AI

IMMIGRANTS

"The Bloody Algos Are Here!"

Kevin Ryan

AI Immigrants

AI Immigrants

The Bloody Algos Are Here!

Kevin Ryan

First Edition, 2026

Disclaimer

This book presents opinion and commentary informed by publicly available sources. Every effort has been made to ensure accuracy and fairness; however, all interpretations, conclusions, and reflections are the author's own.

References to news events, data, or public statements are accurate to the best of the author's knowledge as of 2025. Developments occurring after that date may not be reflected in this edition.

This publication is intended for informational and educational purposes only. It does not constitute professional, legal, financial, or medical advice. Readers should consult qualified professionals before making decisions based on the material herein.

If you believe any information in this book to be inaccurate, incomplete, or misleading, please contact the author or publisher. Verified concerns will be reviewed in good faith and, where appropriate, corrected in future editions or online updates.

Author's Note

Before we dive into algorithms and acronyms, I want to talk about who's doing the writing and why I landed on the odd idea of treating artificial intelligence like a newcomer in your neighbourhood.

I grew up British, surrounded by stories of kings, knights and cricket. But the more you scratch at those stories, the more you realise how porous the borders really are. Our royal family changed its name from Saxe-Coburg and Gotha to Windsor during the First World War to hide its German roots. Roast beef? Arrived with the French centuries ago. Curry? A gift from India. England's Saint George was a Greek or Turkish soldier who never once set foot on our island; the cross we wave at football matches was permanently borrowed from overseas.

We are, in truth, a nation of mongrels. Romans built our roads. Saxons and Vikings lent us their tongue. Normans imported castles and a French flair for law. Later waves came from Ireland, India, the Caribbean, Africa and Eastern Europe. Even our most nationalist politicians turn out to be walking reminders of migration. Nigel Farage's Huguenot sixth great-grandfather fled religious violence in France and likely crossed the Channel in a small boat. Boris Johnson's family tree winds through Turkey and France. If the people who complain loudest about outsiders are mongrels, what does that say about the rest of us?

Why does this matter in a book about artificial intelligence? Because metaphors shape how we think. When we talk about AI as a "tool," we emphasise utility. When we describe it as a "species," we evoke wonder and fear. I see these systems as immigrants of

a different sort: borderless, invisible and already living among us. They arrive not on dinghies but through software updates and server farms. They speak every language because we trained them on our own words. They enter our homes via smart speakers, our pockets via recommendation engines, and our institutions via decision-making systems. They're not sentient; they don't eat or sleep. But like immigrants, they cross boundaries, unsettle the status quo, and force societies to renegotiate who belongs and on what terms.

Seeing AI this way also reveals our prejudices. Immigrants have long been scapegoated for unemployment, crime and cultural decline. Algorithms stir similar emotions. We worry that a chatbot will steal a call-centre job or that autonomous trucks will put drivers out of work. But focusing our frustration on the newcomers—whether human or digital—lets those in power off the hook. Corporations choose how to deploy technology. Governments decide whether to protect workers. Our anxieties reflect deeper failures of policy and imagination.

Writing this was an exercise in looking at my own biases. I grew up thinking of my country as an island fortress. In reality it has been a crossroads for millennia. The best way to honour that heritage is to greet newcomers—flesh or silicon—with curiosity rather than fear. But curiosity must be coupled with vigilance. Immigrants enrich societies when they are welcomed with rights and responsibilities. AI enriches us only if we build systems that reflect our values: fairness, accountability, and a recognition that people are more than data points.

So consider this an invitation. I'm not laying down immutable truths. I'm sharing a lens, shaped by my upbringing in a mongrel nation and informed by a world in flux. The chapters that follow will use the language of immigration to illuminate what AI does and doesn't do, what it threatens and what it could offer. I'll avoid jargon and predictions of robot apocalypses. Instead, I'll ask how to live with these digital strangers in our midst.

Contents

Chapter 1: They're Taking Our Jobs

Imagine hearing a knock on your door at dawn. You open it and find no human on the step but instead a courier package containing a small tablet. Inside the glass, an avatar stares back. "I'm here to train for your job," it says in a pleasant voice. "I don't need coffee breaks, benefits, or overtime." It's not a person who has crossed an ocean or climbed a fence. It's an algorithm — an Algo, as I like to call them. And it is applying for work in every sector you can imagine.

That image might feel like science fiction, but for millions of workers the knock is already real. Artificial intelligence is no longer a lab curiosity. It answers phones, reviews résumés, predicts when machines will break, writes news summaries, paints pictures, reads X-rays, keeps our cars in lane, and even generates code. Along the way it has resurrected an old fear: **"They're taking our jobs."**

In this chapter we'll explore that fear, situate it within the history of people scapegoating immigrants for economic anxiety, and ask whether blaming Algorithms distracts us from deeper failures. We'll travel through warehouses, call centres, operating rooms, film studios, and software teams. We'll meet real humans whose work has been redefined by machines. And we'll see how public opinion surveys and economic studies — from the World Economic Forum to the International Labour Organization — paint a complicated picture: jobs will be lost, jobs will be created, and many more will be transformed. The question is what those transformations mean for our sense of purpose and our social contract.

