IS1108 Digital Ethics and Data Privacy

AY 24/25 Sem 2 — github/omgeta

1. Professional Ethics

Professional ethics are the principles which govern behaviour of people in their work environment such as:

i. Relationships and responsibilities with other stakeholders
ii. Guidelines related to the actions and decisions of individuals who create and use computer systems

They are often codified in professional code of ethics:

i. ACM Code of Ethics
ii. Software Engineering Code of Ethics and Professional Practice

Typical principles include:

i. Honesty
ii. Trustworthiness
iii. Loyalty
iv. Respect for Others
v. Adherance to Law
vi. Do Good, Avoid Harm
vii. Accountability

Personal responsibility involves taking accountability for one's actions, decisions and thoughts such as:

i. Admitting when a program is faulty
ii. Declining a job that one isn't qualified to do
iii. Speaking out when others do wrong

Posner's Principle: negative information should be in the public domain.

Principle of Double Effect: if an action has a good and bad outcome, it may be permissible to perform that action if only the good outcome is intended and reasonable steps are taken to avoid/minimise the bad outcome.

2. AI Ethics

AI Ethics are guidelines which advise on the design and outcomes of AI-enabled systems such as:

i. Privacy and Surveillance
ii. Bias and Discrimination
iii. Role in Human Judgement

Three Generations of AI:

i. Handmade: Knowledge, algorithms and compute power (e.g. expert systems)
ii. Statistical: Data, algorithms and compute power (e.g. ML)
iii. Generative: Knowledge, data, algorithms, and compute power (e.g. self-driving cars, LLM)

Four Types of AI:

i. Reactive: Unable to learn (e.g. spam filters)
ii. Limited Memory: Uses memory to study past data
iii. Theory of Mind: Can socialize and empathise
iv. Self-Aware: Aware of themselves, internal state and of other's emotions

Five Pillars of Trustworthy AI:

i. Fairness: bias-free and no unethical discrimination
ii. Explainable: decisions understandable by humans
iii. Robustness/Security: no unathorised misuse
iv. Transparency: use known to all stakeholders
v. Privacy: privacy and data rights for people

Singapore's Two Principles and Four Pillars:

i. AI Decisions "explainable, transparent and fair"
ii. AI Systems should be human-centric
iii. Internal Governance, Operations Management, Human-Centricity, Stakeholder Communications

Singapore's AI Strategy:

i. Global Hub for AI Solutions
ii. Govern and Manage AI Impact
iii. Generate Economic Value & Improve Lives

3. Automation and Autonomous Systems

Automation is the ability to perform tasks with deterministic results without AI technologies.

Autonomy is the ability of AI-based systems to perform tasks with minimal human intervention.

Assess autonomous systems with impact-autonomy matrix.

Ethical Concerns:

i. Safety: Veering away from programmed learning in unstructured environments
ii. Privacy: Extent to which surveillance functions infringe on personal privacy of others.
iii. Security: Detailing the purpose and types of data collected, and stipulaated level to access to data.
iv. Liability: Right allocation of responsibilities and compensation risks in event of harm.
v. Effects to Incumbent Workforce: Disruptive unemployment consequences.
vi. Autonomy and Independence: Ability of users to exhibit self-determination and assert preferences.
vii. Human Interaction: Compromise of social interactions and human touch.
viii. Objectification and Infantilization: Undermining dignity by subjecting users to control of robots.
ix. Deception: Counterfeiting social engagement and misleading users.
x. Social Justice: Ensuring equity in access to and distribution of autonomous systems.

Framework for Automation:

i. Technical: ensure safety and assurance in function.
ii. Professional Responsibility: encourage best safety practices, especially when there is constant change.
iii. Regulation: enforce relevant regulations.
iv. Oversight: ensure transition from development to deployment is just.
v. Public Acceptance: do not over-intrude into lives.
vi. Ethics: ensure transparency in use and tracking.

4. Personal Data Protection Act (PDPA)

Singapore's PDPA is technology-neutral (electronic and non-electronic data), complaint-based and reasonable. It must be applied to SOP Operations, ICT Control and Policy Management.

Individual: natural person, whether living or dead

Organisation: individual, company, association or body of persons whether or not recognised locally (e.g. businesses, NGOs, agents, freelancers)

Data Intermediary*: organisation which processes personal data on behalf of another organisation (data controller) but does not include any of their employees.

Personal Data/ Personally Identifiable Information (PII) is data whether true or not, about an individual who can be identified with only the data, or with other information the organisation is likely to have. DP Provisions:

i. Notification: Collection must be disclosed.
ii. Consent: Individuals must consent to collection, usage and sharing of data.
iii. Purpose Limitation: Collection must be reasonable and only for consented reason.
iv. Accuracy: Data must be accurate and complete.
v. Protection: Data must be secured*.
vi. Retention Limitation: Data only kept as long as necessary for purpose, or for legal reasons*.
vii. Transfer Limitation: Data overseas follows PDPA.
viii. Access and Correction: Requests need response within 30 days, with reasonable access fees. Applied to data used or disclosed in past year.
ix. Accountability: Data Protection Officer (DPO) is appointed to implement policies.
x. Data Breach Notification: Notify PDPC and individuals in 30 days (3 days if > 500 affected)*.
xi. Data Portability: data extracted must be portable to other organisations

Business Contact Information (BCI) is not provided solely for personal use (e.g. business contacts for business purposes). BCI is not protected under PDPA.

