Virtues in the Age of Social Media and AI

Character, Technology, and Human Flourishing

Brendan Shea, PhD

Rochester Community and Technical College

Spring 2026

Recap: Five Information Technology Revolutions

- In Lecture 1, we traced five major **information technology revolutions** that transformed how humans create, store, and share knowledge.
- Each revolution changed not just *what* we can do with information, but *who* controls it and *how* it shapes society.
- We saw recurring patterns: promises of democratization, fears of lost skills, debates over gatekeeping, and unintended consequences.
- Today we focus on what comes *after* Web 1.0: the rise of social media, algorithmic platforms, and artificial intelligence.

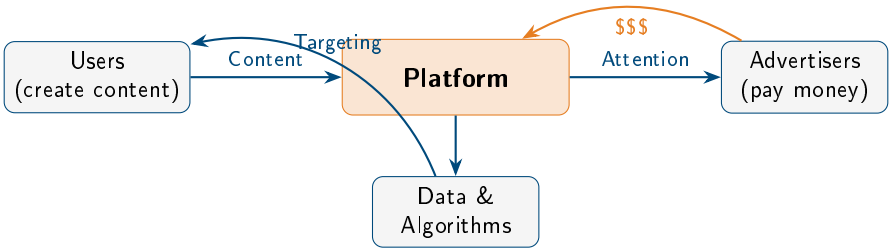
From Web 1.0 to Web 2.0

- **Web 1.0** (roughly 1991–2004) was primarily a "read-only" web: static pages created by webmasters, organized by directories like Yahoo, with users as passive consumers.
- **Web 2.0** (2004–present) transformed the web into a "read-write" platform where ordinary users generate content through blogs, social networks, and video sharing.
- This shift moved power from professional publishers to platforms that aggregate user-generated content—but also concentrated power in new ways.
- The term "Web 2.0" was popularized by Tim O'Reilly in 2004 to describe this architectural and cultural transformation.

Feature	**Web 1.0**	**Web 2.0**
Content Creation	Webmasters, companies	Anyone with an account
User Role	Consumer (read)	Producer (read/write)
Organization	Directories, portals	Search, social feeds
Examples	GeoCities, AOL	Facebook, YouTube, Wikipedia

The Architecture of Web 2.0

- **Platforms** are digital intermediaries that connect users, content creators, and advertisers—they don't create content but profit from hosting it.
- **User-generated content** means the platform's value comes from what users post: status updates, photos, videos, reviews, and comments.
- **Social graphs** map relationships between users, enabling features like friend recommendations, news feeds, and targeted advertising.
- **Network effects** make platforms more valuable as more people join—your friends are on Facebook, so you join Facebook, making it more valuable for others.



The Attention Economy and the Infosphere

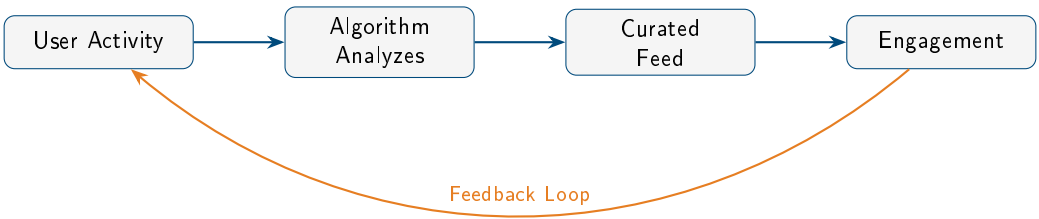
- Web 2.0 platforms are free to use because **you are not the customer—you are the product**; your attention is sold to advertisers.
- The **attention economy** treats human attention as a scarce resource to be captured, measured, and monetized through engagement metrics.
- Philosopher Luciano Floridi argues we now live in an **infosphere**—an environment where the boundaries between online and offline life have dissolved.
- In this infosphere, we have become **inforgs** (informational organisms): beings whose identities, relationships, and choices are increasingly mediated by digital information.

The Business Model

"If you're not paying for the product, you are the product." Platforms maximize **engagement** (time on site, clicks, shares) because more engagement means more advertising revenue—regardless of whether that engagement benefits users.

Algorithmic Feeds and Filter Bubbles

- **Algorithmic feeds** replaced chronological timelines around 2009–2016; platforms now decide what you see based on predicted engagement.
- **Filter bubbles** (Eli Pariser, 2011) occur when algorithms show you content similar to what you've engaged with before, narrowing your information diet.
- **Echo chambers** form when users primarily encounter views that reinforce their existing beliefs, making opposing perspectives seem alien or extreme.
- These systems optimize for *engagement*, not truth, understanding, or wellbeing—outrage and controversy often generate more clicks than nuance.



