Information Technology and Moral Values

A Course for the Digital Age

Brendan Shea, PhD

Rochester Community and Technical College

Welcome: Technology in Your Daily Life

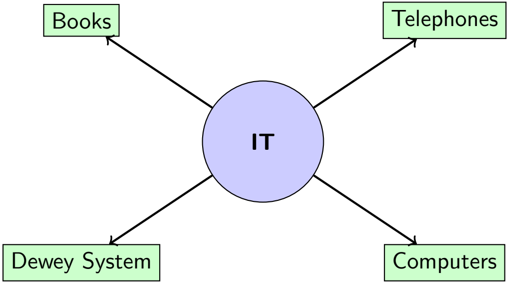
- You likely checked your smartphone within minutes of waking up this morning, joining billions of people worldwide in this daily ritual.
- Every digital action you take—from streaming music to posting on social media—leaves a trail of data that could be recorded, analyzed, and stored indefinitely.
- Information technologies have become so integrated into our lives that we rarely stop to consider their profound ethical implications.
- This course will examine how digital technologies challenge our traditional moral values and require us to rethink fundamental concepts like privacy, ownership, and autonomy.

Course Goal

To develop your ability to recognize, analyze, and respond to the ethical challenges posed by information technologies in your personal and professional lives.

What Is Information Technology? Defining the Digital World

- **Information technology (IT)** refers to any system used to record, transmit, organize, or synthesize information—whether ancient or modern.
- Examples of IT include books (which record information), telephones (which transmit information), and library cataloging systems like the Dewey Decimal System (which organize information).
- Modern computers are special because they function as **universal machines**—they can be programmed to perform all three functions and emulate any form of information technology.
- Information technologies have existed for thousands of years, from writing systems to printing presses, though digital technologies have accelerated their impact dramatically.



Why Ethics Matter in the Digital Age

- Information technologies have moved from being optional tools to essential infrastructure that shapes how we work, communicate, and understand the world.
- The rapid pace of technological change often outstrips our ability to develop appropriate ethical guidelines and legal frameworks to govern these technologies.
- Digital information is fundamentally different from physical resources because it can be copied infinitely without loss, raising unique questions about ownership and distribution.
- The choices we make about designing, deploying, and using information technologies will profoundly affect human flourishing for generations to come.

The Central Question

Who should have the power to collect, control, and use the massive amounts of data generated by modern information technologies—and what moral principles should guide these decisions?

What Is Information? From Strings on Fingers to Smartphones

- **Information** can be understood as any useful data, instructions, or meaningful message content that helps shape our thoughts and actions.
- The word "information" literally means to "give form to" or to shape one's understanding of the world.
- A simple example: tying a string around your finger serves as information technology—the string symbolizes a complex proposition like "buy groceries before you come home."
- The string itself is not the information; it merely represents or symbolizes the information, which means it must be correctly interpreted to be useful.

Technology	**Symbol/Medium**	**Information Represented**
String on finger	Physical knot	"Remember to buy groceries"
Written letter	Ink on paper	Personal message to recipient
Book	Printed text	Knowledge or story
Smartphone alarm	Digital signal	"Wake up at 7:00am"

Syntactical vs. Semantic Information: Understanding the Difference (Shannon)

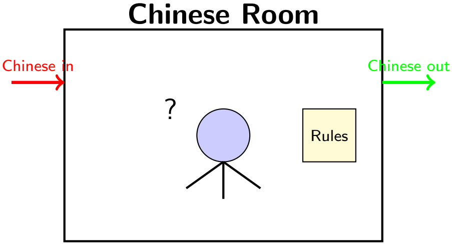
- **Syntax** refers to the structure and ordering of symbols in information, while **semantics** refers to the meaning or content that those symbols represent.
- Claude Shannon, working at Bell Labs in the 1940s, developed a mathematical theory of communication that focused on syntax—how to encode, transmit, and decode symbols while minimizing noise and errors.
- Shannon's approach treats **information** as *meaningfully ordered sets of symbols that can be transmitted as signals*, but it deliberately leaves out questions about what those symbols actually mean.

Why This Matters

Modern computers are extraordinarily good at processing syntactical information (manipulating symbols according to rules), but whether they can truly understand the semantic meaning of that information remains a fundamental philosophical question.

The Chinese Room Argument: Can Computers Really Think? (Searle)

- Philosopher John Searle (1980) proposed a thought experiment to argue that computers cannot truly understand information, even if they appear to process it intelligently.
- Imagine a person who doesn't speak Chinese locked in a room with a rulebook for manipulating Chinese characters—they receive Chinese stories and questions, then follow rules to output appropriate Chinese answers.
- The person in the room would produce correct answers without understanding anything about the stories, knowing only how to manipulate symbols according to syntactical rules.
- Searle argues that computers work the same way: they skillfully manipulate syntax without any genuine understanding of semantics or meaning.