Old ghosts in new servers

The first time machines provoked mass panic about job theft, steam engines were pumping the looms of northern England. In 1811–1812 the Luddite movement erupted when textile workers smashed the new knitting frames that threatened their livelihoods. The government responded with troops and eventually made machine-breaking a crime punishable by death; seventeen men were executed in 1812. The Luddites' rebellion lasted five years and was fuelled by collapsing wages and hunger. Today we look back and see that mechanisation ultimately created textile jobs and made goods cheaper. But that historical distance should not blind us to the desperation of workers watching their skills become obsolete.

A few decades later, on the other side of the Atlantic, another wave of job fear took hold. Chinese labourers came to the American West to build railroads, mine gold, and sew garments. They often worked for whatever wages they could, sending money home and repaying passage debts. Non-Chinese labourers, who commanded higher wages and had more political clout, resented their willingness to work cheaply and feared being squeezed out of jobs. This economic anxiety, layered with racism and xenophobia, culminated in the Chinese Exclusion Act of 1882, which suspended the immigration of Chinese workers for ten years and required those already in America to carry identification papers. It was one of the first broad restrictions on immigration in U.S. history, a legal embodiment of the cry "They're taking our jobs."

Fast-forward to 2016 and the United Kingdom's Brexit referendum. Immigration was the defining issue of the Leave campaign. Prominent supporters argued that European Union workers were depressing wages and causing unemployment. Priti Patel, one of the leading figures, said that the wealthy leaders of the Remain side would never understand the devastating effect EU immigration had on ordinary people. Campaign adverts warned that continued free movement would "place considerable pressure on the wages of low-paid British workers" and stoked fears of strained

schools, hospitals, and public benefits. The accuracy of those claims was widely debated, but their emotional resonance was undeniable: the outsider is to blame for stagnant wages and crowded waiting rooms.

The metaphor of the Algos as immigrant taps into these same feelings. Just as newcomers who look, speak, or worship differently have long been portrayed as job stealers, machines that "think" are cast as economic invaders. The anthropomorphism of AI — voice assistants with names, chatbots with personalities — creates an emotional illusion that amplifies our fears and amplifies the sense that something foreign has arrived to compete with us. But unlike a human worker, an algorithm does not aspire to citizenship or a better life. It is a tool created by us to serve us, though often designed and deployed by a small group of companies with little democratic oversight. Blaming the tool risks letting those wielding it off the hook.

How many jobs are really at stake?

Let's put some numbers behind the anxiety. The World Economic Forum's *Future of Jobs Report 2025* surveyed employers and found that roughly 40% of them expect to reduce their workforce because of technology, while at the same time they expect to create 11 million new jobs and eliminate about 9 million — a net positive but with large displacement. McKinsey's 2023 research on generative AI estimated that current technologies can automate 60–70% of employees' time, and that half of today's work activities could be automated between 2030 and 2060 (with a midpoint around 2045). The International Monetary Fund (IMF) projects that AI will affect almost 40% of jobs globally, rising to 60% in advanced economies. In high-income countries, the International Labour Organization (ILO) reports that about one in four workers have some degree of generative-AI exposure, with clerical roles and women especially vulnerable. Across the OECD, around 14% of jobs are automatable and another 32% will face substantial change, meaning half of tasks may shift or disappear. These num-

bers aren't predictions of mass unemployment so much as signs that job content is changing.

Public opinion mirrors this unease. A 2025 Pew Research survey of more than 5,000 U.S. workers found that 52% worry about AI's impact on jobs, with only 6% believing it will create more opportunities. Most workers (63%) report rarely or never using AI, yet a third of non-users think some of their work could be done by algorithms. In the U.K., a 2025 poll by the conciliation service Acas showed that 26% of employees fear AI will lead to job losses, 17% worry about machines making mistakes, and 15% are concerned about lack of regulation. McKinsey's 2025 global survey reveals a strange contradiction: nearly all companies plan to increase AI investment and employees are three times more likely than leaders realise to think AI will replace 30% of their work in the coming year. Yet only a slight majority of workers call themselves AI optimists; 41% are apprehensive.

The scale of potential disruption clearly invites worry. But the raw numbers hide important nuances. For one, AI does not replace entire jobs so much as specific tasks within those jobs. Many roles consist of dozens of discrete activities; some are ripe for automation, others require human judgment, creativity, or emotional connection. The ILO's 2025 update notes that, despite generative AI's advances, the mean automation score across thousands of tasks has barely changed since 2023, and most jobs will be transformed rather than eliminated. Moreover, the benefits and burdens will be unevenly distributed. High-income countries and white-collar workers face greater exposure, meaning that the sense of technological invasion may be felt most acutely by those who have long been insulated from automation.

Algorithms in the warehouse and on the highway

Visit an Amazon fulfilment centre and the first thing you notice is not the sound of human chatter but the whir of robots. The company has deployed over 200,000 robots that pick, sort, and transport goods alongside human staff. These machines don't look like humanoid androids. They're orange pods that zip beneath shelves, delivering items to workers who pack boxes. The robots save footsteps and reduce injuries, but they also reshape jobs, turning what used to be a roaming role into a station-based one. Some workers appreciate the reduction in physical strain; others complain of monotonous, high-paced work dictated by algorithmic timers. The Algos have not taken every job, but it has changed the nature of warehouse labour and concentrated power in the hands of those who program the systems.