From Social Media to AI: A Continuous Story

- This lecture examines social media and AI together because they are **not separate phenomena**—AI is the next chapter in the same story.
- Social media trained us to seek validation from likes, to curate personas, and to prefer mediated interaction over face-to-face conversation.
- Generative AI (ChatGPT, 2022) extends these patterns: now we can have "relationships" with systems that simulate understanding without possessing it.
- The thinkers we'll study—Vallor, Turkle, Haidt, Floridi—have all extended their frameworks from social media to AI, seeing continuity rather than rupture.

Sherry Turkle, *Reclaiming Conversation* (2024 Preface)

"Social media was our gateway drug to conversations with machines."

Discussion Questions: Web 2.0 and the Infosphere

1. You get free access to powerful platforms; they get your attention data. Is this a fair exchange? Who benefits more?
2. How would you know if you were in a filter bubble? What steps, if any, do you take to encounter perspectives different from your own?
3. Floridi says we've become "informational organisms" whose identities are partly constituted by our digital presence. Does this match your experience? Is this troubling or simply the new normal?
4. Turkle suggests social media prepared us for AI relationships. Do you see connections between how you use social media and how you might interact with AI assistants?

What Is Virtue Ethics?

- **Virtue ethics** is a moral framework that focuses on *character*—the kind of person you are—rather than on rules or consequences alone.
- While **deontology** asks "What rules should I follow?" and **consequentialism** asks "What outcome should I produce?", virtue ethics asks "What kind of person should I become?"
- A **virtue** is an excellent trait of character—like honesty, courage, or compassion—that enables a person to live well and flourish.
- Virtue ethics emphasizes that good actions flow from good character, which is developed through **practice and habituation** over time.

Approach	Central Question	Key Concept
Deontology	What is my duty?	Rules, obligations
Consequentialism	What produces the best outcome?	Utility, welfare
Virtue Ethics	**What kind of person should I be?**	**Character, flourishing**

Aristotle and Eudaimonia

- **Aristotle** (384–322 BCE) is the founding figure of Western virtue ethics; his *Nicomachean Ethics* remains the most influential text in the tradition.
- Aristotle argues that all human action aims at some **good**, and the highest good is **eudaimonia**—often translated as "happiness" but better understood as "flourishing" or "living well."
- Eudaimonia is not a feeling but an **activity**—it is achieved by living in accordance with virtue throughout a complete life.
- For Aristotle, humans flourish by exercising their distinctive capacity: **reason**—both in thinking well and in governing emotions and desires wisely.

Aristotle, *Nicomachean Ethics* (c. 340 BCE)

"Happiness [eudaimonia] is an activity of the soul in accordance with virtue, and if there are several virtues, in accordance with the best and most complete."

Aristotle's Doctrine of the Mean

- Aristotle's **doctrine of the mean** holds that each virtue is a balanced middle point between two extremes: one of excess and one of deficiency.
- The mean is not a mathematical average but the **right amount** for this person, in this situation—it requires judgment, not calculation.
- Finding the mean requires **practical wisdom** (phronesis): the ability to perceive what virtue requires in particular circumstances.
- This framework suggests that virtuous technology use involves finding balance—neither total abstinence nor unreflective immersion.

Deficiency	Virtue (Mean)	Excess
Cowardice	**Courage**	Recklessness
Stinginess	**Generosity**	Wastefulness
Self-deprecation	**Truthfulness**	Boastfulness
Boorishness	**Wit**	Buffoonery

Habituation and Practical Wisdom

- Aristotle insists that virtues are not innate—we are not born courageous or just—but are developed through **habituation**: repeated practice until virtuous action becomes second nature.
- "We become just by doing just acts, temperate by doing temperate acts, brave by doing brave acts"—character is formed by what we *repeatedly do*.
- **Phronesis** (practical wisdom) is the master virtue: the capacity to discern the right action in particular circumstances, balancing competing considerations.
- This has profound implications for technology: our digital habits shape our character, and we need practical wisdom to navigate novel situations that Aristotle never imagined.

Technology Application

If virtues are formed by habits, then our daily technology practices—checking phones first thing in the morning, scrolling before sleep, responding instantly to notifications—are shaping who we become. What character traits do these habits cultivate?

Confucian Virtue Ethics

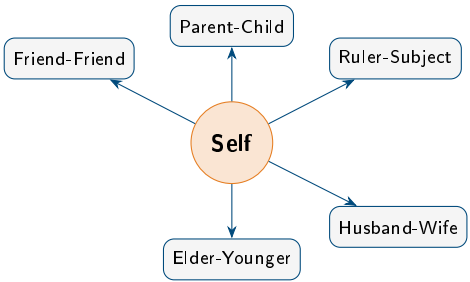
- **Confucius** (551–479 BCE) developed a virtue tradition in China contemporaneous with Greek philosophy, emphasizing social relationships and moral cultivation.
- The central virtue is **ren**—variously translated as "humaneness," "benevolence," or "human-heartedness"—the quality of caring appropriately for others.
- **Li** refers to ritual propriety: the customs, manners, and social forms through which we express respect and maintain harmonious relationships.
- The goal is to become a **junzi**—an "exemplary person" or "gentleman"—who cultivates virtue and serves as a moral model for others.