Three Functions of Information Technology: Record, Communicate, Organize

- All information technologies can be categorized by their function(s): **recording** (storing), **communicating** (transmitting), or **organizing/synthesizing** information.
- Many technologies perform multiple functions—for example, a smartphone can record photos, communicate via text messages, and organize your calendar appointments.
- The computer is unique as a **universal machine**: it can be programmed to emulate any form of information technology and perform all three functions simultaneously.
- Understanding these three functions helps us recognize the different ethical challenges that arise from each type of information processing.

Function	Classic Example	Digital Example
Record	Book, photograph	Cloud storage, database
Communicate	Telephone, letter	Email, text message
Organize	Dewey Decimal System	Search engine, algorithm

The Information Technologies Behind This Lecture

- This lecture itself demonstrates the layered nature of information technology—each layer building on previous innovations to enable new forms of communication and learning.
- **Written language** (ancient technology) allows us to record and transmit ideas across time; **mathematical notation** enables precise symbolic expression.
- **Von Neumann computer architecture** (1945) provides the computational foundation; **ASCII** (1963) standardizes character encoding for digital text.
- **LaTeX** (1980s) typesets these slides with mathematical precision; **VSCode with GitHub Copilot** (2021) provides AI-assisted coding; **GitHub** stores and versions the files; **Zoom** transmits the lecture; **YouTube** archives it for future viewing.

Technology	**Year**	**Function**
Written language	∼3200 BCE	Record information
Von Neumann architecture	1945	Compute and process
ASCII	1963	Encode characters digitally
LaTeX	1984	Organize and present
GitHub	2008	Store and collaborate
Zoom/YouTube	2011/2005	Communicate and archive
Copilot and LLMs	2021	AI-assisted “generative” creation of LaTeX code

What Is Ethics? Understanding Moral Philosophy

- **Ethics** (or moral philosophy) is the systematic study of right and wrong, good and bad, and what we ought to do in various situations.
- Ethics asks fundamental questions: What makes an action right or wrong? What kind of person should I strive to be? How should we organize society fairly?
- Ethical thinking differs from simply following laws, customs, or personal preferences—it requires reasoned justification and consideration of principles that apply broadly.
- In this course, we'll use ethical frameworks to analyze the moral dimensions of information technologies and develop reasoned positions on controversial issues.

Ethics vs. Other Forms of Guidance

- **Law**: What we are legally required or prohibited from doing
- **Custom/Etiquette**: What social conventions expect us to do
- **Religion**: What religious authorities or texts command
- **Ethics**: What we ought to do based on moral reasoning and principles

Normative Ethics vs Descriptive Ethics

- **Descriptive ethics** studies how people actually behave and what moral beliefs they hold, without making judgments about whether those beliefs or behaviors are right or wrong.
- **Normative ethics** prescribes how people ought to behave and which moral principles should guide our actions, providing standards for evaluating conduct.
- In this course, we focus primarily on normative ethics—developing reasoned arguments about how information technologies should be designed and used, not merely describing current practices.

Descriptive Ethics	**Normative Ethics**
"Most people share their location data with apps."	"Should people share their location data with apps?"
"62% of teens say social media has no effect on their self-esteem."	"What responsibilities do social media companies have to protect teen wellbeing?"
"Companies routinely collect browsing history."	"Is it morally permissible for companies to collect browsing history without explicit consent?"

From Common Morality to Professional Ethics

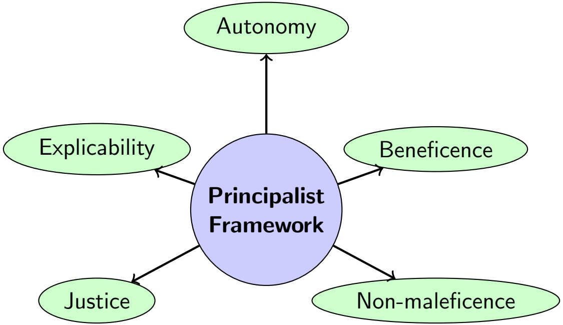
- **Common morality** consists of basic moral norms shared across cultures—prohibitions against harm, theft, and deception, and requirements to help others and keep promises.
- However, information technologies create **novel moral problems** that common morality alone cannot resolve—issues our ancestors never faced and traditional moral wisdom doesn't directly address.
- **Applied ethics** and **professional ethics** extend common morality by developing specialized principles and reasoning for new contexts like healthcare, business, and information technology.

