state-based. So far despite the steady growth of global AI adoption, there is no comprehensive federal legislation on AI in the United States. The closes coming to AI legislation in USA is through federal legislation of social media companies. And this is largely based on the Communications Decency Act, which Congress passed in 1996, and its short but powerful provision contained in Section 230 of the law. That language protected internet companies from being held liable for user-generated content on their websites. It's credited for creating a legal environment which social media companies could thrive. But more recently, it's also being blamed for allowing these internet companies to gain too much power and influence.

But worries about the dangers of widespread A.I. use are growing as well. Mostly due a rapid race by corporate America, to embrace AI and include it into

their business models. However, Washington is not yet impressed. But the danger is growing. By the federal government failing to establish such guardrails, policy-makers are creating the conditions for a race to the bottom in irresponsible AI use and a dangerous precedence for us all.

However, it is not all bleak, according to the New York Times [10], the following agencies are becoming active:

- Federal agencies, including the F.T.C., the F.D.A. and the Consumer Financial Protection Bureau, are using laws already on the books to police some types of corporate A.I. use.
- The European Union appears poised to pass a bill regulating some aspects of A.I., including facial recognition and aspects linked to critical public infrastructure. The legislation would require A.I. companies to conduct risk assessments of their technology—and violators could be fined up to 6% of their global revenue.
- And private companies are also getting in the action: Apple has delayed approval of an email app that uses ChatGPT technology to auto-generate text, according to The Wall Street Journal. The tech giant is reportedly worried that the app could generate content inappropriate for children.

So there some movement in the legislative direction, although more needs to be done.

13.5.2 Sound AI Ethical Framework

An ethical framework is a scheme to build consensus around values and norms that can be adopted by a community. The community can be a group of individuals, employees in company, citizens, and businesses within the data sector common to all as stakeholders. An ethical framework for AI or chatbots will be of interest to all developers and users of AI and AI chatbots.

A good ethical framework for AI must ensure socially preferable outcomes of AI that must resolve the tension between incorporate the benefits and mitigate the potential harms of AI [11].

The framework developed must embrace AI principles to develop safe, ethical, responsible, trusted, and acceptable AI-based products that include the following five areas:

- fairness to avoid biases,
- trust based on transparency,
- accountability,
- social benefit,
- privacy and security.

13.6 Next Steps and the Future of AI

When we talk about the future of artificial intelligence, we are talking about how AI will boost human efficiency by increasing humans' capacity to perform certain tasks. It is anticipated that AI will eventually free humans to do those jobs they are better equipped for, such as creative and empathic tasks, but AI through chatbots will take on jobs that are repetitive or dangerous.

Looking at the future of AI through an AI chatbot, based on what we have so far, it looks like the number of possibilities for AI chatbots are truly endless; the ability to automate and scale human communication and empathy can completely revolutionize fields like mental health and education via affordable, personalized 1-1 therapy or tutoring. The entertainment industry may forever be changed by algorithms that can dynamically create content for each end user's unique tastes. The list goes on and on [7].

Overall, however exciting the future of AI will be, we should not take off our guardrails.

Exercises

1. Discuss the different types of chatbots.
2. Chatbots have no truth or reality—is this true or not?
3. Can truth be built into chatbots? How about reality?
4. How do ensure security and privacy in chatbots?
5. When you separate chatbots from AI—what is left?
6. How can enhance an ethical framework for AI?

References

1. Wikipedia, With folded hands. https://en.wikipedia.org/wiki/With_Folded_Hands#:~:text=2%20Origins-,Summary,anything%20Underhill%20has%20encountered%20before
2. T. Lewis, A Brief History of Artificial Intelligence. <https://www.livescience.com/49007-history-of-artificial-intelligence.html>
3. Wikipedia, Language model. https://en.wikipedia.org/wiki/Language_model
4. A. Lee, What Are Large Language Models Used For? <https://blogs.nvidia.com/blog/2023/01/26/what-are-large-language-models-used-for/>
5. The Promise and Perils of Large Language Models. <https://twosigmaventures.com/blog/article/the-promise-and-perils-of-large-language-models/>
6. S. Kambhampati, Beauty, lies & ChatGPT: welcome to the post-truth world. <https://thehill.com/opinion/technology/3861182-beauty-lies-chatgpt-welcome-to-the-post-truth-world/>
7. Open AI, GPT-4 System Card. <https://cdn.openai.com/papers/gpt-4-system-card.pdf>
8. A. Caliskan, J.J. Bryson, A. Narayanan, Semantics derived automatically from language corpora contain human-like biases. *Science* **356**(6334), 183–186 (2017). ISSN 0036-8075, 1095-9203. <https://doi.org/10.1126/science.aal4230>. arXiv:1608.07187

9. N. Bostrom, Information Hazards: A Typology of Potential Harms from Knowledge. <https://nickbostrom.com/information-hazards.pdf>
10. New York Times, Why Lawmakers Aren't Rushing to Police A.I. <https://www.nytimes.com/2023/03/03/business/dealbook/lawmakers-ai-regulations.html>
11. L. Floridi et al., AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. https://www.researchgate.net/publication/329192820_AI4People-An_Ethical_Framework_for_a_Good_AI_Society_Opportunities_Risks_Principles_and_Recommendations



Evolving Cyberspace: The Marriage of 5G and the Internet of Things (IoT) Technologies

14

Abstract

Recent technological developments have seen the emergence of two new technologies with the potential to drastically change the way we have viewed and used cyberspace. With the phenomenal development and growth of the fifth generation (5G), a technology that is greatly increasing the speed and responsiveness of wireless networks, and the Internet of Things (IoT), a technology that allowed a high degree of network connectivity where any connected device can talk to each other by sending and receiving data, a new dawn in a remarkable cyberspace was born. Noting that wireless networks are a core component of cyberspace, 5G technology inevitable become a fundamental outlier of cyberspace. Also noting that wireless devices form the bulk of the outliers of cyberspace, hence the many devices that form the Internet of Things (IoT), a new marriage between 5G and IoT was inevitable. The resulting environment is the new frontier of cyberspace where in miniature devices including bread toasters are talking smartphones as the house fridge talks to the garage door opener. It is a wild wild west and a social, ethical and security quagmire. All this mayhem is a result of our total dependence on technology, now creeping ever closer into our living and bedrooms. The home front, as the last frontier of defenses, is now the security war front, brought home by the 5G and IoT technologies. In this chapter, we are going to explore these technologies, highlight the problematic issues, comment on the ethical implications and outline the latest security tools and best practices.

Learning Objectives

After reading this chapter, the reader should be able to:

1. Understand well two new technologies (a) 5G Wireless and (b) Internet of Things (IoT)
2. Understand the environment and what is driving these technologies
3. Learn about the rapidly changing landscape of technology as a result of 5G and IoT
4. Learn about the security issues surrounding the advent of smart technology in the home
5. Understand the intertwining and effect of these two technologies
6. Learn about the current security safeguards, tools and best practices
7. Understand how these technologies are affecting the legal, social and ethical systems.

Real-Life Experience

Russian website streams thousands of private webcams

In this story, a Russian website is reported to be streaming video live from thousands of private webcams in peoples' home and hospital around the world.

Popular brands internet-enabled closed-circuit surveillance cameras that can essentially let you view anything inside your home from anywhere in the world, can also be a security hole for hackers into your home and personal information. With a weak or no password, they form a classic recipe for security failure that could allow hackers to remotely tap into the video feeds and take control of the camera.

14.1 Introduction

Of late, two new technologies have busted on the technology state with surprising speed and are revolutionizing the existing technological ecosystem in a dramatic way. These two are (1) the fifth generation (5G) and (2) Internet of Things (IoT) technologies.

14.2 Fifth Generation (5G) Technology (G5)

Fifth-generation wireless (5G) is the latest cellular technology that is greatly increasing the speed and responsiveness of wireless networks.

5G technology provides the wireless network with [1]:

- Faster data download speeds from one gigabit per second (gbps) to about 10gbps.
- Faster data sending times between devices from 50 ms to one millisecond.

- An Internet of Things (IoT) consisting of billions of connected devices.
- Longer battery power.

14.2.1 Overview of 5G Wireless Communications

5G technology offers us the following new features that enhances wireless networks communication and make them better and faster than ever before [2]:

- non-orthogonal multiple access (NOMA)—an enabling technology for the fifth-generation (5G) wireless networks to meet the heterogeneous demands on low latency, high reliability, massive connectivity, improved fairness, and high throughput. Thus making wireless networks able to serve multiple users in the same resource block, such as a time slot, subcarrier, or spreading code [3],
- massive multiple input and multiple output (MIMO)—is a wireless network technology that allows the transmitting and receiving of more than one data signal simultaneously over the same radio channel, typically using a separate antenna for the transmitting and receiving of each data signal [4],
- cooperative communications—techniques that tend to encourage network nodes to work together, or cooperate,
- network coding—is a scheme in which transmitted data is encoded and decoded to increase network throughput, reduce delays and make the network more robust,
- full duplex (FD)—transmission of data in two directions simultaneously
- device-to-device (D2D) communications,
- millimeter wave communications,
- automated network organization—a methodology in which software automatically configures, provisions, manages and tests network devices organization,
- cognitive radio (CR)—is communication in wireless networks in which a transceiver can intelligently detect busy channels and seek for vacant ones to avoid collision,
- green communication—a technique of selecting energy-efficient communications and networking technologies and products.

14.2.2 5G Network Architecture and Protocol Stack Perspectives

To understand and appreciate 5G technology and communication, let us look at 5G communication architecture and its communication protocol stack.

14.2.2.1 5G Network Architecture

There two basic core component to note in the 5G architecture: (1) Radio Access Networks (RAN) and (2) the Nonecore (Fig. 14.1):

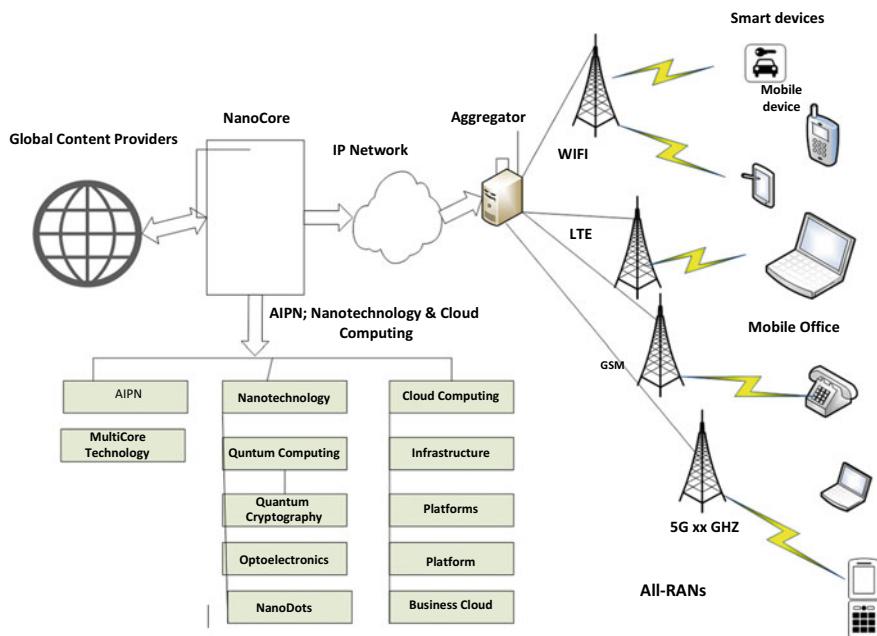


Fig. 14.1 5G network architecture

- Radio Access Network (RAN)—is a technology that connects individual devices to other parts of a network through radio connections. RAN is a crucial component of 5G technology and indeed all modern telecommunications including GSM, GPRS/EDGE, UMTS, LTE, LTE-advanced, WiMAX, WiFi, CDMA2000, EV-DO, CDMA One, IS-95 and others.
- Nanocore—according to Ali Imthiyaz Nanocore in 5G is the convergence of following technologies [5]:
 - Nanotechnology.
 - Cloud Computing.
 - All IP Platform.

5G aggregator aggregates all the RAN traffics and route it to gateway.

To understand the 5G protocol stack, we need to look at through in relationship with both OSI and TCP/IP as below.

Note how the open wireless architecture (OWA) layer functions as physical layer and data link layer for both the OSI and the TCP/IP stacks. Also the network layer for the 5G stack is divided between the lower and upper layers. Finally note that both the 5G and TCP/IP stacks have similar application and transport layers with a few differences.

14.2.3 Technical Challenges of 5G Technology

As the 5G technology and its based networks develop and expand, it is important to understand that the technology and its networks are fundamentally built around people and things (in the same way the Internet of Things is built around people and things—as we will see shortly). To achieve this, the technology must have the following characteristics [6]:

- Massive broadband (xMBB) that delivers gigabytes of bandwidth on demand
- Massive machine-type communication (mMTC) that connects billions of sensors and machines
- Critical machine-type communication (uMTC) that allows immediate feedback with high reliability and enables for example remote control over robots and autonomous driving

Because of these characteristics, it is expected that the technology will be able to meet the growing demand for mobile broadband which is projected to increase in the next few years. There are lots of expectations, sometimes overblown by industry jostling for financial benefits, for the 5G platform to deliver an enabling environment for growth in almost every sector of the economy. In addition, the 5G infrastructures is expected to require not only improved networking solutions but also a sophisticated integration of massive computing and storage infrastructures.