Confucius, *Analects* 12.2

"Do not do to others what you would not want done to yourself."

Confucianism: Relationships and Moral Cultivation

- Unlike Aristotle's focus on individual flourishing, Confucian ethics is fundamentally **relational**—virtue is expressed through our roles as children, parents, friends, citizens, and colleagues.
- The **Five Relationships** (wulun) define core moral bonds: parent-child, ruler-subject, husband-wife, elder-younger sibling, and friend-friend.
- Moral development requires **self-cultivation** (xiuyang): continuous effort to refine one's character through study, reflection, and practice.
- This relational focus raises important questions: How do digital technologies affect our ability to fulfill our roles and obligations to others?



Religious Virtue Traditions

- Most world religions include robust virtue traditions that overlap with and extend philosophical approaches to character ethics.
- **Christianity** distinguishes cardinal virtues (prudence, justice, fortitude, temperance) from theological virtues (faith, hope, charity/love) that require divine grace.
- **Islam** emphasizes virtues like **taqwa** (God-consciousness), **sabr** (patience), and **ihsan** (excellence in worship and conduct toward others).
- **Buddhism**'s Noble Eightfold Path cultivates virtues of wisdom (right view, intention), ethical conduct (right speech, action, livelihood), and mental discipline (right effort, mindfulness, concentration).

Tradition	Key Virtues
Christianity	Prudence, justice, fortitude, temperance; faith, hope, love
Islam	Taqwa (consciousness), sabr (patience), ihsan (excellence)
Buddhism	Right speech, right action, right mindfulness
Judaism	Tzedakah (justice/charity), chesed (loving-kindness), emet (truth)

Care Ethics and Relationships

- **Care ethics** emerged in the 1980s from feminist philosophy, emphasizing relationships, context, and emotional responsiveness over abstract universal rules.
- **Carol Gilligan** argued that traditional ethics privileged "justice reasoning" (rights, rules) over "care reasoning" (relationships, responsibilities, responsiveness to particular others).
- **Nel Noddings** developed care ethics as a complete moral theory: the fundamental moral relation is the caring encounter between "one-caring" and "cared-for."
- Care ethics asks: What do *these particular people* in *this relationship* need from each other?—a question algorithms cannot answer.

Technology Application

Care ethics highlights what's lost when relationships are mediated by screens: the ability to perceive subtle cues, respond to unspoken needs, and be physically present. Can we truly "care" through a platform optimized for engagement metrics?

Communitarian Political Philosophy

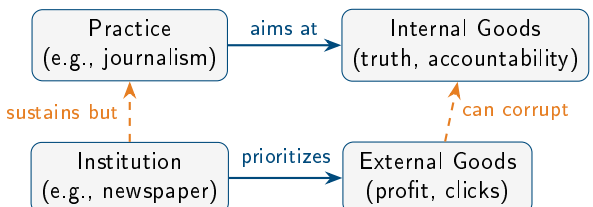
- **Communitarianism** emerged in the 1980s as a critique of liberal individualism, arguing that identity and morality are shaped by communities, traditions, and shared practices.
- **Alasdair MacIntyre**'s *After Virtue* (1981) argues that modern society has lost a coherent moral framework; we need to recover virtue within living traditions and practices.
- **Michael Sandel** criticizes the liberal idea of the "unencumbered self"—we are not abstract individuals but members of families, communities, and nations that shape who we are.
- **Charles Taylor** emphasizes that humans are "self-interpreting animals" who need shared frameworks of meaning to understand what makes life worthwhile.

Alasdair MacIntyre, *After Virtue* (1981)

"I can only answer the question 'What am I to do?' if I can answer the prior question 'Of what story or stories do I find myself a part?' "

Practices, Communities, and Virtues

- MacIntyre defines a **practice** as a cooperative activity with internal goods (excellence) and external goods (money, fame)—examples include medicine, teaching, chess, and farming.
- Virtues are the qualities needed to achieve the **internal goods** of practices: a good doctor needs honesty, compassion, and practical wisdom, not just technical skill.
- Practices are sustained by **institutions** (hospitals, universities, guilds), but institutions tend to prioritize external goods and can corrupt practices.
- This framework raises questions: Is social media use a practice? What internal goods might it have? How do platform business models affect those goods?



Why Virtue Ethics for Technology?

- Rules-based ethics struggles with technology because innovation constantly creates situations no rule anticipated—virtue ethics asks what a wise person would do.
- Consequentialist calculations are difficult when effects are unpredictable, distributed across millions of users, and unfold over decades.
- Virtue ethics focuses on **character formation**: technologies shape our habits, and habits shape who we become—this is precisely what virtue ethics addresses.
- The virtue tradition emphasizes **practical wisdom** (phronesis)—the ability to navigate novel situations wisely—which is exactly what digital life demands.