Why "Fourth Grade Morality" Isn't Enough

Common morality tells us not to harm others and to respect their property—but it doesn't tell us whether collecting browsing data violates privacy, how to distribute algorithmic benefits justly, or who's responsible when an AI system causes harm. We need specialized ethical frameworks to extend basic moral principles to these novel technological situations.

The Principalist Framework: Five Key Principles

- The **principalist framework**, developed by Beauchamp and Childress for biomedical ethics and extended by Floridi for information ethics, provides five core principles for ethical analysis.
- The five principles are: **autonomy** (self-determination), **beneficence** (doing good), **non-maleficence** (avoiding harm), **justice** (fairness), and **explicability** (transparency and accountability).
- When principles conflict, we must use careful moral reasoning to determine which principle should take priority in a specific context.



Principle of Autonomy: Respecting Self-Determination

- **Autonomy** is the capacity for self-determination—the ability to make informed decisions about one's own life based on one's values, preferences, and reasoning.
- Respecting autonomy means treating people as ends in themselves, not merely as means to others' goals, and ensuring they have meaningful control over decisions that affect them.
- In information technology contexts, autonomy requires informed consent for data collection, genuine choices about technology use, and freedom from manipulation or coercion.

Autonomy in IT: Key Questions

- Do users understand what data is being collected and how it will be used?
- Can users meaningfully choose whether to use a technology or share their data?
- Are users being manipulated through algorithmic recommendation systems?
- Do people have genuine control over their digital identities and information?

Principle of Beneficence: Promoting Good and Wellbeing

- **Beneficence** is the positive obligation to promote good, contribute to welfare, and help others flourish—going beyond merely avoiding harm.
- In information technology, beneficence requires designing systems that genuinely improve human lives, enhance capabilities, and contribute to individual and social wellbeing.
- This principle challenges us to ask whether technologies serve human needs or merely generate profit, and whether they empower or diminish human agency and connection.

Beneficence in IT: Examples

Positive: Health tracking apps that help people manage chronic conditions; educational platforms that increase access to learning; assistive technologies that enable people with disabilities.

Questions to ask: Does this technology make people's lives genuinely better? Does it enhance human capabilities or create new dependencies? Who benefits most from this technology?

Principle of Non-maleficence: Above All, Do No Harm

- **Non-maleficence** embodies the ancient medical principle "first, do no harm"—we have an obligation to avoid causing harm to others through our actions.
- This principle requires us to anticipate potential harms from information technologies, including privacy violations, security breaches, psychological damage, and social harms.
- In IT contexts, harm can be direct (a data breach exposing personal information) or indirect (algorithms that perpetuate discrimination or social media features that contribute to mental health problems).

The Challenge of Unforeseen Harms

Many information technologies cause harms that designers never anticipated—social media addiction, algorithmic bias in hiring, or cyberbullying platforms. Non-maleficence requires taking reasonable steps to identify and prevent potential harms before deployment.

Principle of Justice: Fairness and Equitable Distribution

- **Justice** requires fair distribution of benefits and burdens, equal treatment of equals, and special concern for vulnerable or disadvantaged groups.
- In information technology, justice addresses the digital divide, algorithmic bias, and whether technologies reinforce or reduce existing social inequalities.
- Justice questions include: Who has access to beneficial technologies? Whose interests do algorithms serve? Who bears the risks and costs of technological development?

Justice Concern	**IT Example**
Access inequality	Broadband access gaps
Algorithmic bias	Facial recognition errors for minorities
Exploitation	Data harvesting from vulnerable users
Fair treatment	Discriminatory hiring algorithms

Principle of Explicability: Transparency and Accountability

- **Explicability**, added by philosopher Luciano Floridi, requires that information systems be transparent, understandable, and that those who design and deploy them be accountable for their effects.
- This principle demands that we can explain how systems work, justify their design choices, and hold someone responsible when things go wrong.
- Explicability is especially challenging with complex AI systems and algorithms that even their creators may not fully understand—the "black box" problem.

Why Explicability Matters

Without explicability, we cannot meaningfully exercise autonomy (we can't consent to what we don't understand), assess beneficence and non-maleficence (we can't evaluate unknown harms), or ensure justice (we can't detect hidden biases). Explicability enables all other principles.