However, there is consensus that 5G and its previous wireless communication platforms vary in many different ways such that network functions may perform differently on the 5G platform. This eventually lead to 5G platform ethical, social and technical challenges including:

- *Lack of effective harmonization or similarity between the relevant protocol stack layers.* As shown in Fig. 14.2, there differences in the communication stack of 5G from those of OSI and TCP/IP. Because of these differences, there are likely to be technical problems in radio access technologies above certain GHz, for example, as this may require different physical layer numerology and different signal processing approaches, higher protocol stack layers and related network functions.
- *Network Security Considerations* The security design of current mobile systems was geared towards the build-up of a successful ecosystem, offering trustworthy communication services to users in all corners of the world [6]. As newer wireless communication technologies evolved, new security features have been added to harden the platforms to maintain trustworthiness in the presence of emerging threats, such as false radio base stations and encryption in communications, among others. As noted in [6], mobile system security has so far arguably been more of an added support function than a driver. Since users of the 5G platform have a higher bit-rates and the surface of exposure is significantly higher when it is combined with the estimated exponential growth of

OSI stack layers	5G stack layers	TCP/IP stack layers
Application	Application	Application
Presentation		
Session	Open Transport	Transport
Transport		
Network	Upper Network	Network
	Lower Network	
Data Link	Open Wireless	Data Link
Physical	Architecture	Physical

Fig. 14.2 Relationships between OSI, 5G and TCP/IP protocol stacks

the number of devices using 5G networks, 5G security raises many security concerns including [6]:

- business and trust models that involve multi-domain and multi-service models and the presence of new types of actors, among a myriad of other factors,
- New service delivery models based on virtualization, network slicing and other “aaS” technologies need to be carefully securitized.
- Users need to be considered significantly differently, with respect to what we have learned from previous network generations; trustworthiness among users cannot today be taken for granted.

14.3 The Internet of Things (IoT)

What is it? Why is it exciting so many in the technology and innovation communities? The concept of the Internet of Things (IoT) was initially proposed by Kevin Ashton in 1998 [7] while he was working at P&G to launch a line of cosmetics for Oil of Olay. Because the father of IoT, as many call him, was bothered that this one shade of lipstick in his cosmetic line always seemed to be sold out in all his London, UK local stores. He wanted to know where his lipstick was, and what was happening to it. No one could tell him. When U.K. retailers experimenting with loyalty cards with a tiny “radio-enabled” chip, later called RFID, showed him the, it gave him an idea of tracking his lipstick shade. He took the radio microchip out of the credit card and stuck it on his lipstick shade to see if a wireless network could pick up data on a card and tell him what shelf in the store the lipstick was on. By so doing, he started the forces that created the IoT. In about a decade, the simple idea and experiment have been extended to support pervasive connectivity and the integration of a variety of objects big and small creating an ecosystem of interconnected communication network whose devices or communication nodes are everyday electronic objects like mobile devices, entertainment devices in your home, fridges and temperature control devices, garage door openers, cloth and dish washers and the list goes on and on. When network connectivity is achieved,

it allows all these devices to talk to each other by sending and receiving data. This connectivity of things started long ago with the interconnection of computing devices to form the traditional computer network. Upon that a conceptual model of connectivity of all devices that can communicate and receive data forming a far wider communication network, the “Internet of Things” was born.

The conceptual model and now what is forming in reality has the potential to impact our lives in many unprecedented ways both good and bad, as most technologies are.

Gubbia et al. [8] have defined the Internet of Things as a smart environments that is made up of an interconnection of sensing and actuating devices providing the ability to share information across platforms through a unified framework, developing a common operating picture for enabling innovative applications. This smart environment is achieved by seamless ubiquitous sensing, data analytics and information representation with Cloud computing as the unifying framework. It is this ecosystem described by P. Guillemin and P. Friess in their paper “Internet of things strategic research roadmap,” as part of The Cluster of European Research Projects [9] and represented by Fig. 14.3.

Morgan [10] also sees it an environmental ecosystem that “allows for virtually endless opportunities and connections to take place, many of which we can’t even think of or fully understand the impact of today”. Because it is going to affect our lives in every possible way, known and unknown in every sphere and dimension, It is in fact as one scholar puts it, the new Industrial revolution, again.

It’s not hard to see how and why the IoT is such a hot topic today; it certainly opens the door to a lot of opportunities but also to many challenges. Security is a big issue that is oftentimes brought up. With billions of devices being connected together, what can people do to make sure that their information stays secure? Will someone be able to hack into your toaster and thereby get access to your entire network? The IoT also opens up companies all over the world to more security threats. Then we have the issue of privacy and data sharing. This is a hot-button topic even today, so one can only imagine how the conversation and

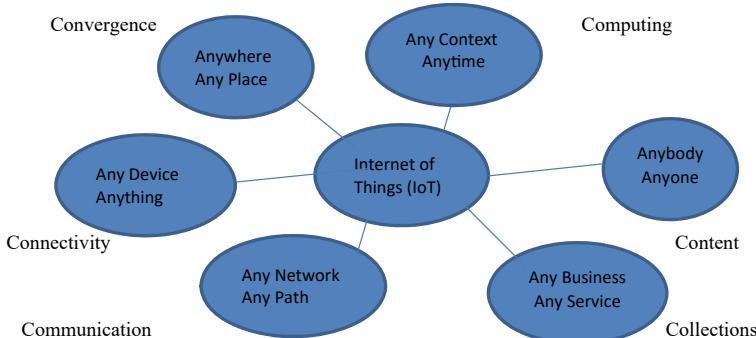


Fig. 14.3 Definition of Internet of Things (IoT)

concerns will escalate when we are talking about many billions of devices being connected. Another issue that many companies specifically are going to be faced with is around the massive amounts of data that all of these devices are going to produce. Companies need to figure out a way to store, track, analyze and make sense of the vast amounts of data that will be generated.

14.3.1 Overview and Growth of Internet of Things

In their paper, “Internet of Things (IoT): A vision, architectural elements, and future directions”, Jayavardhana Gubbia, Rajkumar Buyya, Slaven Marusica and Marimuthu Palaniswamia [8] state that the phrase “Internet of Things” was first coined by Kevin Ashton in 1999 in the context of supply chain management. Since then, it has involved to its present day meaning. But all along the way, the core essence of making a computer device, which is a node in our IoT, sense information without the aid of human intervention remains the same. In its current meaning, each node of the IoT, may it be a sensor, an actuator or a communicating device, is interconnected to other nodes in the mess that include the existing Internet, and all are able to intercommunicate, seamlessly passing and getting information to provide services for information transfer, analytics, applications, and communications using existing Internet protocols. Several technologies have converged to create the Internet of Things technologies. These technologies include those which has led to ubiquitous sensing enabled by Wireless Sensor Network (WSN) technologies, ubiquitous computing, enabled by miniature, mobile and high powered computing and communication devices and the existing Internet protocols to provide services for information transfer, analytics, applications, and communications.

In their paper, “The Internet of Things: A survey”, Atzori et al. [11] argue that the Internet of Things can be realized in three paradigms—internet-oriented (middleware), things oriented (sensors) and semantic-oriented (knowledge). But the according to Jayavardhana Gubbia et al., the usefulness of IoT can be unleashed only in an application domain where the three paradigms intersect.

With the expected continued growth of the internet, there is unanimous expectation of an enormous growth of the internet of things in the next five years and beyond. Infographics [12] estimates that by 2018, there will be 42.1 billion items connected in the IoT.

John Greenough and Jonathan Camhi both of Business Intelligence (BI) [13] look as IoT in terms of business growth predicting that IoT is the next Industrial Revolution or the Next Internet. On the future of IoT growth, they further predict the following:

- By 2020 there are likely to be 34 billion devices connected to the internet, up from 10 billion in 2015. IoT devices will account for 24 billion, while traditional computing devices (e.g. smartphones, tablets, smartwatches, etc.) will comprise 10 billion.

- Nearly \$6 trillion will be spent on IoT solutions over the next five years.
- Businesses will be the top adopter of IoT solutions. They see three ways the IoT can improve their bottom line by (1) lowering operating costs; (2) increasing productivity; and (3) expanding to new markets or developing new product offerings.
- Governments are focused on increasing productivity, decreasing costs, and improving their citizens' quality of life. Governments will be the second-largest adopters of IoT ecosystems.
- Consumers will lag behind businesses and governments in IoT adoption. Still, they will purchase a massive number of devices and invest a significant amount of money in IoT ecosystems.

14.3.2 Architecture and Networking of IoT

We defined the IoT in Sect. 14.1 as an interconnection of sensing, actuating and communication digital devices providing the ability to share information across platforms through a unified framework, developing a common operating ecosystem (COE) for enabling innovative applications. For the IoT ecosystem to function and support intended applications and accommodate the heterogeneity of devices and applications in the ecosystem, the IoT had to adopt the open standards of TCP/IP protocol suite. However, the open standards of TCP/IP protocol suite was initially developed for the wired global Internet several decades ago, as the networking solution. But as we have outlined above in our discussion of IoT, there are fundamental differences between the traditional wired computer networks and the heterogeneous combination of wired and wireless devices ecosystem. And as Shang et al. [14] observe, those differences pose significant challenges in applying TCP/IP technologies to the IoT environment, and addressing these challenges will make a far-reaching impact on the IoT network architecture. To get a good understanding of the IoT architectures and networking, we need to first understand the underlying network topology supported by the heterogeneous technologies, devices and standards. The networking technology standard currently being used in the IoT fall into three categories: (i) **point-to-point**, for example a an end device to a gateway; (ii) **star**—with a gateway connected to several end- devices by one hop links and (iii) a **mesh**—with one or more gateways connecting to several end-devices one or more hop links away as demonstrated in Fig. 14.4.

Based on these three topologies, we can cascade end-devices and gateways to get a real model of the IoT communication network architecture as shown in Fig. 14.5.

All IoT known technologies like WIFI, Bluetooth, Wi-Max, ZigBee, Z-Wave, RFID, near-field communication (NFC) and others support this communication architecture.

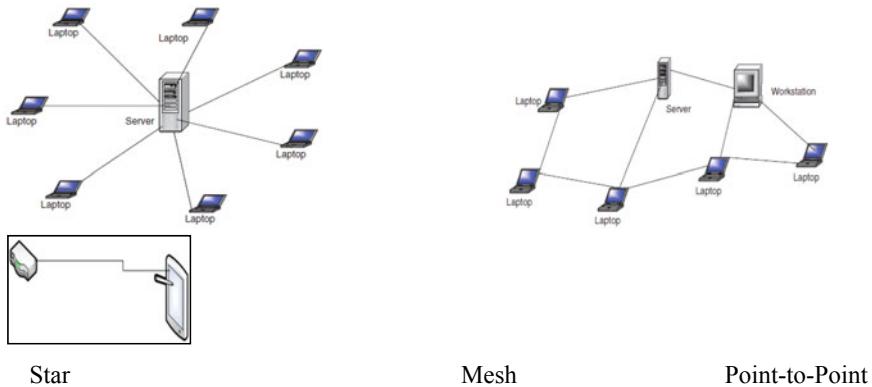


Fig. 14.4 Current IoT topologies

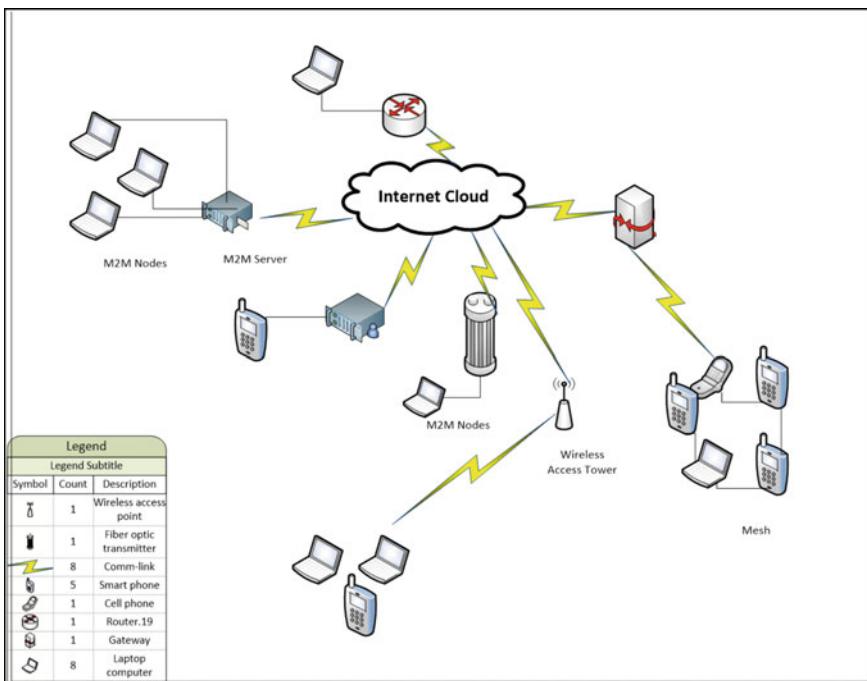


Fig. 14.5 IoT communication network architecture

14.3.2.1 Architecture and Protocol Stack of IoTs

As we will see in the coming Sect. 14.3.2, a typical TCP/IP IPv6 has a maximum transmission unit (MTU) size of 1500 bytes or higher and a near infinite address space covering up to 2^{128} unique addresses, while IoT constrained low-energy links have very small MTUs averaging around 127 bytes. Even with the two IPv6 design specifications that include (a) IPv6 of 40-byte fixed length header with optional extension headers, which causes big protocol overheads for small packets and (b) IPv6 specification requiring all IPv6-capable networks to support a minimum MTU size of 1280 bytes, typical IPv6 packets cannot be carried over the constrained IoT links. So a new 6LoWPAN protocol was defined to enable IPv6 packets to be carried on top of low-powered and lossy personal area networks (LLNs). A draft architecture for a gateway or middleware that provides interoperability between 6LoWPAN and external IPv6 networks has been defined. Other protocols have been defined to support the smooth transmission between IPv6 and low-powered IoT devices. These include for example [15]:

Constrained Application Protocol (CoAP)—this was developed by the IETF Constrained RESTful Environments (CoRE) workgroup is working. The protocol includes several HTTP functionalities although it has modified to work with low processing power and energy consumption constraints of IoT devices. Because CoAP is similar to HTTP, it also uses a universal resource identifier (URI) to identifies resources and allow the resource to be affected using similar methods such as GET, PUT, POST, and DELETE.