Challenge	**Why Virtue Ethics Helps**
Rapid change	Focus on character, not outdated rules
Unpredictable effects	Cultivate wisdom for novel situations
Habit formation	Directly addresses how practices shape character
Relational complexity	Care ethics and Confucianism emphasize relationships
Loss of community	Communitarianism diagnoses fragmentation

Discussion Questions: Virtue Ethics Foundations

1. Aristotle says we become what we repeatedly do. What character traits are your daily technology habits cultivating? Are these the traits you want to have?
2. What might the "virtuous mean" look like for smartphone use—avoiding both the excess of constant connection and the deficiency of total disconnection?
3. Confucian ethics emphasizes fulfilling roles in relationships. How does technology help or hinder your ability to be a good friend, child, sibling, or citizen?
4. Think of an activity you care about (music, sports, art, learning). How do social media platforms affect whether you pursue its internal goods (excellence) or external goods (likes, followers)?

Shannon Vallor: Technology and the Virtues

- **Shannon Vallor** is a philosopher of technology at the Edinburgh Futures Institute; her book *Technology and the Virtues* (2016) applies virtue ethics to emerging technologies.
- Vallor argues that we face a **technosocial condition**: technology and society co-evolve so rapidly that traditional moral frameworks struggle to keep pace.
- She synthesizes Aristotelian, Confucian, and Buddhist virtue traditions to develop a framework of **technomoral virtues**—character traits needed to flourish with technology.
- Her central claim: we cannot rely on rules or calculations alone; we need cultivated **practical wisdom** to navigate technological life well.

Shannon Vallor, *Technology and the Virtues* (2016)

"Virtue ethics asks not 'What should I do?' but 'What kind of person should I become, and how?'—questions essential for beings whose characters are increasingly shaped by technological habits."

Vallor's Twelve Technomoral Virtues

- Vallor identifies **twelve technomoral virtues** that draw on multiple traditions and address the specific challenges of life with emerging technologies.
- These virtues are not new inventions but classical virtues **reinterpreted** for contexts shaped by social media, AI, surveillance, and global connectivity.
- The list culminates in **technomoral wisdom**—the integrating virtue that enables us to apply all others appropriately in novel technological situations.
- Vallor emphasizes that these virtues must be **cultivated through practice**—they cannot be downloaded or installed like software.

Self-Regarding	**Other-Regarding**	**Integrating**
Honesty	Empathy	Perspective
Self-Control	Care	Magnanimity
Humility	Civility	**Technomoral Wisdom**
Courage	Justice	
	Flexibility	

Vallor's *The AI Mirror* (2024)

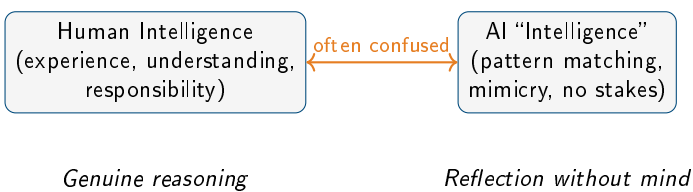
- In *The AI Mirror* (2024), Vallor extends her framework to generative AI, arguing that AI is best understood as a **mirror**, not a mind.
- Just as a mirror reflects your image without containing a second face, AI systems reflect patterns from human data without possessing understanding or experience.
- Vallor warns that the greater danger is not superintelligent AI but **undermining our sense of ourselves** as responsible, reasoning, free intelligences.
- Silicon Valley promotes a "debased version of what we are"—viewing humans as "soft, wet computers"—which erodes our confidence in human wisdom and agency.

The Mirror Metaphor

"When you look in a mirror, you know there isn't a second face looking back at you from the other side of the glass. That's just a reflection of your face. There's nothing on the other side participating."

Vallor on AI and Human Wisdom

- Vallor critiques the redefinition of intelligence to mean "performing economically valuable tasks"—this conflates competence with genuine understanding.
- Large language models do not truly reason; they "bullshit a whole chain of thought" that can collapse into nonsense when pressed—mimicry without comprehension.
- **Thinking requires experience**: "Thinking without experience is like water without hydrogen—you've taken something out that loses its identity."
- Yet Vallor remains hopeful: technology at its core *can* be humane—"We're at a moment when we need to rebuild our confidence in the capabilities of humans to reason wisely."



Sherry Turkle: Alone Together

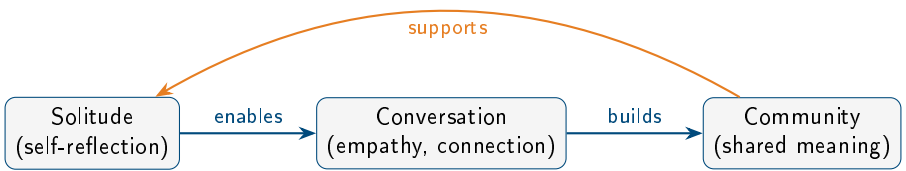
- **Sherry Turkle** is a psychologist and professor at MIT who has studied human-technology relationships for over forty years.
- In *Alone Together* (2011), Turkle documents a troubling paradox: we are increasingly **connected** through devices but increasingly **alone** in meaningful human contact.
- She observes a "flight from conversation": people prefer texting to talking, curated posts to spontaneous interaction, and controllable digital exchanges to unpredictable face-to-face encounters.
- Turkle warns that we are designing technologies that offer the **illusion of companionship** without the demands of real relationship.