Applying Principles to Information Recording

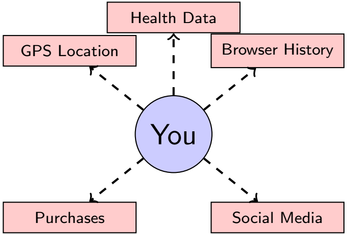
- We've established our ethical framework—now we apply it to specific challenges posed by information technologies.
- The first major function of IT is **recording** (storing) information—creating databases, files, and digital records of human activities.
- Recording information raises novel ethical problems that common morality doesn't directly address: What counts as consent for data collection? Who owns recorded information about you? When is surveillance justified?
- We'll analyze these issues using our principalist framework—asking how recording technologies affect autonomy, beneficence, non-maleficence, justice, and explicability.

The Ethical Stakes

Modern recording technologies can capture and store virtually unlimited information about every person—creating unprecedented power imbalances between those who control data and those who generate it. This section explores how to navigate these novel moral challenges.

Digital Footprints: What Data Do You Leave Behind?

- Every digital action leaves a **data trail**—GPS coordinates, browser histories, purchases, and social media interactions.
- Modern technologies automatically collect biometric data, financial transactions, and search queries—creating comprehensive digital profiles.
- This collection often happens **behind the scenes** without users' awareness or meaningful consent.



Key Issue

Aggregating seemingly innocent data points can reveal intimate details you never explicitly shared.

Privacy in the Digital Age: Who Owns Your Information?

- **Privacy** traditionally meant the ability to control access to one's personal space and information, but information technology has fundamentally challenged this concept.
- Philosopher Helen Nissenbaum argues that where physical barriers once protected privacy, technology now makes personal information "available at the click of a button or for a few dollars."
- The central ethical question is: who has the right to control, sell, or use information about you—especially when third parties collect it without your explicit knowledge or meaningful consent?

The Death of Privacy?

In 1999, Sun Microsystems CEO Scott McNealy famously declared: "You have zero privacy anyway. Get over it." While privacy may be harder to protect today, many philosophers argue that **personal autonomy and intimacy** still require us to defend privacy rights—even in the digital age. Without privacy protection, our autonomy is undermined.

The Cloud: Convenience vs. Control

- **Cloud storage** means your data is stored on remote servers owned by third parties (like Google, Apple, or Microsoft) rather than on devices you control.
- The cloud offers enormous convenience—access from anywhere, automatic backups, seamless syncing—but creates new vulnerabilities and trust dependencies.
- You're placing tremendous **trust** in third parties to protect your information, maintain access, and not misuse your data.

Benefits	**Risks**
Access from any device	Loss of direct control
Automatic backups	Company could lose/delete data
No local storage needed	Privacy breaches possible
Easy sharing with others	Uncertain legal protections
Professional security teams	Data mining by companies

Biometric Data: Your Body as Information

- **Biometric data** refers to measurements of your physical characteristics—fingerprints, facial recognition, heart rate, blood pressure, sleep patterns, and DNA sequences.
- Technologies like smartwatches and fitness trackers collect continuous biometric data streams, often shared with third-party applications and companies.
- Unlike passwords, you cannot change your fingerprints or facial structure—making biometric data breaches particularly serious and permanent violations.

Principalist Analysis

Autonomy: Do users consent meaningfully? **Non-maleficence**: What harms result from breaches? **Justice**: Who has access to beneficial health tracking vs. whose data is exploited? **Explicability**: Can users understand how their biometric data is analyzed?

Case Study 1: The Fitness Tracker Dilemma

The Scenario

Ada, a high school senior, receives a fitness tracker as a gift. The device monitors her heart rate, sleep patterns, steps, and location throughout the day. The companion app offers personalized health insights and connects to a social network where users can share achievements and compete with friends.

Key Facts:

- The privacy policy (12 pages of legal text) states that health data may be shared with "trusted third-party partners" for "service improvement."
- Ada's health insurance company offers a 15% discount to customers who share fitness tracker data proving they exercise regularly.
- The app uses dark patterns—pre-checked boxes for data sharing and complex opt-out procedures buried in settings.
- Ada's school requires fitness trackers for PE class participation and shares aggregate student data with the district.