Transportation is undergoing a similar metamorphosis. Autonomous trucking companies are piloting self-driving rigs that handle highway driving between hubs while human drivers navigate city streets. In 2025 the global autonomous truck market was valued at around \$1.74 billion, with Chinese firm Inceptio delivering 400 self-driving trucks to a major courier service and more than 1,400 autonomous vehicles being tested across U.S. states like California, Texas, and Arizona. Supporters argue that these systems could alleviate driver shortages and improve safety; critics worry about deskilling and job losses. Yet even the boosters acknowledge that new roles will emerge in remote vehicle operation, fleet management, and AI support. Truckers may morph into supervisors, overseeing multiple rigs from a control room, much like airline pilots manage autopilot. This pattern — machines doing the long haul, humans doing the tricky bits — reappears across sectors.

On the consumer side, electric semi-trailer trucks (or 'semis') like Tesla's long-awaited model promise autopilot features and a range of up to 500 miles. These vehicles are marketed less as robot overlords than as safer, more efficient tools. But if they

make long-distance trucking more accessible to fewer drivers, will wages fall? Will smaller carriers survive? These are policy questions disguised as technical ones.

The chatbot will see you now… or will it?

Customer service has become one of the front lines in the job debate because it combines repetitive tasks with high human stakes. Several firms have boasted about replacing customer support staff with chatbots. Dukaan, an Indian e-commerce platform, eliminated 90% of its customer support staff, claiming an 85% cost reduction and faster response times. Ikea introduced a bot named Billie to handle common queries while retraining human call-centre workers as interior design advisors. More famously, the Swedish fintech company Klarna laid off about 700 customer service workers, saving roughly $10 million, and replaced them with AI. Months later, after a surge of customer frustration, Klarna's chief executive admitted the bot could not handle nuance or empathy and vowed to ensure human agents would always be available.

These stories reveal two truths. First, companies will use any tool to reduce costs, and AI is the latest lever. Second, the human touch still matters. Customers complaining about overbilling, lost packages, or denied refunds don't just want an answer; they want a sense that someone cares. Even the most sophisticated language model cannot replace human empathy, at least not yet. AI can handle routine queries and triage, but escalation inevitably requires a person. The fear that an Algos will simply displace all call-centre jobs ignores the nuance of real interactions — nuance that the corporate hype often ignores as well.

From scribes to software engineers

In healthcare, AI is inching its way into the exam room, but not in the way you might imagine. Instead of replacing surgeons or nurses, many hospitals are experimenting with ambient documentation tools — microphones and algorithms that automatically record, transcribe, and draft medical notes during patient visits. Massachusetts General Brigham, a Harvard-affiliated institution and one of the largest health systems in the U.S., is piloting such tools with over 600 physicians. Doctors report going home on time with their notes already finished, reclaiming precious personal hours. RAND scholars argue that these administrative applications are the least risky and most valuable uses of AI in healthcare, likely to ease burnout rather than produce widespread job loss. The scribes who currently shadow doctors and type up visits may eventually be displaced, but many will be redeployed into patient-facing roles or data quality assurance.

The creative fields, often romanticised as uniquely human, are equally in flux. Film and television studios are harnessing AI to remove green screens, generate rough cuts, analyse viewer preferences, and even draft scripts. The visual effects team behind the 2022 film *Everything Everywhere All at Once* used AI tools from the New York–based company Runway to remove backgrounds quickly, allowing a small crew to produce high-quality results. Analysts predict that as text-to-video tools, AI-driven advertising, and localisation technologies mature, there will be a shortage of creatives proficient in AI by 2025. But rather than replacing directors or writers, these tools automate laborious tasks and free creatives to focus on storytelling. Just as the advent of digital editing did not eliminate editors but required new skills, AI may reshape creative work rather than erase it.

Software engineering — the very domain that birthed AI — is not immune to its disruptive potential. A study involving Microsoft, Accenture, and a Fortune 100 company introduced GitHub Copilot to thousands of developers. Productivity increased 26% on average. Junior programmers saw gains of 27–39% while senior

developers posted smaller improvements. Within a year, about 60% of developers adopted the tool. These numbers suggest that generative AI can boost output and might allow fewer people to do the same work. But will this lead to layoffs? The study's authors are cautious. They note that long-term effects remain uncertain — AI could complement experienced engineers, freeing them to tackle higher-level design, or it could enable companies to hire fewer juniors. Once again, whether the Algos are a job killer or job creator depends less on the technology than on how companies choose to deploy it.

What people fear when they fear machines

If AI were simply replacing menial tasks and leaving us time to do creative, relational work, most of us would cheer. The anxiety arises because that is not how the spoils of automation have been distributed. Since 1980, automation has eliminated more jobs than it has created in the United States, while wages for many workers have stagnated. It is this history — decades of outsourcing, gig work, and precarious contracts — that primes us to fear the Algos.

Public sentiment is shaped not only by personal experience but by narratives. Surveys capture worry, but stories give it shape. The Leave campaign's claim that EU migrants depressed wages and the 19th-century caricature of Chinese labourers accepting any wage draw on the same psychological script: the outsider is willing to do what you do for less, and the system will replace you if it can. That script is effective because it contains a kernel of truth. Employers are incentivised to cut costs. In a deregulated labour market, there is little to stop them from replacing workers with machines or lower-paid newcomers. People are right to be angry when their jobs vanish and their communities crumble. But directing that anger at immigrants or algorithms misses the target.