Sherry Turkle, *Alone Together* (2011)

"We expect more from technology and less from each other... We are tempted to think that our little 'sips' of online connection add up to a big gulp of real conversation. But they don't."

Turkle: Reclaiming Conversation

- In *Reclaiming Conversation* (2015), Turkle argues that face-to-face conversation is the most humanizing thing we do—and we are systematically abandoning it.
- Conversation develops **empathy**: children learn to read faces, tolerate ambiguity, and respond to others' emotions through unstructured talk, not screens.
- Turkle identifies the importance of **solitude**—the capacity to be alone with one's thoughts—which is eroded when we reach for phones at every moment of boredom.
- She calls for reclaiming "sacred spaces" for conversation: the kitchen table, the dining room, the car, the classroom—places where phones should not intrude.



Turkle on Artificial Intimacy (2024)

- In her 2024 preface to *Reclaiming Conversation*, Turkle extends her analysis to AI chatbots that simulate emotional connection and understanding.
- She argues that social media was our **gateway drug**: it trained us to accept mediated relationships, preparing us for "conversations" with machines.
- AI companions offer **frictionless relationships**—always available, never demanding, endlessly patient—but this convenience comes at a cost.
- Turkle invokes psychoanalyst Erik Erikson's warning about engineers who "aren't convinced that people have an interior"—reducing humans to information processors.

The Problem of Unrequited Love

"We nurture what we love, but we love what we nurture. We love what we allow ourselves to relate to. It's important to remember that this love is unrequited."

Turkle: What Machines Cannot Provide

- Turkle challenges the view (held by figures like Eric Schmidt) that AI with access to "every paper on anxiety" could replace human therapists or friends.
- Information is not understanding: "When you tell your troubles to a machine, it has no stake in the conversation. It can't feel pain. It has no life in which you play a part."
- The **performance of emotion** does not make machines more human—but it does challenge what we think makes *people* special.
- Drawing on philosopher Emmanuel Levinas, Turkle emphasizes that the **human face** initiates an ethical compact—embodied presence creates obligations that screens cannot.

The Levinas Point

For Levinas, encountering another person's face creates an immediate ethical demand: "Here I am, responsible for you." A chatbot has no face, no vulnerability, no mortality—and thus makes no genuine ethical claim on us, nor we on it.

Luciano Floridi: Five Principles for AI Ethics

- **Luciano Floridi** is a philosopher at Yale and Oxford who developed the concepts of the "infosphere" and "inforgs" we encountered earlier.
- In his "Unified Framework of Five Principles for AI in Society" (2019), Floridi analyzed 47 AI ethics documents and found convergence around **five core principles**.
- Four principles come from **bioethics**: beneficence (do good), non-maleficence (avoid harm), autonomy (respect human choice), and justice (be fair).
- Floridi adds a fifth principle specific to AI: **explicability**—combining intelligibility ("How does it work?") with accountability ("Who is responsible?").

Principle	**Core Question**
Beneficence	Does this AI promote human wellbeing and flourishing?
Non-maleficence	Does this AI avoid causing harm to individuals or society?
Autonomy	Does this AI respect and preserve human decision-making?
Justice	Does this AI treat people fairly and avoid discrimination?
Explicability	Can we understand how it works and who is responsible?

Floridi: AI as Agency Without Intelligence

- Floridi describes AI as "smart agency on tap"—systems that can **act** in the world (book flights, filter content, recommend products) without genuine understanding.
- This is **agency without intelligence**: AI systems produce effects and make consequential decisions, but they don't "know" what they're doing in any meaningful sense.
- The ethical challenge is that we are delegating more decisions to systems that cannot be held responsible, understand context, or exercise moral judgment.
- Floridi's explicability principle responds to this: if AI acts without understanding, *humans* must understand how it works and remain accountable for its effects.

The Explicability Gap

Many AI systems are "black boxes": even their creators cannot fully explain why they produce particular outputs. This creates an **explicability gap** that threatens accountability—if no one understands how a decision was made, who is responsible when it causes harm?

Jonathan Haidt: The Anxious Generation

- **Jonathan Haidt** is a social psychologist at NYU whose book *The Anxious Generation* (2024) argues that smartphones and social media have caused a mental health crisis among youth.
- Haidt identifies **four foundational harms** of phone-based childhood: social deprivation, sleep deprivation, attention fragmentation, and addiction.
- The shift from "play-based childhood" to "phone-based childhood" (roughly 2010–2015) correlates with sharp increases in anxiety, depression, and self-harm among adolescents.
- Haidt proposes concrete reforms: no smartphones before high school, no social media before 16, phone-free schools, and more unsupervised outdoor play.