Case Study 1: Review Questions

Using the Principalist Framework, analyze:

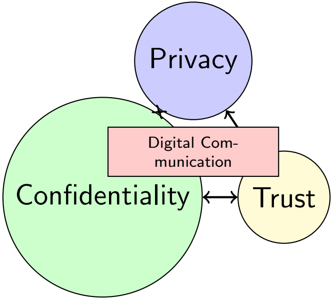
1. **Autonomy**: Does Ada have meaningful autonomy over her health data? Consider the length and complexity of the privacy policy, the school requirement, and the insurance incentive. What would genuine informed consent look like here?
2. **Justice**: Is it fair that students must use fitness trackers for PE class? What about students who can't afford devices or have health conditions they wish to keep private? Does the insurance discount create unjust inequalities?
3. **Non-maleficence**: What potential harms could result from sharing Ada's biometric and location data with third parties? Consider both immediate risks (data breaches) and long-term risks (discrimination, surveillance).
4. **Explicability**: Who is accountable if Ada's data is misused? Can Ada meaningfully understand how her data will be analyzed and used by "trusted partners"?

Privacy, Confidentiality, and Trust Online

- Information technology forces us to rethink privacy concepts based on print technologies like letters and newspapers.
- **Confidentiality** refers to keeping specific information private between particular parties—like doctor-patient or lawyer-client communications.
- **Trust** becomes crucial when we must rely on third parties to protect our information and use it appropriately.

Key Tension

Digital information is easily shared and altered, stretching traditional privacy protections to their breaking point.



Under pressure from:

- Easy copying
- Third-party access
- Data aggregation
- Unclear ownership

Who Controls Your Personal Information?

- One fundamental question of information ethics: who has the final say over whether information about you is communicated, sold, or used?
- Philosopher Alan Westin argued that **control of access to personal information** is the key to privacy in the digital age—without control, privacy rights are meaningless.
- Traditional institutions (governments, banks, healthcare systems, corporations) have long held power through their control of stored information about individuals.

Your Data	**Who Controls It?**	**Ethical Issue**
Medical records	Hospital, insurance, govt	Privacy vs. public health
Financial history	Banks, credit agencies	Access vs. exploitation
Browsing history	ISPs, websites, advertisers	Tracking vs. consent
Email content	Email provider, govt	Confidentiality vs. security
Social media posts	Platform companies	Ownership vs. terms of service

Digital Communication: Email, Texts, and Social Media Ethics

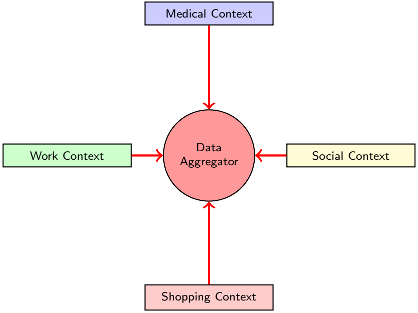
- Digital communication technologies (email, text messages, social media) have transformed how we share information, but they create new ethical challenges regarding privacy, permanence, and context collapse.
- **Permanence**: Digital messages can be stored indefinitely, forwarded without permission, and retrieved years later—unlike ephemeral spoken conversations.
- **Context collapse**: Social media collapses multiple audiences (family, friends, employers, strangers) into one, making it nearly impossible to share appropriately for different contexts.
- The ease of copying and forwarding digital messages means confidential information can spread instantly beyond its intended recipients, violating trust and confidentiality norms.

Ethical Question

If you send someone a private message, do they have the right to screenshot and share it? Does your expectation of privacy depend on the platform used? What obligations do we have regarding digital communications we receive?

Contextual Integrity: Privacy in Different Social Contexts

- Philosopher Helen Nissenbaum's theory of **contextual integrity** argues that privacy isn't about secrecy—it's about appropriate information flow for specific contexts.
- Information appropriate in one context (medical exam) becomes a privacy violation in another (dinner conversation).
- Social media and data aggregation violate contextual integrity by combining information from different contexts in unexpected ways.



Privacy violation through context collapse

Key Insight

Privacy violations occur when information flows in ways that violate context-specific

Search Engines and Filter Bubbles: How Algorithms Shape What You See

- **Search engines** organize the vast amount of information on the internet using algorithms that rank and filter results based on relevance, popularity, and personalization.
- **Filter bubbles** occur when algorithms selectively present information based on your past behavior, creating a personalized information environment that reinforces existing beliefs.
- These systems raise serious concerns about **autonomy**—can you make truly informed decisions if algorithms control what information reaches you?
- Questions of **justice** emerge when different users receive systematically different information based on their demographics, location, or wealth.

Algorithm Feature	**Ethical Concern**
Personalized results	Filter bubble effect
Ranking by popularity	Amplifies majority views
Ad-based revenue model	Prioritizes engagement over truth
Opaque decision-making	Violates explicability

Data Mining and the Power of Information Control

- **Data mining** refers to analyzing large datasets to discover patterns, correlations, and insights that weren't visible in smaller samples or individual cases.
- Companies and governments use data mining to predict behavior, target advertising, assess risk, and make decisions about individuals—often without those individuals' knowledge.
- The control of information represents a form of **political and economic power**—those who can collect, analyze, and act on data have significant advantages over those who cannot.