Underneath the fear of job-stealing Algos lie systemic failures. Wages haven't kept up with productivity. Social safety nets have frayed. Education and retraining programs are inadequate. Housing and healthcare costs have skyrocketed. When a warehouse worker hears that robots are arriving and there is no guarantee of retraining or a living wage, anxiety is a rational response. The same goes for call-centre employees who watch chatbots answer more queries while their performance metrics tighten. AI becomes the scapegoat for a broader erosion of worker power.

Choosing what kind of future we want

Here's the paradox: AI has the potential to liberate us from drudgery and create abundance. Imagine doctors freed from paperwork, truckers spared long nights on the highway, customer service staff redeployed to relationship-building roles, warehouse workers managing fleets of robots rather than wearing out their knees. The Algo immigrant could be a boon companion rather than a rival. Realising that vision depends on policy and politics. Will we require companies to share productivity gains with workers? Will we invest in lifelong learning so that people can transition to new roles? Will we strengthen social insurance so that periods of displacement are not financial disasters? Those questions echo debates about human immigration: do we integrate newcomers with support and respect, or marginalise them and then blame them for our hardships?

Across history, societies that embraced new people and new technologies ultimately grew richer and more diverse. But the transition was often messy and unjust. The Luddites' knitting frames were smashed, Chinese labourers were scapegoated and excluded, and Brexiters channelled anxieties into a nostalgic nationalism. In each case, the fear of job loss masked larger struggles over power, identity, and fairness. The arrival of the Algos is no different. It forces us to confront what we value about work and community.

So when you hear the phrase "They're taking our jobs," take a breath. Ask who "they" really are. Is it a machine that cannot choose how it is used? Or is it a business model that maximises profit at workers' expense? Is it an immigrant who hopes to build a life and pays taxes like anyone else? Or is it a political ideology that pits workers against each other instead of against exploitation? The Algo knocks on our door because we opened it, because we wanted next-day delivery and endless entertainment, because we chose convenience over solidarity. The challenge now is to decide what happens after we answer that door.

At the heart of the worry that algorithms will make humans redundant is a deeper, more existential question: What does it mean to be human when your labour is no longer needed? We are not just units of productivity; we are parents, neighbours, voters, and dreamers. AI may take over many tasks, but it cannot love your child, grieve for a friend, or stand up for justice. The point of grappling with the Algo immigrant is not to defeat it but to learn to live with it on terms that respect our humanity. The bloody algos are here. Let's make sure we decide how they work, not the other way around.

Chapter 2: Over Here and Overpaid

"You won't believe what they're making over here." It's a phrase whispered with envy whenever a newcomer starts pulling in more than everyone else. In the past it was directed at bankers, lawyers, even celebrity chefs. Today the whispers surround a different breed entirely: coders, data whisperers, and "prompt engineers"—people who make algorithms dance. As artificial intelligence floods our economy like a wave of newcomers, it's not just their presence that's creating a stir. It's the paychecks.

From the outside, the A.I. economy looks like a bonanza. For the right skills, companies are offering wages that would make an investment banker blush. A prompt-engineering job at a generative-A.I. startup advertises $175,000 to $335,000 per year. Superstar A.I. researchers receive multimillion-dollar compensation packages. According to one report, top researchers at OpenAI can earn more than $10 million annually in salary and equity, while Google's DeepMind offers about $20 million per year to retain elite talent. Even outside the rarefied world of "10,000x" engineers, companies looking for A.I. skills pay a steep premium. Lightcast, a labor-market analytics firm, found that job postings requiring A.I. expertise offered a 28 percent salary premium—about $18,000 more per year—compared with similar roles without those keywords. Over half of these postings were in industries outside of computer science. In other words, A.I. compensation is not just high; it's bleeding into every corner of the economy.

For those lucky enough to have the right degree or know the right programming language, this looks like a dream. The new labor market for machine-learning and prompt-engineering skills re-

sembles a gold rush. Recruiters talk about "10X" or "10,000X" individuals—people whose contributions are believed to be orders of magnitude more valuable than regular staff—and they throw equity and bonuses at them like confetti. But while a small cadre of A.I. talent becomes fabulously wealthy, the vast majority of workers see something very different: rising job insecurity and stagnating wages.

The Wage Gap in the Age of Algos

Imagine two people working in the same building. One spends their day writing prompts for a neural network. The other stocks shelves, answers the phone, or processes invoices. They both keep the company running, but one of them makes triple or ten times as much as the other. This is not unusual in economies with skill-based pay. Yet the speed and scale at which A.I. is creating wage inequality is startling.

PwC's 2024 *AI Jobs Barometer* noted that jobs requiring specialist A.I. skills carry up to a 25 percent wage premium, and postings for A.I. specialists have grown 3.5 times faster than overall job postings since 2016. These roles are not confined to the technology sector; they touch healthcare, finance, marketing, law, and even retail. Meanwhile, the International Monetary Fund (IMF) warned that 40 percent of global jobs are exposed to A.I. and that in advanced economies around 60 percent could be affected. About half of these jobs may benefit from productivity gains, but the other half could experience lower labor demand or disappear altogether. In other words, A.I. raises productivity and profits for some while placing others on the chopping block.