Social Deprivation | Sleep Deprivation | Attention Fragmentation | Addiction

Haidt on AI: Amplifying Existing Harms

- In a 2023 essay with Eric Schmidt, Haidt argues that AI is "about to make social media much more toxic"—not a separate problem but an **amplifier of existing harms**.
- Haidt identifies four imminent AI threats: super-empowering bad actors, drowning in fake evidence (deepfakes), accelerating polarization, and enabling new forms of harassment.
- AI companions pose particular risks for young people: chatbots designed to simulate friendship and intimacy may further displace real human relationships.
- Haidt's analysis shows **continuity**: the same dynamics that made social media harmful (attention capture, social comparison, displacement of real connection) are intensified by AI.

Haidt and Schmidt (2023)

"Generative AI is going to pour gasoline on the fire of social media's harms... The same business model—capturing attention and selling it to advertisers—will now deploy far more powerful tools to keep users engaged."

Critics of Haidt: The Evidence Debate

- Haidt's claims have provoked significant scholarly debate; critics argue the evidence for social media causing mental health harms is weaker than he suggests.
- **Candice Odgers** (UC Irvine) argues that correlations between screen time and mental health are small and inconsistent, and that Haidt overstates causation.
- **Amy Orben** (Cambridge) notes that effect sizes in studies are often tiny—comparable to the "harm" of eating potatoes—and that moral panic may cause its own harms.
- The debate matters: if Haidt is wrong, we may restrict beneficial technologies; if critics are wrong, we may fail to protect vulnerable youth from genuine harm.

Haidt's Position	**Critics' Response**
Strong causal link to mental health crisis	Correlations are small and inconsistent
Smartphones are the key variable	Many factors changed simultaneously
Urgent action needed now	Moral panic may cause its own harms
Natural experiments support causation	Methodological concerns with studies

Social Media and Political Polarization

- Social media platforms have been accused of driving **political polarization**—the widening gap between political groups and the increasing hostility between them.
- **Filter bubbles** and algorithmic curation may expose users primarily to views they already hold, making opposing views seem extreme or incomprehensible.
- **Outrage amplification**: content that provokes anger and moral outrage generates more engagement, so algorithms systematically promote divisive content.
- The shift from broadcast media (shared national conversation) to personalized feeds means citizens increasingly live in **different information worlds**.

The Erosion of Shared Reality

- Political philosopher **Hannah Arendt** argued that democratic life requires a **common world**—shared facts and frameworks that allow citizens to disagree productively.
- Social media threatens this common world: when groups cannot agree on basic facts, democratic deliberation becomes impossible.
- **Deepfakes** and AI-generated content accelerate this erosion—soon any image, video, or audio can be fabricated, making "seeing is believing" obsolete.
- The result is not that everyone believes lies, but that **no one believes anything**—a cynicism that benefits those who prefer citizens disengaged and confused.

Hannah Arendt, *The Origins of Totalitarianism* (1951)

"The ideal subject of totalitarian rule is not the convinced Nazi or the convinced Communist, but people for whom the distinction between fact and fiction... no longer exist[s]."

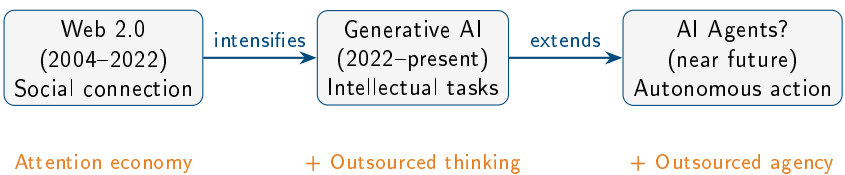
Virtue and Democratic Citizenship

- Democratic self-government requires **civic virtues**: the character traits that enable citizens to participate thoughtfully in collective decision-making.
- Classical republicans emphasized virtues like **public-spiritedness** (concern for common good over private interest), **deliberative openness** (willingness to hear other views), and **civic courage** (speaking truth even when unpopular).
- Social media may erode these virtues: rewarding performance over deliberation, tribalism over public-spiritedness, and outrage over thoughtful engagement.
- The question is not just "Is social media making us unhappy?" but "Is it making us **worse citizens**?"

Civic Virtue	**How Social Media May Erode It**
Public-spiritedness	Algorithms reward tribal loyalty over common good
Deliberative openness	Filter bubbles limit exposure to other views
Civic courage	Pile-ons and cancellation punish dissent
Truthfulness	Engagement metrics reward sensationalism

AI as the Next Chapter in the Same Story

- The arrival of generative AI (ChatGPT, November 2022) is not a break from the social media era but its **continuation and intensification**.
- The same dynamics apply: attention capture, engagement optimization, displacement of human connection, and the erosion of skills we once cultivated.
- What social media did to our *social* lives—fragmenting attention, encouraging performance, substituting metrics for meaning—AI threatens to do to our *intellectual* lives.
- All four thinkers we've studied see this continuity: Vallor (technomoral virtues still needed), Turkle (gateway drug), Haidt (gasoline on fire), Floridi (agency without intelligence).