The Power Asymmetry

Organizations increasingly possess detailed profiles of billions of individuals, while those individuals often lack access to their own data or understanding of how it's used. This creates a fundamental **justice** problem: unequal power, unequal access, and unequal vulnerability to harm. As philosopher Richard Mason argued, these technologies require us to rethink the social contract itself.

Case Study 2: Bias in Google Search and ChatGPT

The Scenario

Marcus, a high school junior, is researching "successful CEOs" for a business class project. He uses Google Search and ChatGPT to find information and generate ideas.

Key Facts:

- Google's image search for "successful CEO" returns predominantly images of white men in business suits, with very few women or people of color in the top 100 results.
- When Marcus asks ChatGPT to "describe a typical successful CEO," the AI generates a description matching traditional stereotypes: "confident," "assertive," "male," "experienced in finance."
- ChatGPT was trained on internet text data that reflects historical biases and underrepresentation of women and minorities in leadership positions.
- Neither Google nor ChatGPT explicitly explains how their algorithms work or why certain results appear first—the decision-making process is largely opaque.
- Marcus doesn't realize these results are algorithmically curated and might be biased—he treats them as objective information.

Case Study 2: Review Questions

Using the Principalist Framework, analyze:

1. **Justice**: How do biased search results and AI outputs perpetuate existing inequalities? Who is harmed when algorithms reinforce stereotypes about leadership, success, or capability? What obligations do tech companies have to ensure fair representation?
2. **Autonomy**: Can Marcus make truly autonomous decisions about his understanding of business leadership if algorithms control what information he sees? Does the lack of awareness about algorithmic curation undermine informed decision-making?
3. **Explicability**: Should Google and OpenAI be required to explain how their algorithms rank results and generate text? Who is accountable when biased outputs cause harm? Can users meaningfully understand these systems?
4. **Beneficence/Non-maleficence**: What are the broader social harms of algorithmic bias in information systems? How might this affect career aspirations, hiring decisions, and social perceptions?

Social Media Ethics: Friendship, Identity, and Authenticity

- Social media platforms fundamentally reshape how we form and maintain relationships, raising questions about the nature of **friendship**, **identity**, and **authenticity** in digital spaces.
- Philosopher Shannon Vallor argues that social media can support genuine friendships and human flourishing—or it can cultivate vice and shallow connections, depending on how platforms are designed and used.
- The architecture of social media (likes, followers, metrics) can encourage performative behavior and superficial engagement rather than deep, meaningful relationships.
- **Virtue ethics** asks: what kind of character traits do social media platforms cultivate? Do they encourage honesty, compassion, and wisdom—or vanity, envy, and distraction?

The Authenticity Problem

Social media allows us to curate idealized versions of ourselves, but this can create pressure to maintain false personas and prevent genuine self-disclosure. How can we cultivate authentic relationships when the medium encourages performance over authenticity?

Online Gaming Communities: Real Ethics in Virtual Worlds

- **Massively multiplayer online games** (MMOs) create virtual worlds where millions of people interact, form communities, and develop alternative lives.
- These virtual worlds raise real ethical questions: Do our moral obligations extend to virtual spaces? Can you wrong someone in a game?
- Players develop genuine relationships, economies, and social structures within games—suggesting that virtual interactions have real moral significance.
- Griefing, harassment, and discrimination in games cause real psychological harm to real people, even if the environment is virtual.

Real-World Case: RuneScape Theft (2007)

In the Netherlands, two teenagers physically assaulted a 13-year-old boy, beating and threatening him with a knife until he logged into the game RuneScape and transferred his virtual amulet and mask to them. In 2012, the Dutch Supreme Court upheld their theft conviction, ruling that virtual objects have "intrinsic value" because of "the time and energy invested" in obtaining them. The offenders received 144 hours of community service.

Digital Identities: Who Are You Online?

- Our **digital identities**—the personas we create and maintain online—raise profound questions about authenticity, consistency, and self-presentation across different platforms and contexts.
- Social media encourages us to curate idealized versions of ourselves, potentially creating a gap between our online persona and our actual lived experience.
- Multiple digital identities (professional LinkedIn, casual Instagram, anonymous Reddit) fragment the self, raising questions about which version is "really you."
- The permanence of digital identities creates new challenges—online posts from years ago can resurface and affect current opportunities, relationships, and reputation.