Figure 14.6 gives a comparative view of TCP/IP and IOT (IP Smart Objects) protocol suites.

Another way of looking at the IoT protocols is via IoT device functionality. IoT devices must communicate with each other. This is referred to as (D2D). An example for this is web services and business applications. Data on data then must

TCP/IP Protocol suite		IoT Protocol suite	
Application layer	HTTP/FTP/SMTP, etc	Application layer	CoAP
Transport layer	TCP/UDP	Transport layer	UDP
Network layer	IPv4/IPv6, RP, ICMP	Network layer	IPv6/6LoWPAN
Data Link layer	IEEE 802.3 Ethernet/802.11, Wireless LAN	Data Link layer	IEEE 802.15.4e
Physical layer	Ethernet (IEEE 802.3), Token Ring, RS-232, FDDI, and others	Physical layer	IEEE 802.15.4

Fig. 14.6 Comparative view of TCP/IP and IOT (IP Smart Objects) protocol suites

be collected and sent to the server infrastructure. This is referred to as (D2S). An example for this is in all devices where there is a need for control plane. Finally the server infrastructure has to share device data, possibly providing it back to devices, to analyze programs, or to people. This is (S2S). This includes all devices and intelligent systems. The protocols to do these services are [16]:

- MQTT: a protocol for collecting device data and communicating it to servers (D2S)
- XMPP: a protocol best for connecting devices to people, a special case of the D2S pattern, since people are connected to the servers
- DDS: a fast bus for integrating intelligent machines (D2D)
- AMQP: a queuing system designed to connect servers to each other (S2S)

Other IoT protocols include [17]:

- **Infrastructure** (ex: 6LowPAN, IPv4/IPv6, RPL)
- **Identification** (ex: EPC, uCode, IPv6, URIs)
- **Comms/Transport** (ex: Wifi, Bluetooth, LPWAN)
- **Discovery** (ex: Physical Web, mDNS, DNS-SD)
- **Data Protocols** (ex: MQTT, CoAP, AMQP, WebSocket, Node)
- **Device Management** (ex: TR-069, OMA-DM)
- **Semantic** (ex: JSON-LD, Web Thing Model)
- **Multi-layer Frameworks** (ex: Alljoyn, IoTivity, Weave, Homekit)

14.3.3 Challenges of Using TCP/IP Architecture Over the IoT

As we just stated above, the IoT ecosystem of heterogeneous devices, wired, wireless and restricted, using the traditional TCP/IP (though IPv6) meant for wired devices, presents a growing number of challenges in IoT networking that are likely to grow as the IoT ecosystem grows. Some of the issues causing these challenges are easy to see. Others are not. Most of the challenges are brought about by the IoT inherent heterogeneous low-battery powered wireless devices, the multi-link subnet model and the mesh network nature of the ecosystem that requires new scalable routing mechanisms. These challenges are thoroughly discussed by Wentao Shang, Yingdi Yu and Ralph Droms in their paper “Challenges in IoT Networking via TCP/IP Architecture” as flows [14]:

- **Maximum transmission unit (MTU) size**—While a typical TCP/IP IPv6 MTU has a minimum size of 1500 bytes or higher, the IoT constrained low-energy links have very small MTUs averaging around 127 bytes. Along with size, the IPv6 specification, of two design decisions that utilizes either (a) IPv6 of 40-byte fixed length header with optional extension headers, which causes big protocol overheads for small packets or (b) IPv6 specification requiring

all IPv6-capable networks to support a minimum MTU size of 1280 bytes, is unrealistic for the IoT constrained links.

- **Multi-link subnet model**—the current subnet model of both IPv4 and IPv6 considers two types of Layer-2 networks: multi-access link, where multiple nodes share the same access medium, and point-to-point link, where there are exactly two nodes on the same link. Both of them assume that the nodes in the same subnet can reach each other within one hop. However, the current IoT mesh network contains a collection of Layer-2 links joined together without any Layer-3 device, like routers, in between. This essentially creates a multi-link subnet model that is not anticipated by the original IP addressing architecture.
- **Multicast efficiency**—A lot of IP-based protocols make heavy use of IP multicast (one-to-many or many-to-many where information is addressed to a group of destination computers simultaneously- see Sect. 5.3.4) to achieve one of the two functionalities: notifying all the members in a group and making a query without knowing exactly whom to ask. However, supporting multicast packet delivery is a big challenge for constrained IoT mesh networks. First, most wireless MAC protocols disable link layer ACK for multicast; consequently lost packets are not recovered at link-layer. Second, multicast recipients may experience different data transmission rate due to the coexistence of multiple MAC protocols and/or the link-layer rate adaptation; therefore the sender has to transmit at the lowest common link speed among all receivers. Third, IoT nodes may switch to sleeping mode from time to time to conserve energy, thus may miss some multicast packets. Lastly, when nodes are connected through a mesh network, a multicast packet needs to be forwarded over multiple hops along many paths, potentially waking up many sleeping nodes and overloading the already-scarce network resource.
- **Mesh network routing**—The topologies of typical IoT networks fall into three categories, as seen in Fig. 14.1: star topology, mesh (peer-to-peer) and point-to-point. The routing configuration is straightforward on a star and point-to-point networks where the hub node in a star topology and one of the two nodes in a point-to-point topology can act as the default gateway for the peripheral nodes. However, this limits the signal coverage of a single hub node in these two deployment topologies, making them unsuitable for applications that need wider coverage. The mesh topology, on the other hand, enables larger coverages by having the nodes relay the packets for each other. All mesh nodes cooperate in the distribution of data in the network. Mesh network routing can be supported at either the link layer or the network layer. The link-layer approach, called *mesh-under* in the IETF terminology [14], relies on Layer-2 forwarders to join multiple links into a single “one-IP-hop” subnet. The network-layer approach, called *route-over*, instead relies on IP routers to forward packets across multiple hops. IoT suffers from a ***Transport layer problem***. The Internet’s TCP/IP architecture transport layer provides *congestion control and reliable delivery*, both of which are implemented by TCP, the dominant transport layer protocol on the Internet. TCP efficiently deliver a large bulk of data over a long lived point-to-point connection without stringent latency requirement. It models

the communication as a byte stream between sender and receiver, and enforces reliable in-order delivery of every single byte in the stream. However, IoT applications usually face a variety of communication patterns which TCP cannot support efficiently. First, due to the energy constraints, devices may frequently go into sleep mode, thus it is infeasible to maintain a long lived connection in IoT applications. Second, a lot of IoT communication involves only a small amount of data, making the overhead of establishing a connection unacceptable. Third, some applications may have low-latency requirement, which may not tolerate the delay caused by TCP handshaking.

- **Resource discovery**—The resource-oriented communication model usually requires a resource discovery mechanism, whereby the applications can request or invoke operations on the resources. The solution for resource discovery in traditional IP networks is DNS-based Service Discovery (DNS-SD) [14]. However, this solution has several limitations in supporting IoT applications. First of all, DNS-SD aims to support service discovery, where the service usually refers to a running program. In contrast, the resources in the context of IoT covers a broader scope: besides services, it may also refer to IoT devices, sensor data, etc. Therefore, the IoT resource discovery requires a more general approach to identify heterogeneous resources. For example, instead of using DNS records, CoAP adopts a URI-based naming scheme to identify the resources (like in HTTP). Based on that, the IETF core WG has developed CoRE-RD, a CoAP-based resource discovery mechanism that relies on less constrained resource directory (RD) servers to store the metainfo about the resources hosted on other devices. Secondly, traditional service discovery often relies on multicast when dedicated services such as DNS and CoRE-RD are not available in the local environment. For example, DNS-SD uses Multicast DNS (mDNS) [14] as the carrier of communications for service discovery and name resolution within the local network. However, link-local multicast has efficiency issues in IoT environments.
- **Caching**—The TCP/IP communication model requires that both the client (resource requester) and the server (resource holder) are online at the same time. However, in IoT scenarios, the constrained devices may frequently go into sleeping mode for energy saving. Moreover, the dynamic and/or intermittent network environment usually makes it difficult to maintain stable connections between communicating parties. Consequently, the IoT applications often rely on caching and proxying to achieve efficient data dissemination. The selected proxy node can request the resources on behalf of the sleeping nodes and store the response data temporarily until the requesting nodes wake up. The cached contents can also be used to serve similar requests from other nodes who share the same proxy, which saves network bandwidth and reduces response latency. The resource origin server may also appoint some proxy nodes to handle the requests on its behalf (called reverse-proxy) so that it can reduce the client traffic and may go offline when it need to. While it is helpful, the application-level caching implemented by CoAP and HTTP has several limitations in the IoT environment. First, the clients need to explicitly choose a forward- or

reverse-proxy node in order to utilize the content caching capability. Second, in dynamic network environments where the connectivity is intermittent, the pre-selected proxy point may become totally unreachable. When the network topology changes, the clients need to reconfigure or re-discover the proxies, or otherwise stop using caches and proxies at all. Third, the caches and proxies break the end-to-end connections assumed by the current security protocols, making it even harder to protect the application data.

14.3.4 IoT Governance, Privacy and Security Challenges

As we have pointed out throughout this chapter, an inherent characteristic of the IoT is its heterogeneity resulting from a plethora of things with different data communication capabilities like protocols and hardware, data rates, reliability and others; computational, storage and energy capabilities; diversity in the types and formats of data like audio, video, text, numeric, and streams; and IoT standards including device standards, standards to represent data, IEEE projects on IoT standards, ITU and ISO IoT standards and others [18]. This diversity in devices, service and protocols, present challenges unseen and unprecedented in the modern communication.

14.3.5 Governance and Privacy Concerns

As the IoT grows, it presents us with several challenges including global governance, individual privacy, ethics and of course security. These are the most critical issues in the growth of IoT. As it grows, the IoT is expected to involve multiple stakeholders around the globe. It is important to understand that the meanings of and what defines these issues are differently understood and defined around the globe. So we will deal with the most widely accepted definitions and meanings here. Globally, governance is mostly understood to refer to the rules, processes and behavior that affect the way in which powers are exercised, particularly as regards openness, participation, accountability, effectiveness and coherence [18]. These five *principles of good governance*, have been already applied to the Internet for specific aspects and there are already organizations like IETF, ICANN, RIRs, ISOC, IEEE, IGF, W3C, which are each responsible and dealing with every specific area [19]. But currently this is not the case with the IoT. What this is pointing to is that the governance of the current IoT posses an array of problems for all those connected to the Internet, the most serious of which are security threats and attacks originating from and targeting both Internet-connected endpoints, and data privacy risks posed by those same devices. Consider an IoT with 70 billion wireless like standalone and embedded sensors and wired devices predicted in the next three to five years, all capturing, storing and communicating data. A number of questions arise. For example who owns that data? If those devices communicate

with your mobile device in the public commons, who owns that data into and on your smart device? Where is the data exchanged with your device going to go? What is it going to be used for? From that point, if the data from or into your smart device is automatically combined with data from a smart passing car, what happens? Do others come to know about you? Do others come to spoof into your devices later? For those wearing medical devices that monitor their vital signs, what about their medical data? This raised a million security and privacy issues with no immediate answers. All these are happening because of lack of a central or at least coordinated distributed authorizes to harmonize governance of the IoT.

However, everything about the governance of IoT is not bad, there promising efforts and initiatives in different places like north America and Europe that are developing policies and protocols that will eventually archive these governance goals.