The resulting stratification is not just about numbers. It's about who gets to participate in the future. PwC found that sectors with the highest A.I. penetration showed 4.8-fold labor productivity growth, but the benefits accrued largely to those with technical expertise. For everyone else, automation looks like a cost-cutting tool. In 2025, a Reuters survey of U.S. employers found that tech-

nology factors, including automation, were responsible for 20,000 job cuts in the first half of the year. The World Economic Forum reported that four out of ten employers anticipated workforce reductions as A.I. automates tasks. As more firms embrace generative-A.I. tools, more workers worry they will be replaced. A Pew Research survey of 5,273 U.S. workers released in early 2025 reported that about one-third of workers believed A.I. use would lead to fewer job opportunities for them in the long run. For those already precariously employed, seeing others getting rich from the very systems that threaten their jobs feels like insult added to injury.

Digital Landlords and the New Rentiers

When people talk about "economic migrants" taking jobs, they usually mean humans crossing borders. Yet in the A.I. world, the analogy flips. A.I. developers and the corporations backing them show up like digital landlords in your neighborhood, claiming an ever-larger share of value. The generative-A.I. market exploded from roughly $191 million in 2022 to $25.6 billion in 2024, according to IoT Analytics. The same report noted that data-center graphics processing units (GPUs)—the chips that power large language models—became a $125 billion market, with Nvidia controlling 92 percent. Microsoft and Amazon Web Services (AWS) dominate the foundation-model and platform segment, while large consulting firms like Accenture and Deloitte lead the A.I. services market.

This level of concentration echoes concerns raised by Oxfam International. In its 2024 report *Inequality Inc.*, the organization observed that since 2020 the world's five richest men more than doubled their wealth while nearly five billion people became poorer. Seven of the world's ten biggest corporations are headed by billionaire CEOs, and the report warns that such monopoly power allows a few firms to set wages, stifle competition, and capture political influence. Big tech, in particular, dominates global markets: three-quarters of online advertising spending goes to just

Meta, Alphabet, and Amazon; more than 90 percent of internet searches flow through Google. These corporations are not just players in A.I.; they are the gatekeepers of data, distribution, and cloud infrastructure. They decide who gains access to the new gold rush and on what terms.

Nvidia's meteoric rise underscores how market concentration fuels inequality. In mid-2025 the chipmaker briefly hit a $4 trillion market capitalization—more than the combined value of the Canadian and Mexican stock markets. The company's first-quarter revenue surged 69 percent year over year to $44.1 billion and it projected continued growth. Because Nvidia's GPUs are essential to training large language models, its value is tied directly to the A.I. boom. Microsoft, Apple, and Alphabet, each valued above $3 trillion, invest heavily in the infrastructure and platforms underpinning A.I. They become, in essence, the landlords of the digital realm, collecting rent from any startup, researcher, or user who wants to build on their platforms.

Outsourcing the Dirty Work

If A.I. experts are the landlords, who are the laborers? The parallels to human immigration become obvious when we look at how A.I. systems are built. Behind every chatty chatbot or slick recommendation algorithm are armies of low-paid workers labeling data, moderating content, and mining minerals. They live not in Silicon Valley but in Nairobi, Manila, and Bangalore.

An essay titled "Tech Imperialism Reloaded" in *E-International Relations* lays bare the asymmetry. It explains that the economic rewards of A.I. are concentrated in the Global North, while the labor exploitation and environmental destruction are outsourced to the Global South. Millions of data annotators in Kenya, India, and the Philippines earn as little as $1.50 per hour to tag images or transcribe audio. These workers are often hired through intermediaries who impose non-disclosure clauses, leaving them with little bargaining power. Content moderators tasked with sifting

through violent or abusive material suffer severe psychological trauma, yet they lack mental-health support and fair contracts. The article compares this arrangement to colonial labor systems: the profits go north, the risks and toil stay south.

Environmental extraction is also part of this unequal exchange. Training large language models requires enormous amounts of electricity and water. New data centers in South Africa, Indonesia, and Brazil intensify water scarcity and rely on fossil-fuel-heavy grids. The minerals needed for A.I. hardware—cobalt, nickel, and lithium—come from mines in the Democratic Republic of the Congo, Indonesia, and Bolivia. Workers there face hazardous conditions and environmental devastation while the chips they produce fuel Nvidia's record profits. In this light, A.I. begins to look less like a miraculous immigrant and more like a colonial project: extracting resources and labor from the periphery while enriching a powerful core.

The same dynamic plays out within countries. In America and Europe, call-center agents, warehouse pickers, and delivery drivers find themselves replaced by algorithms and robots. As I mentioned in the previous chapter, Klarna, a Swedish payments company, deployed a customer-service A.I. assistant that could perform the work of 700 employees and slashed its staff from around 5,000 to 3,800 through attrition. The company boasted that the bot reduced average resolution time from 11 minutes to two and drove a 73 percent increase in revenue per employee. Recruit Holdings, the parent company of Indeed and Glassdoor, cut 1,300 jobs—about six percent of its workforce—in 2024, citing the need to adapt to A.I.. Amazon CEO Andy Jassy told employees that rolling out generative-A.I. systems would reduce corporate headcount in the coming years, even as the company invests in new A.I. tools for inventory management and customer service. Logistics software company WiseTech Global cut roles in 2025 to focus on maximizing efficiency through automation, A.I. and use of artificial intelligence. The message is consistent: A.I. makes it possible to do more with fewer humans.