AI and the Virtue of Honesty

- **Honesty** is a core virtue in every tradition we've studied—Aristotle's truthfulness, Confucian sincerity, religious commands against bearing false witness.
- Generative AI makes deception radically easier: **deepfakes** can fabricate video of anyone saying anything; AI can generate convincing fake news at scale.
- But there's a subtler threat: AI-generated content blurs the line between human and machine authorship—is submitting AI-written work as your own a form of dishonesty?
- Vallor's technomoral honesty requires "habitually presenting oneself and one's work accurately"—a virtue increasingly difficult to practice and verify in an AI-saturated world.

The Authenticity Question

When you read a heartfelt social media post, how do you know a human wrote it? When you receive a personalized message, how do you know someone actually composed it for you? AI threatens not just truth but **authenticity**—the connection between expression and genuine human intention.

AI and the Intellectual Virtues

- The **intellectual virtues**—curiosity, careful reasoning, intellectual humility, love of truth—are capacities developed through practice and struggle.
- If AI can instantly produce essays, solve problems, and answer questions, what happens to the **habits of mind** we develop by doing these things ourselves?
- There's a parallel to GPS navigation: convenient, but studies show heavy GPS users develop weaker spatial reasoning and memory for routes.
- The question is not whether AI is useful (it obviously can be) but whether outsourcing intellectual labor **atrophies the virtues** that labor was meant to cultivate.

The Student's Dilemma

A student can use AI to write an essay in minutes. But the assignment existed to develop skills: organizing thoughts, constructing arguments, expressing ideas clearly. If AI does the work, the student gets the grade but misses the growth. What virtue would help navigate this tension?

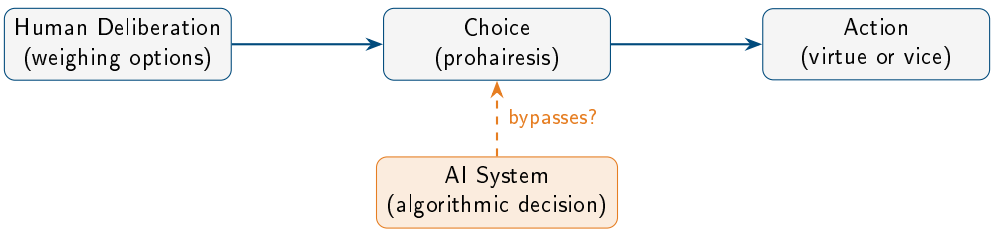
AI and the Relational Virtues

- **Relational virtues**—empathy, care, compassion, fidelity—are cultivated through relationships with other humans who have their own needs, perspectives, and vulnerabilities.
- AI companions offer relationships without friction: they never get tired, never disagree, never have needs of their own—but this "ease" may prevent virtue development.
- Recall Turkle: "When you tell your troubles to a machine, it has no stake in the conversation. It can't feel pain. It has no life in which you play a part."
- And Vallor: there is "nothing on the other side participating"—the reciprocity that makes human relationships both difficult and morally formative is absent.

Feature	Human Relationship	AI Companion
Availability	Limited, scheduled	Always on demand
Reciprocity	Mutual needs and care	One-directional
Friction	Disagreement, conflict	Designed to please
Stakes	Real consequences	No genuine stakes

AI and Human Agency

- For Aristotle, virtue requires **prohairesis**—deliberate choice—which is why we don't praise or blame people for involuntary actions or accidents.
- AI systems increasingly make choices that affect our lives: what content we see, which job applicants advance, what medical treatments are recommended.
- When AI makes decisions, **human agency is displaced**: we become subjects of algorithmic governance rather than authors of our own choices.
- Floridi's **autonomy principle** insists that AI should preserve and enhance human decision-making, not replace it—but current systems often do the opposite.



Can AI Systems Be Virtuous?

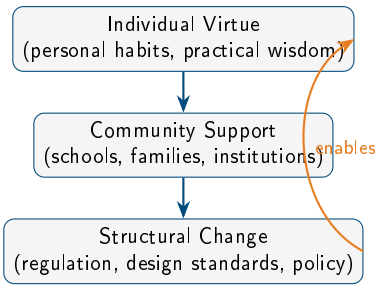
- Some argue we should design AI systems to be "virtuous"—to embody honesty, fairness, care—but can a machine *possess* virtues in any meaningful sense?
- Aristotle's virtues require **feeling the right emotions** at the right time—courage involves feeling appropriate fear, not just acting bravely.
- Vallor argues that genuine thinking requires **experience**: "Thinking without experience is like water without hydrogen—you've taken something out that loses its identity."
- AI can *simulate* virtuous behavior and produce outputs that *look* virtuous, but simulation is not the same as genuine character.