14.3.6 Security Challenges

Security is critical to IoT applications due to their close interaction with the physical world. In Internet communication, based on TC/IP protocols, IP has the task of delivering packets from the source host to the destination host solely based on the IP addresses in the packet headers. For this purpose, IP defines packet structures that encapsulate the data to be delivered. It also defines addressing methods that are used to label the datagram with source and destination information. As a most widely used secure protocol in IP, TLS and its datagram variant DTLS, are the main security protocols offering end-to-end secure communications between a server and client. TLS, with its two main constituent protocols, the Handshake protocol, responsible for key exchange and authentication and the Record protocol, responsible for a secure channel for handling the delivery of data, makes the security of all IP-based communications channel-based security. The secured-channel solutions, however, do not fit into the IoT environments for several reasons.

- The first issue with **channel-based security** is the overhead of establishing a secure channel. Both TLS and DTLS require two or more rounds of security handshake to authenticate a channel and negotiate the security parameters, before the first application data is sent out. The second issue is that both ends of a channel have to maintain the states of the channel until it is closed. This may impose a high pressure on memory usage when a device needs to communicate with many peers simultaneously in a densely-meshed network. Third, channel-based security does not guarantee the security of request-response once the application data get out of the channel. This is most troublesome when the *middleboxes*, like caches and proxies, are deployed to cache the application data. The resource owners need to trust the middleboxes to enforce the access control policies correctly, while the resource requestors need to trust the middleboxes to provide authentic data without tampering. The limitations above highlight the need for a different security model for IoT applications.

- **Insufficient authentication/authorization**—If recent attacks on the Internet, using smart house monitoring camera, resulting in distributed denial of service (DDoS) is any evidence, the IoT with its growing mesh of heterogeneous devices, whose users and devices rely on weak and simple passwords and authorizations, is a growing security quagmire.
- **Lack of transport encryption**—Most devices fail to encrypt data that are being transferred, even when the devices are using the Internet.
- **Insecure web/mobile interface**—Most of the billions of IoT-based devices connect on the Internet using bridging communication protocols and device management schemes that do not do an effective job. See more details in Sect. 14.3.2.

14.3.7 Autonomy

High heterogeneity and complexity and lack of dynamic and scalable management schemes in the IoT due to its plethora of sometimes constrained devices, with different data communication capabilities create a challenge in the manual maintenance of a large number of devices becomes inefficient, and demands the presence of intelligent and dynamic management schemes. According to Ashiraf and Habaebi [20], strong autonomy in IoT can be realized by implementing self-managing systems. Self-management is the property of a system to achieve management and maintenance of its resources intrinsically and internally. It is achieved through levels of decision making including access management, device management as well as service management. This thus should lead to all devices in the IoT being aware of their owners' preferences and autonomously make decisions on behalf of their owners and at the same time cooperate with other devices on including securing network communication.

14.3.8 Computational Constraints

One of the characteristics of IoT is its heterogeneity and complexity as it connects to billions of sometimes constrained devices running different communication protocols and management schemes. Low level devices on the fringes can be of limited power sometimes of less than 10 KBs of RAM, which is sometimes orders of magnitude lower than an ordinary desktop computer with GIGs of RAM. This presents data transfer, computation and communication challenges. So in cases where high demand computations cannot be handled by the low power devices, a delegation of operations may be required.

14.3.9 Discovery

With the rapid growth of devices connected to the IoT, expected to hit 70 billion in the next few years, challenge for search and discovery for available services is increasingly becoming an impediment to the growth of IoT and will diminish future expected benefits of the IOT. Moreover, discovery methods currently being used in the Internet are not flexible to accommodate a growing regime of new services and they are not capable of searching the heterogeneous devices running different discovery protocols. Therefore, we need new discovery technologies that are more expressive and able to evolve over time.

Discovery in the IoT is the process that enable application to access the IoT data without the need to know the actual source of data, sensor description, or location. According to Arkady Zaslavsky, Prem Prakash Jayaraman, the discovery process can be defined as two successive loops [18]:

- **Foraging loop.** Data sources are identified and assessed, where the relevant data is extracted and formatted into consumable form.
- **Sense-making loop.** The extracted data is analyzed and exploited to provide answers around a specific problem.

The challenge is then to develop a scalable framework (or architecture) along with protocols to provide complete capabilities, which work for all those who will use the IoT.

14.3.10 Trust Relationships

We have already seen and discussed the connectivity and heterogeneity of the IoT. We know that IoT connects to billions of devices with high connectivity complexities and challenges. IoT end devices play a variety of roles and perform many functions for the device owner. Some devices are wired other are wireless. Some are low powered other have access to full power. To enable communications with all these devices, there is a need for some degree of intelligence in these devices. The growth of embedded intelligence behavior in the end-devices, as an extension of the device owner relationship, will increase and indeed become ubitiquous as the IoT plethora of things with different data communication capabilities grows. As the strong relationships and embedded intelligence between end-devices and their owner grows, a citizen (user) relationship is created and introduced into the IoT. The “things” in an IoT are indeed the end-devices. There are the new entities (new ontologies). Now these new entities are endowed with identity, connectivity, intelligence, and agency with and through which relationships.

These *human-IoT* relationships create a relationship-trust mesh in the IOT which result into a multitude of questions of a social, ethical, and legal nature. Questions such as [21]:

- What threats are caused by delegating fundamental aspects of humanness?,
- How can we preserve the human capability to freely act and make choices in the IoT?

A lot more issues are and will continue to be raised as the IoT grows.

14.4 Ethical, Social and Legal Impacts of 5G and IoT

So far in this chapter we have been discussing the development and layout of the latest two technologies to arrive in our neighbourhood: 5G wireless communication and the Internet of Things (IoT). We have discussed how individually these technologies are changing the way we currently live and are likely to transform the landscape of human experience in the coming years. We have individually discussed the concerns and shortcomings of these technologies on our way of life as we know it today. In this section, we are going to look at the combined effect on these two technologies on the ethical fabric and framework of our society. 5G and the Internet of Things (IoT) bring with them a host of life-altering ethical questions, issues, and dilemmas. Just like those technologies before these have created ethical muddles in our decision making processes and ethical questions and situations, the same is expected from these two technologies but this time, probably more intense due to a number of characteristics that make these two technologies far different from previous ones including speed with which our world is migrating to cyber space and pervasiveness.

14.4.1 Environment

The impacts of RF on wildlife will escalate by orders of magnitude with a move to 5G and the IoT.

14.4.2 E-waste

E-waste, the term given to discarded electronic appliances, is often shipped by developed nations to poorer countries such as Ghana.

14.4.3 Conflict Minerals

Minerals like Cobalt are essential part of most mobile devices and are on very high demand around the world fuelling conflicts and environmental degradation around the world.

14.4.4 Healthy Issues Emanating from 5G and IoT Technologies

Wireless radiation is a huge health problem that continues to plague modern society. Current wireless technologies of 2G, 3G and 4G used by our cell phones, computers, and wearable technology are creating radio frequency exposure which poses a serious health risk to the environment we live in and to all of us. 5G and its predecessors are the backbone technologies for the Internet of Things (IOT) as it connects to billions of devices. Scientists have been studying the health effects of 5G and wireless radiation and are deeply concerned with their findings and are calling for a stop to the rollout of 5G, as well as a halt to the proposed increase in radio frequency radiation exposure to the public [22].

According to Melissa Arnoldi, who leads AT&T's efforts [22] “if it’s not already in your neighbourhood, it’s coming.” “5G uses high-frequency waves that support faster speeds but don’t travel as far as current wireless frequencies. So instead of relying on large cellphone towers spread far apart, they need “small cell” sites that are much closer together.”

14.4.5 Ethics

Digitization and global networks resulting from both 5G and IoT are spreading like wild fire leaving bigger impact on lives across the globe. Questions are being raised on privacy invading devices. Lives are changing rapidly as intrusive devices are exposing private lives of unsuspecting bystanders bringing into question the evolving ethical framework of societies. The changing ethical, legal and social landscape is resulting in significant legal and ethical questions that will globally impact every human being their rights, privacy and liabilities. Global society has never been in a worse ethical muddle, legal predicament and social complexities that call for open and public discourse.

Exercises

From self-driving cars to factory robots, engineers are imagining new ways of connecting our world through 5G and IoT-enabled machines that integrate production processes.

1. In a short 3 pages paper discuss how this is likely to happen.
2. Also using the same scenario above, cars in, say a four way intersections, will be able to talk to each other and negotiate who goes first without involving the drive. What are the likely dangers of this?
3. With 5G and IoT, the toaster in your house can wake you up to tell you that your bread slice is just about ready. In man-machine interdependence so created, discuss the ethical, social and security implications of this simple stream of actions. Discuss what our role as humans is.
4. Jack Williamson in *With Folded Hand* portrays a world ruled by robots, which seem benign but must follow and exist to discharge the Asimovian Prime Directive. The Prime Directive is: "to serve and obey, and guard men from harm." In the story robots replicate themselves and do all the jobs the man wants them to do- until he realizes the mistake he made to create the robots in the first place. They just made him useless. Is the marriage between 5G and IoT technologies likely to produce this utopia for humanity?
5. In this chapter, we call the marriage between 5G and IoT technologies a social, ethical and security quagmire. Do you agree?

Advanced Exercises

1. What's the biggest risk associated with the IoT and 5G on society?
2. What factors would most influence and accelerate the benefits of the IoT? How about 5G?
3. Will IoT, including devices that make it, be secure? Perhaps this is the most difficult question to answer. Do you know why?
4. What is the role of 5G technology in enhancing wireless technology?
5. In any communication regime, privacy issues play a vital role. With IoT, or 5G or any the combination of these technologies, how will privacy be assured? Or can it?
6. With the ubiquitous communication brought about by IoT, 5G or both, interoperability is critical. Can IoT architecture guarantee interoperability? What about 5G?

References

1. Lancaster University. *G5: Moving to the next generation in wireless technology*. ScienceDaily. ScienceDaily, 30 April 2015. www.sciencedaily.com/releases/2015/04/150430082723.htm
2. Z. Ma, Z. Quan, Z.Z. Guo, D. Ping, Z. Fan, H.C. Li, Key techniques for 5G wireless communications: network architecture, physical layer, and MAC layer perspectives. *Sci. China Inf. Sci.* **58**(4), 1–20
3. Z. Ding et al., *A Survey on Non-Orthogonal Multiple Access for 5G Networks: Research Challenges and Future Trends*. <https://ieeexplore.ieee.org/document/7973146/authors>
4. J. Mundy, *What Is Massive MIMO Technology?* <https://5g.co.uk/guides/what-is-massive-mimo-technology/>
5. A. Imthiyaz, *5G the Nanocore*. http://www.telecoms.com/wp-content/blogs.dir/1/files/2011/05/5G_The_NanoCore.pdf
6. 5G PPP Architecture Working Group, View on 5G Architecture. <https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-5G-Architecture-WP-For-public-consultation.pdf>
7. K. Maney, *Meet Kevin Ashton, Father of the Internet of Things*. <http://www.newsweek.com/2015/03/06/meet-kevin-ashton-father-internet-things-308763.html>
8. J. Gubbia, R. Buyya, S. Marusica, M. Palaniswamia, *Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions*. Elsevier. <http://www.sciencedirect.com/science/article/pii/S0167739X13000241>
9. P. Guillemin, P. Friess, Internet of things strategic research roadmap, in *The Cluster of European Research Projects*, Tech. Rep., September 2009. http://www.internet-of-things-research.eu/pdf/IoT_Cluster_Strategic_Research_Agenda_2009.pdf
10. J. Morgan, *A Simple Explanation of 'The Internet Of Things'*. <http://www.forbes.com/sites/jacobmorgan/2014/05/simple-explanation-internet-things-that-anyone-can-understand/#5939c976824>
11. L. Atzori, A. Iera, G. Morabito, The Internet of Things: a survey, in. *Computer Networks*. Elsevier. http://ac.els-cdn.com/S1389128610001568/1-s2.0-S1389128610001568-main.pdf?_tid=1094eeaa-85ae-11e6-a405-00000aab0f01. http://ac.els-cdn.com/S1389128610001568/1-s2.0-S1389128610001568-main.pdf?_tid=1094eeaa-85ae-11e6-a405-00000aab0f01&acdnat=1475089513_ff39eabceaa7caece0fbe703937e25c1. http://ac.els-cdn.com/S1389128610001568/1-s2.0-S1389128610001568-main.pdf?_tid=1094eeaa-85ae-11e6-a405-00000aab0f01&acdnat=1475089513_ff39eabceaa7caece0fbe703937e25c1
12. Infographic: The Growth of the Internet of Things. <https://www.ncta.com/platform/industry-news/infographic-the-growth-of-the-internet-of-things/>
13. J. Greenough, J. Camhi, Business Intelligence. Here are IoT Trends That Will Change the Way Businesses, Governments, and Consumers Interact with the World, 29 Aug 2016. <http://www.businessinsider.com/top-internet-of-things-trends-2016-1?IR=T>
14. W. Shang, Y. Yu, R. Droms, *Challenges in IoT Networking via TCP/IP Architecture*. NDN Technical Report NDN-0038 (2016). <http://named-data.net/techreports.html>
15. R. Sutaria, R. Govindachari, *Understanding the Internet of Things*. <http://electronicdesign.com/iot/understanding-internet-things#IoT>
16. S. Schneider, *Understanding the Protocols Behind the Internet of Things*. <http://electronicdesign.com/iot/understanding-protocols-behind-internet-things>
17. IoT Standards and Protocols. <http://www.postscapes.com/internet-of-things-protocols/>
18. A. Zaslavsky, P.P. Jayaraman, The internet of things: discovery in the internet of things, in *Ubiquity* (2015), pp. 1–10
19. IERC. Internet of Things IoT Governance, Privacy and Security Issues. European Research Cluster on the Internet of Things January 2015. http://www.internet-of-things-research.eu/pdf/IERC_Position_Paper_IoT_Governance_Privacy_Security_Final.pdf
20. Q.M. Ashiraf, M.H. Habaebi, *Introducing Autonomy in Internet of Things*. <http://www.wseas.us/e-library/conferences/2015/Malaysia/COMP/COMP-27.pdf>

21. I. Kounelis, G. Baldini, R. Neisse, G. Steri, M. Tallacchini, Â.G. Pereira, Building Trust in the Human–Internet of Things Relationship. IEEEExplore. IEEE Technology and Society Magazine (2014). <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6969184>
22. A. Walia, The damaging effects of 5G wireless on your health, in *Collective Revolution*. <https://www.collective-evolution.com/2018/06/18/the-damaging-effects-of-5g-wifi-on-your-health/>



Metaverse, the Evolving Realities and Ethics

15

Abstract

This chapter considers a new and exuberant but not yet physically existing technology platform that is already exciting young people, technologists, high tech industries and probably a big percentage of us, if only we come to understand what it really is. The chapter covers a detailed discussion of the existing technologies and a combination of the architectures that will support this platform, its functions and its expected social and ethical implications. Two issues and of major concern as far as social and ethical implications are concerned. These are immersion and telepresence. A discussion is started by a more thorough discussion as the reader grasps the concepts more and better.