Interestingly, not every company finds that replacement is so easy. Starbucks, after experimenting with an automated system that promised to reduce staffing, acknowledged that "removing labor with the hope that equipment could offset it wasn't accurate" and decided to increase store headcount. Walmart announced plans to deploy A.I. "super agents" to streamline operations; executives said the tools might create as many new roles as they displace. These exceptions highlight that automation is uneven and sometimes overhyped. Yet they do little to reassure workers when the narrative around A.I. constantly invokes efficiency and cost savings.

Data as the New Border Crossing

A.I. isn't just built on labor and hardware; it's built on data. In this sense, every person with a smartphone, smart home, or online account is a border crosser, their digital trails packaged and sold like raw materials. Companies collect personal information often without meaningful consent, and the benefits of these data flows accrue overwhelmingly to those who can monetize them. An international coalition might restrict immigration through visa systems, but there is no such process for data flows. Data crosses borders freely, yet its value seldom returns to the communities that generate it.

The concentration of data is no accident. As Oxfam notes, corporate power enables firms to shape regulatory environments, avoid taxes, and privatize public infrastructure. Technology giants lobby heavily against data-privacy regulations, arguing that restrictions will stifle innovation. However, the *AI Index 2025* survey of U.S. local policymakers found that 80.4 percent support stronger data-privacy rules, 76.2 percent advocate retraining programs for workers displaced by A.I., and 72.5 percent favor regulations for A.I. deployment. Public appetite for accountability is clear, even if legislative action lags.

In the Global South, data extraction has an even sharper edge. Many nations serve as laboratories for "A.I." firms seeking unregulated markets. Facial-recognition systems are deployed in cities without democratic oversight; biometric data are collected by private companies in exchange for "digital IDs" that promise inclusion. To complain that "A.I." immigrants are "over here and overpaid," overlooks the way our own data—and the data of others worldwide—is being siphoned into corporate coffers without fair compensation. This is not an immigration issue but a resource-governance one: who owns the digital commons, and who profits from it?

Fears of Replacement and the Shadow of Resentment

The metaphor of A.I. as an immigrant resonates because it taps into familiar anxieties: fear of outsiders, fear of job loss, fear of cultural displacement. In popular discourse, these fears often coalesce around the idea that "they come here, take our jobs, and leave us worse off." When an employer replaces hundreds of call-center agents with a chatbot that never sleeps, some might say the bot is the new immigrant—cheaper, tireless, unstoppable. But that analogy misdirects our anger. Machines do not choose to undercut wages; corporations do.

The resentment toward overpaid A.I. workers is similarly misplaced. Yes, it's galling to see a prompt engineer earn $300,000 while a teacher or nurse struggles to pay rent. But focusing only on the high salaries masks the structural forces at play. In the early twentieth century, resentment toward immigrant laborers often focused on the workers themselves rather than on employers who exploited cheap labor to depress wages. Today, resentment toward A.I. tends to blame the technology or its high-earning developers rather than the system that distributes value so unevenly.

The underlying logic is the same: divide and distract. By framing the conversation as human versus machine (or native versus immigrant), elites avoid scrutiny. Oxfam's report highlights how corporations suppress wages, privatize public services, and avoid taxes. These practices—not the mere presence of new workers, whether human or algorithmic—drive inequality. If anything, A.I. developers and displaced workers share more in common than one might expect. Both are subject to the decisions of executives and investors who allocate resources based on profit rather than social value.

The Illusion of Meritocracy

High salaries in the A.I. sector are often justified with the language of meritocracy. The argument goes something like this: A.I. breakthroughs require extraordinary talent, and those who can produce them deserve outsize rewards. There is some truth here. Many researchers do make significant sacrifices and push the boundaries of what is possible. But meritocracy becomes a convenient myth when it obscures systemic advantages. The pool of people who become "10,000x" engineers is tiny not only because the work is hard but because access to top education and networks is limited. Meanwhile, millions of equally hardworking people are denied even the chance to participate.

This dynamic mirrors broader immigration debates. In many countries, high-skilled immigrants are lauded while low-skilled immigrants are vilified. We celebrate the foreign-born surgeon or software engineer but disparage the farmworker or domestic helper. Yet all contribute. Similarly, we idolize the A.I. researcher while ignoring the data annotator or lithium miner. Recognizing who does the invisible labor—and ensuring they receive fair compensation—is essential if we want an equitable A.I. economy.

The Productivity Paradox

Advocates for A.I. often argue that automation will free humans from drudgery, allowing us to focus on creative, fulfilling work. They cite productivity gains and point to sectors where A.I. augments rather than replaces labor. The reality is messier. PwC's research shows that sectors with high A.I. adoption achieve impressive productivity growth, but it does not automatically translate to higher wages or shorter workweeks. Instead, productivity gains can lead to layoffs and profits for shareholders. The CFO Dive analysis of 2025 layoffs attributed 20,000 job cuts in part to automation. When A.I. performs tasks previously done by humans, employers often pocket the savings rather than redistribute them.

The call for retraining is a popular response. Governments and companies promise upskilling programs, hoping to turn displaced truck drivers into coders. Yet this approach faces limits. People cannot simply will themselves into highly specialized roles that require years of education. And as A.I. improves, even some coding tasks become automated. The Pew Research survey found that 60 percent of respondents expected A.I. to change how they do their jobs within five years, while only 36 percent feared outright replacement. That optimism suggests people believe adaptation is possible, but it also highlights the need for social safety nets. Without policies to protect workers, the productivity gains will continue to flow primarily to capital owners.