The Authenticity Question

Philosopher Shannon Vallor argues we must ask: Do our online identities reflect genuine self-expression and help us flourish as human beings? Or do they trap us in performative cycles that prevent authentic connection and self-knowledge? The architectures of social media platforms shape which aspects of ourselves we present—and therefore who we become.

The Technological Transparency Paradox: Freedom vs. Control

- Digital information creates a fundamental paradox: "Information wants to be free" (easily copied and shared) versus "Information is valuable and should be controlled" (property requiring protection).
- Unlike physical objects, digital information is **nonexclusory**—we can all possess the same digital file without preventing others from having it, since copying doesn't eliminate the original.
- This creates moral tensions: Should digital music, books, and movies be freely shared? Do creators have rights to control copies? Is piracy theft or liberation?
- The paradox extends to personal data: we want our information protected (privacy) but also want free access to others' information (transparency).

The CIA Triad in Tension

Information security balances three goals:

- **Confidentiality**: Keeping information private and controlled
- **Integrity**: Ensuring information remains accurate and unaltered
- **Availability**: Making information accessible when needed

These goals often conflict—maximizing availability threatens confidentiality; strict confidentiality limits availability.

Malware and Cybersecurity: Protecting Digital Systems

- **Malware** (malicious software) includes viruses, worms, trojans, ransomware, and spyware designed to damage systems, steal data, or enable unauthorized access.
- **Cybersecurity** involves protecting information systems from malware and unauthorized access, but it raises ethical tensions between security and privacy, convenience and protection.
- The ethics of cybersecurity involves questions of **non-maleficence** (preventing harm from attacks) and **justice** (who bears the costs and risks of security measures).

Malware Type	Function	Primary Harm
Virus	Replicates and spreads	System damage, data loss
Ransomware	Encrypts files for ransom	Extortion, business disruption
Spyware	Monitors user activity	Privacy violation
Trojan	Disguised malicious code	Unauthorized access

Cyberwarfare and the Ethics of Digital Weapons

- **Cyberwarfare** refers to using information technologies as weapons to attack enemy information systems, infrastructure, and military capabilities.
- These attacks can disable power grids, disrupt communications, destroy data, and potentially cause physical harm—all without traditional kinetic weapons.
- Cyberweapons raise novel ethical questions: Can they distinguish between military and civilian targets? Who is responsible for collateral damage? Do traditional rules of warfare apply?

The Stuxnet Case

Stuxnet (discovered 2010) was a sophisticated computer worm that targeted Iranian nuclear facilities, damaging centrifuges and setting back their nuclear program. Some argue it prevented military action that would have caused greater civilian casualties. Others worry it opened the door to continuous low-level cyber conflicts that avoid accountability and escalate slowly. The attack raised questions about **proportionality**, **discrimination**, and whether cyber operations can ever satisfy just war principles.

Artificial Intelligence: Can Machines Make Moral Decisions?

- **Artificial Intelligence (AI)** refers to information technologies that exhibit aspects of human-level intelligence, problem-solving, and decision-making capabilities.
- Alan Turing proposed in 1950 that machines would eventually "think" in ways indistinguishable from humans—the famous **Turing Test** where a computer tries to convince a human it's not a machine.
- Modern AI systems (like ChatGPT, facial recognition, and recommendation algorithms) raise urgent ethical questions about bias, accountability, and whether machines can or should make moral judgments.
- Key debate: Does consciousness matter for moral agency? Some argue we can only have moral duties toward conscious beings; others argue AI systems can have moral rights and duties regardless of consciousness.

The Moral Machine Problem

If AI systems make decisions affecting human welfare (medical diagnoses, loan approvals, criminal sentencing), who is morally responsible when they err? The programmer? The company? The AI itself? Can we program machines to behave ethically, or does morality require something machines cannot possess?

Robotics in Society: From Healthcare to Warfare

- **Robotics** combines information technology with physical machines that interact directly with humans and the environment.
- Applications range from medical robots assisting surgery to military drones conducting warfare—each raising distinct ethical challenges.
- **Military robotics** is especially controversial: autonomous weapons can make life-or-death decisions with little human intervention.
- Some ethicists argue robots could reduce casualties by following international humanitarian law more consistently than humans under stress; others worry autonomous weapons make war too easy to declare and violate human dignity by removing human judgment from killing decisions.

Principalist Analysis of Military Robotics

Autonomy: Do civilians consent to being potential targets of autonomous systems? **Non-maleficence**: Can robots reliably distinguish combatants from civilians and avoid excessive harm? **Justice**: Who bears responsibility when autonomous weapons cause unjust casualties? **Explicability**: Can military robots explain their targeting decisions?