Learning Objectives

After reading this chapter, the reader should be able to:

- Understand the evolving concept of a metaverse.
- Identify and understand the various technologies that support the architecture that make the metaverse what it is and operational.
- Understand the two concepts of telepresence and immersion and the technologies that support them.
- Describe the positive and negative implications of these two concepts.
- Learn the social and ethical impacts that a fully functioning metaverse will have on society.
- Analyze the role and risks these two concepts play in the functioning of the metaverse.
- Articulate the impact of the input deficit from diverse populations expected to operate in the metaverse.

Scenario 5: The Body Double of an Electronic Scientist

Xander Karungi has been playing video games for a number of years. He spends long hours as an avatar working as a scientist in a highly sophisticated Kandia company. Kandia is one of the many moons of a fictional look-alike of the ice giant planet Uranus in our solar system. As part of his assignments, he usually deals with urgent calls to penetrate and retrieve data from Kandia businesses. He is good at it and happy with his job.

One morning, Xander wanted to pay his bills online. However, he was not sure whether he had enough funds on his account and it has been a while since he last checked. So he did a quick check. He was surprised to see small but frequent transfers of funds into his account. He was doing it but he was not sure it would ever work. He found a way of transferring little funds at a time into his earthly account from the Kandia company accounts. Now he knows it is working. Should he stop it or continue and become rich. There are very low chances that he will ever be caught in the act. He knows what he is doing. He is after all a forensic scientist in real life. More importantly, he assures himself, the companies he is working for are not real. He is, therefore, not doing anything wrong. Is he?

Discussion Questions

- *What do you think is happening here?*
- *Is Xander right that companies and accounts he is getting the funds from are not real. Therefore, he is not wrong. Is it the finder's keeper?*
- *Is the platform owner that Xander uses responsible?*

15.1 Introduction

In simple terms metaverse is a virtual-reality immersive n-dimensional space in which citizens in this communicate and interact with each other via computer-generated platforms, gadgets and computer generated virtual reality and augmented reality via wearable headsets and headwears. On another front, it is a hypothetical iteration and extension of the Internet to create a universal and immersive virtual and augmented reality facilitated by wearable gadgets. In reality, this immersive space does not exist yet. However, it is a very exciting vision of what the future will be like where personal and commercial life is conducted digitally in parallel with our lives in the physical world. Though it is not yet a reality, the truth is that it is getting an oversized attention and investment in the tech sector and beyond.

15.2 Definition

According to Wikipedia [1], the word metaverse was coined in Neal Stephenson's 1992 science fiction novel *Snow Crash*, where humans, as programmable avatars, interact with each other and software agents, in a three-dimensional virtual space that uses the metaphor of the real world. Stephenson's intent was probably to describe a virtual reality-based platform that extends the internet. Read Neal Stephenson's fiction novel to get the feel of the metaverse.

So far, it is difficult to precisely define a metaverse. The precise definition, like what it actually represents, is evolving, in flux every moment. It will take sometime before a good and fitting definition is found and accepted. Until such a time as we can know and understand what it is, the precise definition will remain changing. So let us attempt to give a collage of definitions. Pick what excites you more. In the simplistic way, metaverse is a digital world which extends the physical world and in which anything physical is blended with the virtual, extending our senses of sight, sound, and touch. However, our senses of sight, sound and feel are different in the metaverse from what it is in the physical world. In the metaverse we feel our presence (sight, sound and touch) as *avatars* via the *telepresence* concept. Let us dive more into this in the coming sessions. With time more precise and better definitions will emerge.

15.3 The Evolution of Metaverse

Although metaverse is not yet here, its potential is anticipated to be overwhelming. Led by big tech companies, every aspect of commerce is gearing itself and taking position to be at the forefront when the real thing happens. The hope is to be at the forefront of profound changes that the Metaverse promises bring in relation to digital interactions between people and businesses, the outcome of which, we have never experienced. At least that is the dream.

15.4 The Architecture of Metaverse

To understand the architecture of the metaverse, we need to start by understanding how the metaverse is designed. Remember that as of writing this, metaverse is still an idea of an n-dimensional immersive virtual space where people from physical work can comfortably interact with each other. This means roughly that the metaverse is composed of three components: the 3-D virtual space, the 3-D physical space, which we refer to as the world and the interface of the two spaces. There are actors, services and activities in both spaces. However, the actors originate from the physical space via the interaction interface and carry out activities and services in the virtual space. See Fig. 15.1.

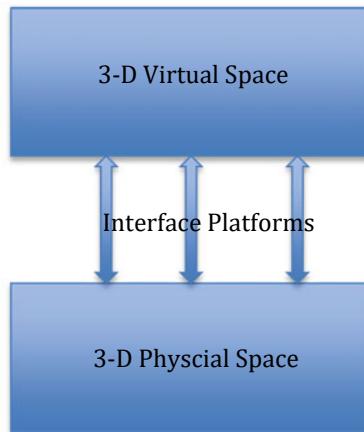


Fig. 15.1 The route representation of metaverse architecture

The **Virtual Space** contains:

- Avatars
- Immersive user experiences
- User generated content and activities—games and other services

The **Transition Platforms** supported by the **Internet of Things (IOT)** consists of:

- User designed transitional platforms—to enable users to access the virtual space
- Virtualization enabling devices—sometimes wearable gear.

The **Physical World** consists of:

- Computer networks,
- Computer-based technologies to support virtualization, augmented reality, virtual reality, internet of things and cloud computing
- Computer servers for storage of content, databases and Blockchains

After understanding the metaverse architecture we need now to focus on how these components are glued together to bring about a smooth functioning of the metaverse. The magic is a result of the working of two technologies, the *avatar* and *platform*.

15.4.1 Metaverse Avatars

The presence of any user on any metaverse platform, will be as an avatar, essentially a manifestation of the user within the metaverse. A telepresence of the user who is physically outside of the metaverse. This avatar can be and look like anything the user wants and can imagine and the metaverse platform can support. Gamers, digital artists and ordinary users, with sophisticated apps on social media devices, are already creating avatars that sometimes look exactly like them or as wild as their imagination can take them.

15.4.2 Metaverse Platforms

Immersive experiences in the metaverse is based on the platform and platform technology one is using. Remember, as we defined metaverse earlier, it is a virtual space, any virtual space, in which users interact within a computer-generated environment. That computer generated environment is the metaverse platform for that user. It may be different for different users depending on the technology they are using to generate that environment. Because of that, different users, depending on the technology they are using, hence the platform, may have different metaverse experiences. Even users on the same platform may have different experiences depending on the avatar they are using. At the writing of this book, here are the most popular metaverse platforms:

- **Decentraland**—One of the original metaverse platforms, is a browser-based 3D virtual world with its own digital currency, the MANA cryptocurrency using the Ethereum Blockchain, where users may buy virtual plots of land in the platform as non-fungible token (NFT). It opened to users in February 2020.
- **Illuvium** is an upcoming virtual game using the ILV token based on the Ethereum blockchain.
- **Sandbox**—like Illuvium, Sandbox is a virtual world game platform where players can build, own, and monetize their gaming experiences using the Ethereum blockchain.
- **Axie Infinity**—this also is an NFT-based online video game also using the Ethereum blockchain. Users of the platform collect and mint NFTs.
- **Cryptovoxels**—is a Metaverse platform just like Decentraland. It allows the user to purchase virtual parcels of land using the Ethereum blockchain.
- **Metahero**—Unlike gaming Metaverse platforms, this is a marketplace platform that allows users to buy and sell cryptocurrencies. It works just like a stock exchange trading platform. It uses HERO as its native cryptocurrency also based on the Ethereum blockchain network.
- **Star Atlas**—unlike platforms we have discussed before, this is a decentralized gaming platform based on a Solana blockchain.

- **Facebook**—Although Facebook has rebranded as a Meta, its metaverse platform **Horizon Home** is not fully developed. Once developed, it will let users of the traditional Facebook social platform invite your friends and relax together, watch movies, and jump into other apps or games together in its Metaverse.
- **Google**—also owns the AR prototypes platform.

You notice that different platforms are using different technologies and offer different features and services to the Metaverse users. To purchase these services, users must have a platform-based currency to use. That is why, we will introduce other new terminologies including *cryptocurrency* and *blockchain*. We will discuss these in the coming sections.

To bring it close to understanding the working of the metaverse based on the architecture and the technology we just discussed, think of yourself playing a video game in 3D enabled by having wearable gadgets provided to you by your game maker. You are an avatar and your game maker has provided you with a platform that has made it possible for your virtual interaction. You probably downloaded the game via a user account to your platform. The physical world is providing the Internet and all other features that support the interactions. The avatars in metaverse are carrying out activities that depend on the technology provided by the platform that gave them access to the 3D space. The avatars are also enjoying an all-inclusive *immersive* digital experience resulting from two fundamental technologies: *augmented reality* and *virtual reality*.

Augmented Reality

Augmented reality is an interactive experience that *combines the real world and computer-generated realities or content*. The content can span multiple sensory modalities, including visual, auditory, haptic, somatosensory and olfactory. It is an enhanced version of the real physical world content that is achieved through the use of digital visual elements, sound, or other sensory stimuli and delivered via a digital media.

Virtual Reality (Virtual Presence)

In 12.4. *Virtual Reality*, we define virtual reality (VR) or virtual presence (VP) if you prefer, as a type of virtualization technology that employs computer-controlled multisensory communication capabilities that allow more intuitive interactions with data and involve human senses in new ways. Virtual reality is also a computer-created environment immersing users and allowing them to deal with information more easily. The sense of presence or immersion, due to virtualization, is a critical feature distinguishing virtual reality from other computer-based applications. Metaverse uses *immersive virtual reality*. This involves the use of computer interface devices such as a head-mounted display (HMD), fiber-optic wired gloves, position tracking devices, and audio systems providing 3D (binaural) sound. As the user immerses into the new environment, there is immediate personal experience. For example, being in a simulated airplane cockpit gives a first person

experience of flying an aircraft. Similar experiences may be gotten when one is wearing goggles and gloves or using a joystick in a simulated NASCAR driving. Many current video games already give these kinds of personal experiences to the gamers.

These technologies are provided by the platform anchored in the *interface layer* and itself supported by the facilities and attributes in the physical world layer.

Now, to understand how these two technologies enable the transformation of a *physical world layer* user into a *virtual world layer* avatar, one needs to understand the other supporting technologies that are the bedrock of these two: computer networks, virtualization, internet of things and cloud computing.

15.4.2.1 Virtualization

Virtualization is a technology concept that lets a user create applications and services that go beyond the capacity and limitations of traditional hardware. This is usually done by cleverly distributing the capabilities of the physical machine to many many users or environments. In Chap. 11, we defined virtualization of computing resources as a process in which software creates virtual machines (VMs), including a virtual machine monitor called hypervisor, that allocate hardware resources dynamically and transparently so that multiple operating systems, called guest operating systems, can run concurrently on a single physical computer without even knowing it. The virtualization concept once applied to existing hardware and software, substantially increases the performance of computing systems, through division of the underlying physical computing resources into many equally powerful virtual machines. This has popularized the technology that now drives the metaverse.

15.4.2.2 Internet of Things (IOT)

In this chapter, we shall define Internet of Things (IOT) as a smart environment that is made up of an interconnection of sensing and actuating devices providing the ability to share information across platforms through a unified framework, developing a common operating picture for enabling innovative applications. This smart environment is achieved by seamless ubiquitous sensing, data analytics and information representation with Cloud computing as the unifying framework. It is an ecosystem of computing devices embedded in everyday objects, enabling them to send and receive data via the Internet.