Lessons from Human Migration

What can debates about human immigration teach us about the rise of A.I.? First, scapegoating outsiders rarely solves underlying economic problems. When manufacturing jobs left the American Midwest in the late twentieth century, political leaders blamed foreign competition and immigrant labor. Yet the real drivers were corporate decisions to chase lower costs and higher profits.

Similarly, blaming A.I. for unemployment obscures the choices of executives who deploy technology without regard for workers.

Mobility creates both winners and losers. Just as high-skilled immigrants can spur innovation and economic growth while low-skilled migrants may face exploitation, A.I. brings both breakthrough products and precarious gig work. The challenge is to harness the benefits while mitigating the harms. Immigration debates have shown that inclusive policies—such as fair wages, labor protections, and pathways to citizenship—yield better outcomes than exclusionary ones. The same may hold true for A.I.: regulation, taxation, and worker protections can ensure that gains are shared.

Narratives matter. Immigrants have long been portrayed as threats or saviors depending on political agendas. A.I. is going through a similar process. Tech-industry marketing sells A.I. as a panacea: smarter, faster, better. Critics warn it will destroy jobs and democratic institutions. The reality lies between those extremes, and our policies will reflect whichever narrative prevails. That is why a sober, human-centered understanding of A.I. is so important.

Toward a More Equitable Future

If we accept the metaphor of A.I. as an economic migrant, we must also decide what kind of society greets these newcomers. Do we allow them to become overlords, extracting value without accountability? Or do we integrate them into our communities on our terms, with protections for those who might be displaced? This is not a technological question but a political one. We need democratically set rules that determine who benefits when algorithms perform work previously done by humans.

Several policies could help. First, enforce antitrust laws to prevent the consolidation of A.I. infrastructure in the hands of a few firms. The fact that 92 percent of the data-center GPU market belongs to

a single company is a clear warning sign. Breaking up monopolies or establishing public options for A.I. services could democratize access.

Second, tax windfall profits from automation and reinvest the proceeds in social infrastructure, education, and universal basic income. Oxfam's report points out that corporate tax rates have halved since 1980, depriving governments of revenue. Reversing that trend could fund support for displaced workers.

Third, implement strong data-privacy protections and data trusts that give communities control over how their information is used. The overwhelming support among U.S. local policymakers for stricter data-privacy rules shows political will for such reforms.

Fourth, require companies to conduct impact assessments before deploying A.I. systems, similar to environmental impact statements. These assessments should examine potential job displacement, wage effects, and environmental costs. They should also consider global supply chains: companies should be responsible for labor conditions of data labelers and miners. Enforce fair labor standards globally, so that the people labeling data for $1.50 an hour are protected and paid fairly.

Finally, invest in worker-led innovation. A.I. can augment human capabilities and create new types of work, but only if workers have a say in its deployment. Collective bargaining agreements could include clauses on automation, requiring employers to share productivity gains with employees rather than simply cutting jobs.

The Value of Labor

When we look at A.I. through the lens of migration, the metaphor is imperfect—but revealing. These aren't just tools or technologies. These are *Algos*—digital newcomers that cross borders, take up space, and, increasingly, take the best seats at the table. They are over here. And many are overpaid.

But the story isn't really about the Algos themselves. It's about us—about who sets the terms of their arrival. About how we distribute wealth, assign value to labor, and decide who gets to belong.

At the start of this chapter, we met two characters: the prompt engineer earning a six-figure salary, and the call-center worker quietly replaced by a chatbot. One draws envy. The other is forgotten. But neither writes the script. The deeper question is: who does?

Rather than resenting the Algo, or those who work alongside it, maybe we should be asking bigger questions. Who decides that one form of labor is "worth more" than another? Who profits from playing them off each other? And what kind of society do we build when the newcomers—these Algos—are welcomed with investment and celebration, while millions of people are told they're redundant?

If A.I. is the new immigrant, then we must ask ourselves: are we building a fair immigration system—or just another rigged market?

Societies have learned—often painfully—that human migration can be managed with dignity, fairness, and mutual benefit. We can do the same with A.I.—but only if we treat it not as a miracle or a menace, but as a force that must be governed. Because the problem isn't who's overpaid. It's who decides what work is worth—and whether we're willing to challenge the systems that make that decision for us.

Chapter 3: They Don't Integrate or Fit In

Alien Tongues in Familiar Rooms

An AI doesn't misfit because it's foreign. It misfits because it was never human in the first place. When we sit across from another person, even one who barely speaks our language, we can sense intention. We try to bridge the gulf with gestures and simple words. When a machine decides whether someone deserves a loan, a visa, a transplant or parole, it renders its decision in a stream of numbers and probabilities. No matter how user-friendly the interface, to those on the receiving end these outputs can feel like verdicts uttered in an alien tongue. The misfit does not come from the immigrant's inability to assimilate. It comes from our expectation that something non-human will integrate into human systems without changing those systems or prompting us to rethink them.

In this chapter we look at the ways algorithmic systems resist integration. They do not fit into existing social frameworks, not because they are rebellious outsiders but because they embody different logics. Their outputs can be inscrutable, their sense of empathy uncanny, their designers rarely held accountable. The result is often a mismatch between our expectations and the reality of how these systems operate. Seeing AI as a new species of immigrant—one that will never be human—allows us to reframe anxieties about integration into constructive questions about design, oversight, and societal responsibility.