5. Digital Ethics by Design

Digital ethics is the implication of technological on the social, political and moral space of business.

Design is the organisation approach to responsible use of technology by applying digital ethics.

Five Concepts of Digital Ethics:

i. Ethical Behavior: base off ethical frameworks
ii. Transparency and Privacy
iii. Technology Advancements
iv. Data Driven Insights
v. Security and Compliance

Need for Digital Ethics:

i. Human rights and safety at risk from data theft
ii. Lack of trust in public institutions using technology
iii. Lack of trust in the legal protections
iv. Legal unreset
v. Lack of community buy-in

Good Practices:

i. Manage data with integrity
ii. Incorporate ethics in decision-making
iii. Observe relevant government-wide arrangements for trustworthy data access, sharing and use
iv. Monitor data inputs and adopt risk-based approach to automation
v. Be specific about purpose of software, especially concerning personal data and human rights
vi. Publish open source code
vii. Be accountable and proactive in risk management

Framework for Digital Ethics:

i. R&D Process integrating ethical design decisions
ii. Define stakeholders, application domain and training for ethical considerations
iii. Identify ethical requirements and values
iv. Elaborate design principles governing R&D
v. Form methodology to perform ethical assessments

6. Human-Computer Interaction (HCI)

HCI is the designing, implementing, and evaluating of interactive interfaces used by people and computers to create products easy to learn, effective and enjoyable.

Automation should be developed to:

i. Balance human control and computer automation
ii. Be reliable, safe and trustworthy
iii. Reduce unexplainable errors
iv. Reduce over-reliance on technology

Brain Control Interface (BCI) create communication between human brain and a computer, with ethical issues:

i. Accountability: humans and machines integration
ii. Privacy: of brain signals
iii. Misinterpretation: of signals to incorrect results
iv. Security: risk of hacking

7. Equity, Accessibility and Inclusion

Digital divide is the growing gap, primarily digital skills and knowledge, between those with differing access to computers and the internet. Digital equity is when individuals have equal opportunity to use digital tools.

Six Pillars of Digital Inclusion:

i. Affordable broadband Internet
ii. Internet-enabled devices
iii. Access to digital literacy training
iv. Quality technical support
v. Applications and online content

Promoting Digital Equity:

i. Provide digital skills training to low-income
ii. Improve online accessibility of social services
iii. Empower low-income households

Risks from Increased Digital Access:

i. Cyberbullying
ii. Supervision vs Surveillance
iii. Opaque decision-making

8. Computing for Social Good

Technology can be applied to social good in:

i. Sustainability
ii. Transport
iii. Assistive Products (should be made as accessible as possible)
iv. Poverty, Hunger, Clean water

Corporate Social Responsibility (CSR) refers to companies taking action to give back to society. It encourages businesses to conduct in an ethical manner and work towards a positive impact through sustainable growth.

Environmental, Social and Corporate Governance (ESG) refers to the central factors in measuring sustainability and social impact of an investment into a company.

Environmental factors:

i. Natural resource use
ii. Carbon emissions and energy efficiency
iii. Pollution or waste

Social factors:

i. Workforce
ii. Human rights
iii. Diversity
iv. Supply chain

Governance factors:

i. Board Independence
ii. Boad diversity
iii. Shareholder rights
iv. Corporate ethics
v. Management compensationi

9. Intellectual Property (IP) Rights

IP is any unique asset created and used for business, such as art, designs and website content.

Three Types of IP Rights:

i. Copyright: Arts, music, performance, software
ii. Patent: Inventions, technology
iii. Trademark: Marks of a registered business such as term, symnbol, slogan

Respecting IP Rights:

i. Always credit the owner/author of the original work
ii. Do not gain profit by free-riding others' works
iii. Do not plagiarise works, with or without permission

Requirements for Copyright Protection in Singapore:

i. Protected Work: "authorial work" (e.g. art)
ii. Singaporean: author Singaporean or residing in Singapore, or first published in Singapore
iii. Expressed Tangibly: work in some material form
iv. Original

Copyright in Computer Science:

i. Preparatory materials (e.g. flowcharts, schemas)
ii. Computer programs (e.g. including source code)
iii. Databases and other works

Copyright in IOT & AI:

i. Architecture of Project
ii. Proprietary Regression Model (AI)
iii. Complex code logic
iv. Specific compilations or arrangement of code

Copyright in Data Analytics:

i. Database structure
ii. Data contents and arragement of contents (only if there are multiple ways)
iii. Reports/works regenerated to display reports

Software involves both the process and the product, which makes it harder to determine if it can or cannot be patented.

i. UK: Inventions are patentable, where inventions are any manner of new manufacture or new method or process
ii. Singapore: Patent can be applied if the invention is new (responsibility of creator to check), involves an inventive step (proven through documentation), and has industrial application (shown capability)