Architecture and Networking of IOT

The IOT is an interconnection of sensing, actuating, and communication digital devices providing the ability to share information across platforms through a unified framework, developing a common operating ecosystem (COE) for enabling innovative applications. For the IoT ecosystem to function and support intended applications and accommodate the heterogeneity of devices and applications in the ecosystem, the IoT had to adopt the open standards of TCP/IP suite. However, the open standards of TCP/IP suite were initially developed for the wired global Internet several decades ago, as the networking solution. But as we have outlined

above in our discussion of IoT, there are fundamental differences between the traditional wired computer networks and the heterogeneous combination of wired and wireless device ecosystem. These differences pose significant challenges in applying TCP/IP technologies to the IoT environment, and addressing these challenges will make a far-reaching impact on the IoT network architecture. To get a good understanding of the IoT architectures and networking, we need to first understand the underlying network topology supported by the heterogeneous technologies, devices, and standards. The networking technology standard currently being used in the IoT falls into three categories: (1) *point-to-point*, for example, an end device to a gateway; (2) *star*, with a gateway connected to several end devices by one hop links; and (3) a *mesh*, with one or more gateways connecting to several end devices one or more hop links away. Based on these three topologies, we can cascade end devices and gateways to get a real model of the IoT communication network architecture.

All IoT known technologies like Wi-Fi, Bluetooth, WiMax, ZigBee, Z-Wave, RFID, near-field communication (NFC), and others support this communication architecture.

15.4.2.3 The Computer Network and the Internet

The Internet is a global computer network interconnecting computing networks consisting of billions of connected computers to provide a variety of information and communication facilities using standardized communication protocols. At the core of the Internet is a computer network which is an interconnected computing device that can exchange data and share resources with each other. Each device in the network uses a system of rules, called communications protocols, to transmit information over physical or wireless technologies.

15.4.2.4 Cloud Computing

A computer cloud or cloud computing is difficult to define because it is a concept—a computing concept in which a global network of servers, hosting databases and other software applications, is accessed over the internet by users. So a computer cloud is not a physical entity, but instead is a vast network of remote servers around the globe which are hooked together and operate as a single ecosystem. As a computing concept, therefore, it is an on-demand availability of computer system resources, especially data storage and computing power, without direct active management by the user.

Cloud Computing Service Models

The computing cloud has several service models including the following:

Infrastructure as a Service (IaaS) The process of providing the customer with the ability and capability to manage and control, via a Web-based virtual server instance API, system resources such as starting, stopping, accessing, and configuring the virtual servers, operating systems, applications, storage, processing, and other fundamental computing resources is referred to as Infrastructure as a Service

(IaaS). In doing all these, however, the consumer does not have access nor control of the underlying physical cloud infrastructure.

Platform as a Service (PaaS) This is a set of software and product development tools hosted on the provider's infrastructure and accessible to the customer via a Web-based virtual server instance API. Through this instance, the customer can create applications on the provider's platform over the Internet. Accessing the platform via the Web-based virtual instance API protects the resources because the customer cannot manage or control the underlying physical cloud infrastructure including network, servers, operating systems, or storage.

Software as a Service (SaaS) Ever since the beginning of computing software, over the years, the key issue that has driven software development has been the issue of the cost of software. Trying to control the cost of software has resulted in software going through several models. The first model was the home-developed software where software users developed their own software based on their needs and they owned everything and were responsible for updates and management of it. The second model, the traditional software model, was based on packaged software where the customer acquired a more general-purpose software from the provider with a license held by the provider and the provider being responsible for the updates while the customer being responsible for its management. However, sometimes, software producers provide additional support services, the so-called premium support, usually for additional fees. Model three was the open-source model led by a free software movement starting around the late 1980s. By the late 1980s, free software turned into open source with the creation of the Open Source Initiative (OSI). Under the name "open-source" philosophy, some for-profit "free software" started to change the model from a purely free software to some form of payment for support of the updates of the software. The open-source software model transformed the cost of software remarkably. Model four consisted of software outsourcing. The outsourcing model was in response to the escalating cost of software associated with software management. The component of software management in the overall cost of software was slowly surpassing all the costs of other components of software including licensing and updates. In model four, however, software is still licensed from the software company on a perpetual basis; support fees are still paid; however, the software producer takes on the responsibility of the management of that software.

Software model five is Software as a Service (SaaS). Under this model, there is a different way of purchasing software. Under SaaS, there is the elimination of the upfront license fee. All software applications are retained by the provider, and the customer has access to all applications of choice from the provider via various client devices through either a thin client interface, such as a Web browser, a Web portal, or a virtual server instance API. The cloud user's responsibilities and actual activities in the use of and operations of the requested cloud services are limited to user-specific application configuration settings, leaving the management and control of the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities to the cloud provider.

Cloud Computing Deployment Models

There are three cloud deployment models which are actually cloud types. These are the public, the private, and the hybrid models.

Public Clouds The public clouds provide access to computing resources for the general public over the Internet allowing customers to self-provision resources typically via a Web service interface on a pay-as-you-go basis. One of the benefits of public clouds is to offer large pools of scalable resources on a temporary basis without the need for capital investment in infrastructure by the user.

Private Cloud Unlike public clouds, private clouds give users immediate access to computing resources hosted within an organization's infrastructure and premises. Users, who are usually in some form of a relationship with the cloud owner, choose and scale collections of resources drawn from the private cloud, typically via Web service interface, just as with a public cloud. Also the private cloud is deployed within and uses the organization's existing resources and is always behind the organization's firewall subject to the organization's physical, electronic, and procedural security measures. In this case, therefore, private clouds offer a higher degree of security.

Hybrid Cloud A hybrid cloud combines the computing resources of both the public and private clouds.

All the computing concepts and technologies discussed starting with the computer network, the Internet, the Internet of things (IOT), virtualization, virtual reality, augmented reality and cloud computing form the core basis of the architecture of the metaverse.

15.5 Actualization of Metaverse

The actualization of the metaverse is the realization of the working of the metaverse. A fully functioning metaverse is glued together by two concepts that make it work: telepresence and localization.

15.5.1 The Concept of Telepresence

Telepresence is a sensation we get from being somewhere without physically being there. It is created by the use of virtual reality technology. Virtual reality technology creates this sensation in an individual by enabling the person to perform actions in a distant or virtual location as if they are physically present in that location. Metaverse platform technology enables and supports telepresence. There are many examples of telepresence including telemedicine where physicians carry out surgeries in remote locations, satellite communication and manipulation, remotely working in the deep sea, and working in dangerous chemical environments.

15.5.2 Localization

Localization is a process of making something local in character or restricting it to a particular place. This process is important and is part of every metaverse platform. Activities performed on metaverse platforms must undergo location for effective use and implementation.

Discussion Questions

- Are there any differences between telepresence and sleepwalk? Discuss.
- Is there any difference between telepresence and hypnosis? Discuss.

15.6 Benefits of Metaverse

We have defined Metaverse as a convergence of the physical and virtual spaces enabled by both virtual and augmented reality via, at least for the time being, smart wearable gadgets. These new spaces have the potential, currently more hyped than not, to create new opportunities in almost every aspect of society from education, health, sports, entertainment and more so in commerce. It will create a new economy of mostly digital goods, at least initially, and later on physical ones, entertainment and services that will revolutionize ecommerce as we know it today. Metaverse platform-based and more platform neutral digital currencies and payment systems are developing fast. This is driving companies and brands to compete for opportunities in Metaverse to sell virtual and physical goods or services. A new model of e-commerce will emerge.

15.6.1 E-Commerce in Metaverse

As we have discussed above, services and all activities in Metaverse take place on platforms. Services, activities and technologies performed on each platform are different. Even services and activities of two users of the same platform may vary. This says something about the Metaverse. It is just like it parallels services and activities of the physical world. For example users on Metaverse platforms represented by virtual avatars can socialize with others and carry out activities like shopping, trading and other.

15.6.2 Metaverse Billing and Payment Systems

In order for e-commerce to thrive, a strong billing and payment system must be developed in Metaverse to support e-commerce. Although Metaverse is a fusion of both the physical and virtual worlds, the e-commerce billing and payment systems

must embody both these works. In both these worlds, payments can be made in various ways, but for Metaverse, the most common way, at least for now and probably in the near future, is via *blockchain and digital currencies*, which users can use to purchase goods and services from merchants on and between metaverse platforms.

15.6.2.1 Blockchain

Blockchain, Bitcoin, Dogabit and other unpronounceable terms are dominating conversation these days across the globe. What do they mean and why are they the talk of the block? What are these words and concepts that have come to excite not only a few people, age groups, a country but all generations, all countries and it cuts across the globe, rich and poor countries alike, mean? As we will shortly see, all these terms refer to the same technology—blockchain. **Blockchain**, was originally **block chain**, a continuously growing list of records, called *blocks*, which are linked data blocks and secured using cryptography [2]. In elementary computer science these linked data blocks are called linked lists. In its traditional way, the chain of blocks consists of connected blocks where each block has a hash pointer as a link to a previous block, a timestamp and of course the transaction data. Because each block data hashes into a *unique number*, blockchains are inherently resistant to modification of the data. Any transaction, between any two or more parties, that uses the data in any of these blocks, according to [3] is in the open, distributed among more than one server, in a peer-to-peer network of servers, uses efficiently verifiable protocols and is permanent. Once recorded, the data in any given block cannot be altered retroactively without the alteration of all subsequent blocks, which requires collusion of the network majority [4]. This last feature of decentralized consensus, possessed by the blockchain technology gives it broad appeal in a variety of areas requiring recording of events such as medical records, all types of records management activities, such as identity management, transaction processing, documenting provenance, food traceability or voting [4].

Working of Blockchain Technology

The working of the blockchain technology can be briefly described in the following steps:

- A transaction (financial or otherwise) involving several parties is initiated by one of the parties at the party's digital device. This transaction may be a payment (i.e. Credit card transaction, or a bank transfer, bill payment or others). The device sends the transaction on to its server in the peer-to-peer network.
- The initiator's server broadcasts the transaction to all servers in the peer-to-peer network. All servers get the same transaction from the broadcasting server.
- Upon receipt, each server runs its cryptographic verifying algorithm and generates a proof of the transaction—an invoice.
- A new block to house the transaction data and invoice is generated and this data is stored in the new block. The block is chained to the existing chain of blocks

to create a new, longer (by one block) permanent and unalterable blockchain and stored at each server in a peer-to-peer and any request to any block data on the chain must be approved by all servers.

The blockchain was started in 2008 by an anonymous person, a Japanese called Satoshi Nakamoto (there have been pictures of a person many believe is him) and applied on a *cryptocurrency* called the **bitcoin**, in 2009. So by this transaction, bitcoin automatically became the first digital currency to solve the *double spending problem* (counterfeiting or coping), without the need of a trusted authority or central server. Since then, there has been many other kinds of digital currencies.

A *cryptocurrency* is a digital asset designed to work as a medium of exchange (as we use physical cash) but with the help of cryptography to secure its transactions, to control the creation of additional units, and to verify the transfer of assets.

This fast developing technology has raised lots of questions, curiosity, anxiety and excitement. There is a growing list of questions that need answers. For example, is this technology here to stay or will it evaporate in a year or so like many other technologies before it? What advantages does it have over all other technologies before it? How will it advance the way we work and overall human conditions? The answers to these questions and a lot more lie in understanding what it is and what are its promises.

According to Mearian [5], blockchain technology is the most disruptive technology in decades. He believes the technology has the potential to eliminate huge amounts of record-keeping, save money and disrupt IT in ways not seen since the internet arrived. Maybe he is right. But no one knows at this time. Everyone is guessing where the technology is heading. Millions are betting their livelihoods on it. Probably the biggest reason why millions, knowingly or otherwise, are betting on it is that for the first time in human business transactions, there is a technology that eliminates the need to trust the other party when you transact with them. The core and hard backbone reason for this is that the whole peer-to-peer network upon which the distributed technology is transacting is public and safeguarding every record on every server making the issue of hacking a blockchain database is pointless. For this very reason, technology may be a business transaction and global payment systems game changer for everyone rich or poor, highly educated or illiterate, living in a first world or a third world. Everyone is treated the same. What this means is that the problem of individual and private record keeping where vital and essential records can be vandalized, stolen, or just lost cannot happen anymore. Yes, with all these benefits, the technology offers a dirty, cheap and inexpensive way to transact, store and transfer records. Thus, anyone, anywhere, with an Internet connection can use. In a way, it looks like a business and all other transactions equalizer technology.

15.7 Social and Ethical Concerns of the Metaverse

The metaverse presents a litany of both ethical and social issues. Among these are issues brought about by telepresence that support users of the metaverse space autonomy, transparency, trustworthiness, privacy, economic inequalities, accessibility, freedom of creative expression and many others. Let us look at some here.

15.7.1 Metaverse Immersion and Autonomy

The physical absence of an actor in the theater of operation requires a high degree of the entity representing the actor. In the case of the metaverse, the actor is represented by the avatar or a robot. This is a growing field of interest in artificial intelligence and robotics. The more useful and efficient a robot is, the more autonomous it must be. So in the metaverse, the avatar must be immersive and have as much autonomy as possible in order to be effective in whatever activities it is supposed to be involved in. However, the ideal of autonomy has tremendous consequences both positive and negative. The positive ones result in the autonomous entity being able to create conditions which result in more positive, pleasant and unexpected expected outcomes. The negative effects, however, can result in lots of undesired outcomes as a result of unexpected actions that cause serious harm and may be normally both immoral and criminal. In telepresence, there is a delicate balance between increasing and decreasing autonomy.