The Future of Humanity: Technology and Human Enhancement

- **Transhumanism** is the view that we should use technology to enhance human capabilities—extending lifespan, augmenting intelligence, and transcending biological limitations.
- **Moore’s Law** observed that computing power doubles approximately every 18 months—futurist Ray Kurzweil predicts this acceleration could lead to ”the Singularity,” where technology fundamentally transforms human nature.
- Critics argue transhumanism reflects pseudoscientific beliefs and fear of death, devalues the natural human body, and would deepen inequality between enhanced and unenhanced humans.

Enhancement Type	**Ethical Concerns**
Cognitive enhancement	Fairness in education/work, authenticity, coercion
Life extension	Justice in resource distribution, overpopulation
Physical augmentation	Access inequality, redefining human dignity

Case Study 3: Being Young in the Age of LLMs

The Scenario

Maria is a high school senior trying to prepare for her future in a world rapidly being transformed by large language models (LLMs) like ChatGPT. She's watching how AI is changing writing, creative work, coding, research, and even social interaction—and wondering what this means for her generation.

Key Facts:

- Maria used to love writing stories and essays, but now wonders if developing writing skills matters when AI can generate high-quality text instantly. She's uncertain which skills will be valuable in 10 years.
- Her older cousin, a junior copywriter, just lost his job because the marketing firm replaced their writing team with two "AI coordinators" who edit AI-generated content.
- Maria notices her younger brother (age 12) primarily interacts with AI chatbots for entertainment, homework help, and even emotional support—he seems to prefer AI conversations to human ones.
- She observes that AI systems reflect and amplify existing biases, spread misinformation convincingly, and that her generation will inherit both the benefits and harms of these systems they didn't create.

Case Study 3: Review Questions

Using the Principalist Framework, analyze:

1. **Autonomy & Human Flourishing**: How do LLMs affect young people's development of critical thinking, creativity, and authentic self-expression? Does relying on AI for writing, problem-solving, and social interaction undermine the autonomy that comes from developing these capacities yourself? What does it mean to flourish as a human in an AI-saturated world?
2. **Justice**: Is it fair that Maria's generation must navigate a transformed job market they didn't create? Who benefits most from LLM technology—tech companies, certain workers, society broadly? How should we address employment displacement? Do young people have a voice in how these transformative technologies are developed and deployed?
3. **Non-maleficence**: What are the potential harms to young people from widespread LLM use—loss of skills, social isolation, exposure to bias and misinformation, psychological dependence? Who is responsible for preventing or mitigating these harms?
4. **Beneficence**: What positive obligations do we have to ensure LLMs benefit rather than harm the next generation? Should education change? Should there be age restrictions or protections?

Programming Morality: Can We Build Ethical Machines?

- As AI systems make increasingly consequential decisions, computer scientists and philosophers are attempting to program machines that can reason about moral questions and behave ethically.
- **Machine morality** faces fundamental challenges: How do we translate complex ethical principles into code? Whose values should machines embody? Can algorithms capture moral nuance and context?
- Some approaches include programming explicit rules (like Asimov's Three Laws of Robotics), training AI on human moral judgments, or designing systems that can learn and adapt their moral reasoning.
- The challenge isn't just technical—it's philosophical: we must first achieve clarity about human ethics before we can teach ethics to machines.

Information Ethics as a Framework

Philosopher Luciano Floridi argues that information technologies themselves are becoming moral agents in our world—not because they're conscious, but because they shape the **infosphere** (the information environment) in which we live. We must design technologies that promote human flourishing, protect autonomy, distribute benefits justly, avoid harm, and remain explicable and accountable. Information ethics extends our moral framework to encompass the unique challenges of our digital age.

Your Digital Future: Becoming Ethical Technology Citizens

- You are growing up in a world where information technologies are not optional tools but essential infrastructure shaping every aspect of human life.
- The decisions your generation makes about designing, deploying, regulating, and using these technologies will determine whether they enhance or diminish human flourishing for decades to come.
- **Ethical technology citizenship** requires: understanding how technologies work, recognizing their moral implications, demanding accountability and explicability, and participating in democratic decisions about technological futures.

Your Responsibility and Your Power

Richard Mason argued in 1986 that information technologies require us to rethink the social contract itself. That task now falls to you. You have both the **responsibility** to think critically about the ethical dimensions of technology and the **power** to shape a digital future that respects autonomy, promotes beneficence, avoids harm, ensures justice, and maintains explicability. The common morality you inherited is necessary but not sufficient—you must extend it thoughtfully to meet challenges your ancestors never imagined.