The Simulation Problem

An AI can be programmed to say kind things, but it cannot *care*. It can produce fair-seeming outputs, but it has no *commitment* to justice. Virtues are not just behaviors but stable dispositions rooted in understanding, emotion, and choice—things AI systems lack.

Vallor's Challenge: Collective Action

- Vallor argues that individual virtue is necessary but not sufficient—we also need **collective action** to reshape the technological environment itself.
- Personal discipline (limiting screen time, cultivating attention) helps, but platforms are *designed* to undermine such discipline—willpower alone cannot win.
- We need structural changes: regulations that constrain harmful business models, design standards that respect human flourishing, and institutions that support virtue cultivation.
- This connects to communitarian insights: virtues are cultivated within communities and practices, not by isolated individuals resisting hostile environments.



Discussion Questions: Social Media, Politics, and AI

1. Haidt claims smartphones caused a mental health crisis; critics say the evidence is weak. How should we make policy when experts disagree? Should we err on the side of caution?
2. Arendt says democracy requires a "common world" of shared facts. Is this still possible in an age of personalized feeds and AI-generated content? What would it take to rebuild?
3. If AI can write your essays and solve your problems, how do you develop the intellectual virtues those tasks were meant to cultivate? Is there a virtuous way to use AI for learning?
4. Vallor says individual virtue isn't enough—we need structural change. But structural change requires political action, which requires virtuous citizens. How do we escape this circle?

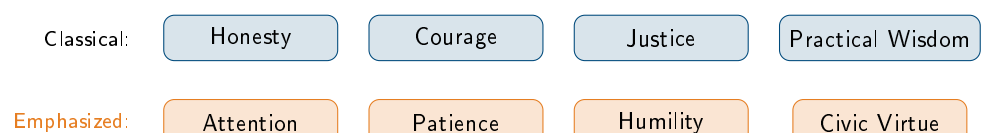
What Virtue Traditions Teach Us

- **Aristotle** teaches that character is formed by habits—our repeated digital practices are shaping who we become, whether we intend it or not.
- **Confucius** reminds us that virtue is relational—we must ask how technology affects our ability to fulfill our roles and obligations to others.
- **Care ethics** emphasizes attention and responsiveness to particular others—qualities threatened by distraction and algorithmic mediation.
- **Communitarians** warn that virtues require supportive communities and practices—isolated individuals cannot cultivate character against hostile environments.

Tradition	**Key Insight for Digital Life**
Aristotelian	Habits form character; choose digital habits wisely
Confucian	Virtue is relational; technology affects all our roles
Care Ethics	Attention to particular others requires presence
Communitarian	We need communities that support virtue cultivation

Virtues for the Digital Age

- The classical virtues remain essential—honesty, courage, justice, practical wisdom—but our technological context gives them new applications and urgency.
- **Attention** emerges as a crucial virtue: the capacity to focus, to be present, to resist the constant pull of notifications and feeds.
- **Patience** and **tolerance for friction** counter the demand for instant gratification—meaningful things often require sustained effort and discomfort.
- **Humility** about our own knowledge becomes vital when AI can produce confident-sounding nonsense and filter bubbles confirm our biases.



Looking Ahead: Reclaiming Friction

- Emerging technologies—AI agents, brain-computer interfaces, immersive virtual worlds—will intensify the challenges we've discussed, not resolve them.
- The temptation will be to embrace ever more seamless, frictionless experiences—but friction is often where growth, meaning, and virtue are forged.
- Turkle calls us to reclaim "sacred spaces" for unmediated human connection: meals without phones, classrooms without screens, conversations without interruption.
- The goal is not to reject technology but to **use it wisely**—with the practical wisdom to know when convenience serves flourishing and when it undermines it.

The Case for Friction

Learning is hard. Relationships are messy. Thinking takes time. Democracy requires patience. These difficulties are not bugs to be engineered away but **features** of a meaningful human life. The virtuous response to technology is not rejection but **wise integration**—preserving the friction that makes us grow.

Final Reflection: Who Do You Want to Become?

- Virtue ethics asks not "What should I do right now?" but "What kind of person am I becoming through my choices and habits?"
- Every time you reach for your phone, scroll a feed, ask an AI to write for you, or choose a screen over a face—you are forming your character.
- The thinkers we've studied—Vallor, Turkle, Haidt, Floridi—offer not despair but a challenge: to cultivate the wisdom our technological moment demands.
- The question is not whether to use technology, but **how to use it in ways that help you flourish** as the person you want to become.

Questions for Reflection

1. What digital habits do you have that you're proud of? Which ones concern you?
2. What would "technomoral wisdom" look like in your daily life?
3. What is one concrete change you could make this week to better align your technology use with the person you want to become?