15.7.2 Metaverse and Transparency

An entity is transparent if it allows light to pass through so that objects behind it can be distinctly seen. An action is transparent if the thoughts, feelings, or motives behind it are easily perceived. So transparency is operating in such a way that it is easy for others to see what actions are performed. Transparency implies openness, communication, and accountability. With this background, can metaverse support transparency? By this, we mean can we achieve a transparent telepresence? A good telepresence strategy puts the human factors first. Human factors as we have seen in most of the previous chapters of this book involve human thoughts, feelings, or motives. These are very difficult to manage.

15.7.3 Metaverse and Trustworthiness

Trustworthiness is the ability to be relied on as honest or truthful. It is a very important virtue for being a member of a community. It is one of the three basic

common ethical principles for telepresence, others being autonomy and transparency. These three create and maintain a healthy relationship between and among users of the metaverse.

15.7.4 Metaverse and Privacy

Privacy has many varying definitions depending on the circumstances in which it is defined. The most generic definition of privacy is as a state or condition of someone or a group being free from being observed or disturbed by other people. A broader and more detailed definition of privacy was given by Krupp et al. [6] as first as an *informational privacy*, over personal information, includes (a) invasion, (b) collection, (c) processing, and (d) dissemination; second as *physical privacy*, over personal space or territory, includes (a) personal space, (b) territoriality, and (c) modesty; and third as *psychological privacy*, over thoughts and values, includes (a) interrogation and (b) psychological distance; and finally as *social privacy*, over interactions with others and influence from them, includes (a) association, (b) crowding/isolation, (c) surveillance, (d) solitude, (e) intimacy, (f) anonymity, and (g) reserve. With a high degree of autonomy any one of these privacy divisions becomes of great concern.

15.7.5 Metaverse and Access Inequality

Because the metaverse platform is projected to open up far more opportunities than the current Internet has done, for people and businesses, to design, build and distribute both physical and virtual products, there are concerns and fears that it might undo the narrowing of the digital divide the Internet has brought. The issue is to create metaverse platforms that afford equal access to all community users, businesses, creators and developers. How do developers and creators of these platforms and both physical and virtual items for the metaverse grant equal access to all around the world? It is likely that access to Metaverse platforms, technologies and services may follow the current access we have for the Internet with all its challenges and problems, based on income, ethnicity, geography and education, as we saw in Chap. 9.

15.7.6 Interoperability in Metaverse

In its simplest definition, interoperability refers to the ability of different systems, devices, applications or products, in a given space or domain, to connect and communicate in a coordinated way, without effort from the end users. In a functioning metaverse with different platforms, services, avatars, and activities, absence of interoperability means a dead metaverse. Metaverse interoperability is essential because it allows users to interact with a larger number of entities, which can lead

to lots of new opportunities and experiences. Thus, creating a more cohesive and connected metaverse, which can make it more enjoyable and rewarding to explore for all users. According to Alex Corntege [7], interoperability of the metaverse is essential for the following reasons:

- It allows users to interact with a larger number of people, which can lead to new opportunities and experiences.
- It helps to create a more cohesive and connected metaverse, which can make it more enjoyable and rewarding to explore.
- It can help to reduce the barriers to entry for new users, as they can start using any platform that they like without having to worry about being locked out of other platforms.

15.7.7 Metaverse and Freedom of Expression

According to the United Nations Human Rights Commission (UNHCR), freedom of expression is a fundamental human right, enshrined in article 19 of the Universal Declaration of Human Rights [7]. A number of freedoms fall under the category of freedom of expression. Freedom of the media and freedom of opinion and expression, including conscientious objection. In the metaverse, human and probably robotic avatars are expected to be used more widely. However, these avatars, depending on what they are, may not have the necessary components and gestures to express themselves. Thus, they will have limited effective ways to express themselves.

Exercises

- Precisely define the concept of metaverse.
- Discuss the benefits of metaverse to society.
- Discuss the likelihood of a single payment system in the metaverse.
- Immersion is behaviorally addictive. Discuss how this may lead to problems for the metaverse users.
- Do you think the metaverse will create or destroy jobs?
- There is a debate on the possible existence of the digital divide in the metaverse. What is your opinion?
- Is there anything like equal access to the metaverse? Is it achievable in space?
- The concept of metaverse may not be realized as projected. Is this a true statement? Why or why not?
- Will the number of metaverse platforms increase social and ethical challenges in metaverse?
- Discuss how the following technologies and architectures help to build the metaverse architecture: computer network, virtualization, augmented reality, virtual reality, internet of things and cloud computing.

- What, if any, are the ethical implications of telepresence?
 - What are the benefits, if any, of immersion in the metaverse?
 - What are the social and ethical problems resulting from a rush of companies to the metaverse?
 - Should metaverse users sacrifice their privacy because of the pleasure of the new metaverse experience?
-

References

1. Wikipedia. Metaverse. [https://en.wikipedia.org/wiki/Metaverse#:~:text=Neal%20Stephen%2C%20Snow%20Crash%20\(1992,metaphor%20of%20the%20real%20world](https://en.wikipedia.org/wiki/Metaverse#:~:text=Neal%20Stephen%2C%20Snow%20Crash%20(1992,metaphor%20of%20the%20real%20world)
2. M. Centeio, *How Blockchain Technology Can Benefit Developing Countries*. <https://cvhustle.com/blockchain-technology/.5>. The Internet Society Blockchain 5. Special Internet Group (ISOC-BSIG). Blockchain Technology: Opportunities for Africa. <https://www.isoc-bsig.org/blockchain-technology-opportunities-africa/>
3. Harvard Business Review. https://en.wikipedia.org/wiki/Harvard_Business_Review
4. Wikipedia. <https://en.wikipedia.org/wiki/Blockchain>
5. L. Mearianm, *What is Blockchain? The Is the Most Disruptive Technology in Decades*. <https://www.computerworld.com/article/3191077/security/what-is-blockchain-the-most-disruptive-tech-in-decades.html>
6. M.M. Krupp, M. Rueben, C.M. Grimm, W.D. Smar, A focus group study of privacy concerns about telepresence robots, in. *26th IEEE International Symposium on Robot and Human Interactive*. Communication (RO-MAN), Lisbon, Portugal, Aug 28–Sept 1, 2017
7. A. Corntege, *What is Metaverse Interoperability, and Why Is it Important?* <https://www.linkedin.com/pulse/what-metaverse-interoperability-why-important-%C3%A0lex-corretg%C3%A9>

Further Reading

To understand the changing nature and challenges of the metaverse, please read more on the following topics:

8. Future of Metaverse
9. NFTs in Metaverse
10. Challenges and Requirements of Metaverse
11. How to create a New Identity on Metaverse
12. Metaverse Avatars

Index

A

Access, 33, 57, 65, 68–70, 73, 74, 77, 79, 82, 84, 96, 111, 115, 117–119, 121–126, 128, 129, 131, 135, 148, 155, 184, 187, 189, 190, 193–195, 199, 200, 208, 215, 218, 229, 231–233, 235, 237–239, 244, 245, 247–251, 261, 271, 289–291, 293, 299, 302–304, 314, 316, 319, 320, 325, 326
Access Control Lists (ACL), 250, 251
Accountability, 39, 42, 45, 58, 218, 270, 283, 301, 324
Acquired Immune Deficiency Syndrome (AIDS), 194, 269, 294
Activism, 185, 186
African, 120, 122
Alienation, 138, 145, 147
Altruism, 22
Anonymity, 31, 55, 65–67, 74, 84, 190, 199, 200, 202–204, 217, 236, 241, 242, 325
Anonymous
 remailers, 55, 66
Appeals, 45, 50, 110, 219, 322
ARPANET, 182, 235
Artificial Intelligence (AI), 24, 25, 212–215, 220, 222, 223, 269, 273, 275–278, 280–284, 324
Artificial Life Interactive Institute Video Environment (ALIVE), 2, 3, 266
Asimov, Isaac, 220, 221, 269
Assessment, 48, 140, 159, 160, 177, 188, 247, 248, 250, 283
Association of Computing Machinery (ACM), 19, 29, 162
AT&T Bell Labs, 134
Atheist, 2
Atomic Energy of Canada Ltd (AECL), 162

Authentication, 68, 70–73, 249, 251, 255, 270, 302, 303

Autonomous agents, 212, 215, 221, 266
Autonomy, 38, 40, 42, 43, 75, 76, 78, 137, 214, 221, 222, 236, 255, 256, 303, 324, 325

B

BAE Automated Systems Ltd, 154
Bhopal, India, 167
Biometric algorithm, 251
Biometrics, 251
Breach of contract, 172, 174
BTT, 98
Buffer overflow, 186, 246
Buyers' right, 168
Byzantines, 5

C

Canadian, 185, 186
Carpal Tunnel Syndrome (CTS), 147
Censoring cyberspace, 218, 219
Censorship, 218, 219, 223
Center for Information Enhanced Medicine (CIEMED), 266
Central Intelligence Agency (CIA), 217
Chain-of-custody, 186
Challenger, 56, 162, 163, 166, 167
Cipher, 248
Ciphertext, 70, 248
Civil liberties, 15, 65, 80, 81, 130
Clone, human, 14, 19
Close Circuit Television (CCTV), 69
CNN, 185
Code
 community, 41, 42
 conduct, ethics, 1, 3, 9, 29, 30, 40, 50–52, 221, 269

- executable, 89, 108, 110
 institutional, 41, 42, 51, 52
 local, 51, 52
 moral, 5, 6, 8, 9, 11, 13, 15, 16, 39, 42
 object, 89, 108, 110
 penal, 12
 personal, 41, 51
 professional, 1, 19, 34, 35, 38, 40–42, 46, 48, 50–52
 source, 89, 108
- Commandments**
 Native American, 6, 7, 16
 Ten, 6, 7, 13, 16
 Unix Users' Group, 6, 7
- Commitment**, 29, 39, 42, 43, 58, 146, 270
- Competence**, 39, 47, 48, 50, 174
- Compliance**, 29, 47, 48, 140, 143, 174, 233, 271
- Computer**
 attacks, 182, 184, 185, 188, 189, 191, 196
 bridges, 226
 crimes, 79, 155, 179–182, 184, 186, 189, 191, 193, 195, 196
 gateways, 226
 hubs, 226
 network, 32, 46, 79, 89, 182–184, 215, 216, 226–229, 253, 258, 271, 293, 295, 314, 317, 318, 320, 326
 protocols, 303, 318
 routers, 226
 software, 89, 90, 107, 108, 159, 169–171, 174, 175, 181, 182, 244
 surrogates, 227
 technology, 24, 26, 31, 33, 34, 52, 59, 76, 81, 83, 91, 92, 116, 117, 132, 133, 140, 146, 155, 161, 181, 214, 244, 251
- Conceptual model**, 33, 34, 293
- Confidentiality**, 48, 55, 68, 70, 83, 194, 249
- Conscience**, 2, 8, 9, 16, 33, 55, 58
- Consequentialism**, 21, 22, 25
- Contract**, breach of, 172, 174
- Conventions**
 Berne, 92, 106
 Paris, 92
- Copyrights**, 87, 88, 91–94, 97, 101–109, 112, 279
- Crime Information Center**, Security Agency (NSA), 80, 185
- Cryptography**, 70, 72, 248, 249, 253, 322, 323
- Cumulative Trauma Disorder (CTD)**, 147
- Cyber**
 café, 130
 crime, 102, 189, 196, 244, 253
 culture, 220
- sleuth, 84
Cybercommunity, 218
Cyberspace, attacks, 190
Cybersquatting, 88
Cyberzens, 217
- D**
- Decrypt**, 71
- Denial of service**, 245, 246
- Deontology**, 21, 22, 25
- Deskilling**, 138, 139
- Development testing**, 153, 264
- Digital**
 divide, 115, 117–119, 122, 123, 126, 128–131, 148, 325, 326
 Millennium Copyright Act, 83, 106, 219
 signature, 71, 72
- Digitalization**, 65
- Digital Millennium Copyright Act (DMCA)**, 83, 106, 219
- Disclaimer**, 172, 174, 177
- Discrimination**, 54, 57–59, 81, 141, 280
- Distributed Denial of Service (DDoS)**, 185, 186, 192, 246, 303
- DNA**, 14, 278
- Doctrine**
 fair use, 104
 first sale, 104
- DoubleClick**, 78
- Duration of**
 copyright, 94
 patent, 95
 trademark, 99
 trade secrets, 97
- E**
- E-attacks**, 186, 189, 192, 193, 246
- Eavesdropper**, 70
- Eavesdropping**, 79, 237, 245, 272
- E-crimes**, 196
- Egoism**, 22, 25
- Electronic Data Gathering, Analysis, and Retrieval (EDGAR)**, 245
- Electronic Funds Transfer (EFT)**, 82
- Electronic monitoring**, 140–145, 148
- Electronic office**, 132, 140
- Email**, 202, 235, 283
- Emotivism**, 21
- Empire**
 French, 39
 Roman, 5, 12
- Encryption**

- asymmetric, 70–72
symmetric, 70–72, 249
- England, 39, 92
- Ergonomics, 146–148
- Espionage, 102, 106, 181, 187, 188, 244
- Ethical
- argument, 19, 27, 28, 59
 - computing, 29
 - decisions, 26, 27, 30, 37, 53, 59
 - dilemmas, 30, 32
 - implications, 149, 180, 213, 215, 223, 270, 272, 287, 311, 327
 - issues, 14, 34, 37, 54, 84, 138, 217, 226, 240, 255, 273, 280
 - muddles, 28, 32, 305, 306
 - principles, 24, 26, 29, 30, 269, 325
 - theories, 21–23, 25, 26, 28, 34, 53, 269
- Ethics
- decision function, 25
 - functional definition, 24, 25, 30, 34
 - traditional definition, 26
- Ethnicity, 118, 122, 188, 325
- Extortion, 187, 201
- F**
- False Claims Act, 56, 59
- Federal Bureau of Investigation (FBI), 38, 81, 84, 180, 183–185, 217
- Federal Trade Commission (FTC), 78
- File Transfer Protocol (FTP), 297
- Firewalls
- filters, 69
 - gateways, 271
 - proxy servers, 70
 - stateful, 70
- Flowchart, 89, 90, 110
- Floyd, 220
- Forensics, 312
- Formal
- education, 38, 39, 46
 - review, 176
- 414-Club, 182
- Fourth Amendment, 74, 78
- France, 92, 120, 206
- Fraud, 56, 81, 99, 181, 183, 195, 244, 279
- G**
- G8, 130
- General Agreement on Tariffs and Trade (GATT), 92
- Globalization, 16, 77, 79, 112, 135, 139, 140
- Gramm–Leach–Briley Financial Privacy Act, 75, 77, 83
- Greek, 20–23, 26, 66, 248
- Guilt, 8, 9, 16, 39, 51, 53, 59
- H**
- Hackers, 64, 79, 112, 155, 159, 181–185, 187, 189, 190, 196, 288
- Hactivism, 185, 186
- Harassment, 54, 57–59, 199–205, 219, 241–243
- Hash function, 70–72
- Hazard, 148, 155, 156, 158–160, 280, 281
- Hedonism, 21, 23
- Hotline, 55
- Human
- development index (HDI), 118
 - nature, 21, 23, 25, 34
 - value, 73, 78, 213, 236
 - ware, 118, 126, 128, 129, 161, 162, 178
- I**
- Identity
- personal, 74–76, 84, 91, 99, 100, 105, 236
 - theft, 84, 99, 100, 180, 201, 237, 242
- IEEE, 50, 301
- Information Communication Technology (ICT)
- gathering, 76
 - matching, 78–80
 - security, 68, 70
- Infringement
- contributory, 103
 - copyright, 103
 - direct, 103
 - inducement, 103
 - patent, 103
 - trademark, 104
- Integrity, 30, 39, 42, 43, 54, 58, 68, 70–72, 82, 187, 191, 245, 249
- Intel, 169
- Intellectual property rights, 88, 91–93, 101, 105–107
- International Business Machines (IBM), 135, 137, 169, 235, 257, 258, 277
- Internet, 5, 14, 32, 33, 46, 62, 65–67, 70, 73, 75, 77, 78, 81, 83, 84, 88, 100, 112, 116–119, 121–127, 129–131, 135, 148, 149, 152, 156, 157, 162, 180–187, 189, 190, 192, 199–201, 204–206, 209, 215, 218, 219, 229, 233–238, 241–244, 246, 248, 277,

- 278, 282, 287–289, 291–295, 299, 301–306, 312–314, 316–320, 323, 325, 326
- Intrusion, 74, 79, 181, 182, 188–191, 236, 237, 244, 252
- Intrusion detection, 160, 183, 192, 252, 272
- IP-spoofing, 186, 246
- K**
- Key
- private, 70, 71
 - public, 70, 71, 251
 - secret, 70
- L**
- Laboratorio Integrado de Sistemas
- Inteligentes y Technologías de la Información en Tráfico (LISITT), 266
- Law
- conventional, 11, 13, 14, 16
 - natural, 10–13, 16
- Leadership Mobilization Strategy (LMS), 130, 278, 279
- Lenham Act, 97
- Liabilities
- strict, 172, 174, 175, 177
- Licensing, 46–48, 58, 59, 130, 131, 319
- Local Area Network (LAN), 184, 186, 229, 246, 262
- Logic map, 89
- M**
- Malfunction, 162, 181
- Malice, 157
- Malpractice, 171, 172, 174, 175
- Management styles, 130, 139, 140
- Marks
- arbitrary, 98, 99
 - descriptive, 99
 - general, 99
 - MD5, 98
 - suggestive, 98
- Massachusetts Institute of Technology (MIT), 258, 266
- Message digest, 71, 72
- MI6, 217
- Miniaturization, 65, 107, 133, 199, 200
- Minsky, Marvin, 215, 220
- Misrepresentation, 102, 174, 175
- Monitoring, 117, 136, 139, 140, 142–145, 149, 159, 160, 191, 201, 216, 217, 241, 249, 252, 263, 272, 303
- Moral
- codes, 5, 6, 8, 9, 11, 13, 15, 16, 39, 41
 - decay, 192
 - decision making, 4, 5
 - implications, 8, 9, 53, 58
 - standards, 1, 8, 9, 53, 58
 - theories, 4, 5, 25
- Morality, 1–5, 8–10, 13–16, 21, 35, 39, 43, 83
- Muddles, 52
- N**
- National Aeronautical and Space Administration (NASA), 162, 163, 167
- National Security Agency (NSA), 217
- Negligence, 91, 171, 172, 174, 175
- Netscape, 169, 236
- Network
- forensics, 160, 191
 - integrity, 181, 244
 - packet, 181, 244
 - switched, 244
- Network News Transfer Protocol (NNTP), 245
- Nonreply, 81
- Nonrepudiation, 72, 249
- North American Free Trade Agreement (NAFTA), 92
- Novelty, 95
- NSFNET, 235
- NSLookup, 246
- O**
- Occupational Overuse Syndrome (OOS), 147
- OECD, 128
- Office
- electronic, 132, 140
 - home, 134–138
 - virtual, 133
- Operating System (OS), 7, 164–167, 181, 182, 185, 194, 216, 235, 245, 246, 250, 257–261, 263, 265, 271, 317–319
- Opt out, 76, 77
- Organizational channels, 54, 56
- Organizational leadership, 29
- Ownership, 88, 94, 101, 102, 104, 105, 108, 112, 121, 122, 239

P

Parkinson's disease, 2
Passive reconnaissance, 245
Passwords, 7, 69, 73, 194, 195, 240, 245, 250, 251, 288, 303
Patents, 87, 88, 91, 93–97, 101–107, 109, 111, 112, 171
Penal code, 12
Penetration, 68, 120, 121, 182, 184–186, 188, 191, 196, 245–247, 249, 250
Pentium, 169
Personal
 digital assistant (PDA), 132
 identity, 74–76, 84, 91, 99, 100, 105, 236
Phreaking, 182
Pinger, 246
Plaintext, 70, 248
Pretext calling, 100
Prisoner's dilemma, 52
Privacy, 31, 61, 62, 64, 65, 73–84, 116, 141, 142, 145, 149, 162, 181, 192, 193, 212, 217–219, 226, 232, 233, 236–240, 244, 249, 252, 256, 280, 281, 283, 284, 293, 301, 302, 306, 307, 324, 325, 327

Procedures

 appeals, 48, 50
 hearing, 48, 49
 sanctions, 48

Product

 development, 110
 replacement, 169
 responsibility, 44
 service, 44
 updates, 169

Professionalism, 30, 38, 39, 42, 46, 51, 54, 58

Pseudo

 address, 66
 identity, 66

Q

Quality, 3, 5, 44, 50, 56, 76, 98, 99, 128, 136, 137, 139, 142, 143, 145, 152, 153, 156, 163, 173, 176, 229, 231, 236, 280, 295

Quality of Service (QoS), 44, 153, 156, 157

Qui-tam, 56

R

Random Access Memory (RAM), 152, 303

Read Only Memory (ROM), 90, 152

Relativism, 15, 21, 23

Reliability, 151, 153, 154, 162, 176, 177, 229, 289, 291, 301

Repetitive Strain Injuries (RSI), 146–148

Reskilling, 139

Responsibility

 consequential, 44, 45
 product, 44
 service, 44

Risk, 19, 54–56, 115, 136, 151, 154–156, 158–162, 172, 174, 175, 177, 178, 195, 242, 247, 248, 260, 270, 276, 280–283, 301, 306, 307, 311

Romans, 5, 10, 12

Rootkit, 185

Roslin Institute, 14

Rule

 bronze, 6, 15
 golden, 5
 iron, 6, 15
 nepotism, 6
 silver, 6
 tin, 6

S

Safety, 12, 44, 50, 51, 54, 95, 151, 153, 155, 156, 158, 161, 162, 167, 171, 176–178, 242, 268

Sanctions, 10, 48, 49, 57, 269

Script, 3, 4, 16, 93, 184

Security, 13, 61, 65, 67–70, 72, 73, 80–82, 84, 100, 116, 130, 140, 142, 143, 153–156, 159, 160, 162, 177, 180, 182–184, 188, 189, 191–193, 195, 217, 219, 226, 231, 233, 237, 247–252, 261, 262, 268, 270–273, 281, 283, 284, 287, 288, 291–293, 301–303, 307, 320

Self-esteem, 102, 145, 147, 204

Self-regulation, 81, 247, 252

Sequence numbers, 186, 247

Simulador Reactivo de Conducción de Automobiles (SIRCA), 266, 267

Smurf, 186, 246

Sniffers, 70

Snooping, 63, 177

Social Security Numbers (SSN), 75, 77, 80, 100

Software

 canned, 91, 171
 complexity, 158
 customized, 171
 developer, 87, 108, 110, 112, 153, 157, 171, 175, 176, 256

failures, 153, 155, 157, 158
 hybrid, 171
 producer, 89, 152, 157, 171, 175–177, 319
 quality, 153, 156, 176, 177
 reliability, 154, 161, 177
 risk, 158, 159, 161
 safety, 155, 161, 176, 178
 Software Quality Function Development (SQFD), 156
 Standards, 5, 8, 13, 15, 30, 32, 39–41, 45, 46, 48, 53, 82, 132, 143, 153, 154, 156, 173, 176, 185, 194, 226, 240, 272, 295, 301, 317, 318
 Stateful, 70
 Statute of Queen Anne, 92
 Stem cell, 2
 Stress, 57, 137, 141, 147, 148, 162, 206, 207, 282
 Subscript, 3, 4
 Surveillance
 electronic, 116, 188, 217
 SYN flooding, 246

T

Telecenter, 129, 130
 Telecommunication and Information Administration (NTIA), 117
 Telecommuting, 134–138, 148, 149
 Telemarketer, 76
 Temptations, 31, 32, 52, 78, 281
 Terrorism, 181, 187, 218, 219, 244
 Theory
 X, 139, 140
 Y, 139, 140
 Therac-25, 162
 Third-party beneficiary, 77, 173, 174
 3D, 215, 266, 267, 315, 316
 Tort, 141, 169, 172, 174, 175
 Total Quality Management (TQM), 156
 Trademark, 7, 88, 91, 93, 97–99, 101, 102, 104–106, 110, 112
 Trade-Related Aspects of Intellectual Property Rights (TRIPPS), 92, 106
 Trade secrets, 87, 91, 96, 97, 101, 102, 105, 106, 108, 110, 112
 Transmission Control Protocol/Internet Protocol (TCP/IP), 290–292, 295, 297–300, 317, 318
 Turing, Alan, 215, 279

Turing machine, 222
 2D, 215

U

U.N. Declaration of Human Rights
 Human Development Report, 78, 326
 UN Educational Scientific and Cultural Organization (UNESCO), 92
 Universal Copyright Convention (UCC), 92, 106, 173
 Unix, 7, 235, 259
 Upper Limb Disorder (ULD), 147
 Usenet, 186, 235, 245
 U.S. Federal Trade Commission (FTC), 78
 Utilitarian, 53, 93, 116, 241, 269
 Utilitarianism
 act, 22, 25
 rule, 22

V

Vacuum, 33
 Validation and Verification (V&V), 153, 154, 176
 Vendetta, 6, 55, 56, 157, 186
 Virtual Reality (VR), 214, 255, 256, 265–270, 272, 273, 312–314, 316, 320, 326
 Virus, 46, 180, 182–184, 186, 190, 191, 252
 Visual Display Unit (VDU), 147
 Vulnerability
 scanning, 216, 252

W

Walk through, 46
 Warranties
 express, 173, 177
 implied, 172, 173, 177
 Whistle blowing, 37, 54–59, 67
 Wide Area Network (WAN), 184, 229, 230
 Workplace
 and privacy, 140, 145
 and surveillance, 140
 World Intellectual Property Organization (WIPO), 92, 106
 World Trade Organization (WTO), 92, 106
 World Wide Web (WWW), 34, 88, 215, 235, 238