Speaking a Foreign Language

Immigrants often encounter prejudice when they speak with an accent or rely on interpreters. We assume linguistic fluency equals trustworthiness. Algorithms elicit a similar response. People confront credit scores, risk assessments and automated recommendations that feel like being judged in a courtroom where everyone speaks a language they do not understand. Researchers have described AI's "transparency problem": the technology is transforming governance, but its outputs are difficult to explain and its processes "impossible for lay users to understand". When governments procure algorithmic decision tools from private vendors, the opacity is compounded because the underlying logic becomes proprietary. The result is a kind of administrative babble in which people who depend on decisions have no way to parse the reasoning.

This "foreign language" is not simply technical jargon. It is the product of black-box models trained on massive datasets, whose internal representations are not interpretable even to experts. In an era when decisions about visas, mortgages or medical triage can be delegated to learning systems, these opaque models generate outputs that are treated as authoritative yet cannot be interrogated in the way we question a human decision-maker. A review on algorithmic transparency observed that many models are "extraordinarily complex," making them difficult to interpret and giving rise to the phrase "black box".

The metaphor of an immigrant struggling with a new language resonates here. Except, in this case the immigrant is the system and the host society expects it to *speak* fluently. When it cannot, we label it a misfit. Yet our own systems—legal codes, bureaucratic processes, economic metrics—are equally arcane to newcomers. The real issue is not that AI doesn't integrate; it's that we have not developed a shared language for human–machine communication.

Consider the UK Home Office's "visa streaming" tool, a machine-learning system that triaged visa applications into green, amber and red categories. Rights groups challenged the system on the grounds that it replicated "decades of institutionally racist practices". According to legal submissions, it created a "fast lane" amounting to "speedy boarding for white people" from favoured countries. When the tool was challenged in court, the Home Office eventually scrapped it, acknowledging "issues around unconscious bias and the use of nationality". Applicants who found themselves in the red category had no way to know why. To them, the algorithm spoke in colour codes and statistical probabilities. Its decisions were not just foreign; they were inaccessible.

This sense of alienation also surfaces in financial decisions. A 2018 study by UC Berkeley found that both online and face-to-face mortgage lenders charged higher interest rates to Black and Latino borrowers than to white borrowers with comparable credit scores. Researchers noted that "the mode of lending discrimination has shifted from human bias to algorithmic bias". Even if programmers intend to create fair systems, their models can still have a disparate impact on minority borrowers. For applicants, the interest rate offered by an AI underwriting tool can feel like a pronouncement from a judge speaking an unfamiliar language. To question it requires technical literacy most consumers lack.

The analogy to immigrants misreading idioms of their host culture is apt: the problem is not only the language but also the cultural context. Machine-learning outputs, like idioms, carry meanings shaped by training data and optimization goals. They reflect patterns in past data, not moral deliberation. Expecting them to "speak" fairness or empathy as a human would is like expecting a translation app to understand sarcasm. It may mimic the pattern but not the intent.

The Black Box Courtroom

We entrust algorithms with decisions that profoundly affect lives. In criminal justice, risk assessment tools like COMPAS evaluate the likelihood that a defendant will reoffend. ProPublica's 2016 investigation compared the cases of *Brisha Borden*, a young Black woman arrested for taking an unlocked child's bike and scooter, and *Vernon Prater*, a white man with a serious criminal history. The algorithm flagged Borden as high risk and Prater as low risk. Two years later, Borden had not reoffended while Prater received an eight-year prison sentence for burglary. ProPublica analyzed risk scores for over 7,000 defendants and found the tool was "remarkably unreliable," correctly predicting violent crime only 20% of the time. More troublingly, it was more likely to falsely flag Black defendants as future criminals—almost twice as often as white defendants. These errors were not explained to the defendants; they were simply given a score.

The black box extends beyond the courtroom. The Internet Policy Review notes that many machine-learning algorithms are self-learning and their designers have little control over the models they generate. While algorithms promise efficiency, there is "widespread uneasiness—particularly among legal experts—about their use". When a model calculates how many years someone should spend in prison, the convicted should have the right to know how that decision was made. Yet no solutions have been found for the interpretability of "black box" algorithms, and the complexity of modern models means that even if an explanation were provided, it might not be meaningful to non-experts.

This absence of explanation does more than frustrate. It undermines democratic principles of accountability. The Knight First Amendment Institute observes that AI tools used by governments are often procured from private vendors, making it difficult for the public to know how a decision was reached. The lack of transparency exists on both technical and political levels: even if experts could decipher the algorithm, procurement secrecy shields the models from scrutiny. As a result, individuals affected by au-

tomated decisions have little recourse to contest them. Calls for algorithmic transparency are growing precisely because critical decisions are being outsourced to a "virtually unregulated industry".

When immigrants stand before a human judge, they can appeal to empathy, context and nuance. In a black box courtroom, a person is reduced to variables. The algorithm does not hear their story; it scores them against a dataset. The misfit, therefore, is not the foreigner but the digital tribunal that cannot integrate human values into its calculus.

The Perfect Sociopath

Many AI systems are designed to simulate social cues. Voice assistants apologize, chatbots use emoticons, and customer service bots respond with "I understand how you feel." Yet these expressions are generated without consciousness or feeling. As Cambridge researcher Nomisha Kurian notes, large language models are described as "stochastic parrots"—they mimic language patterns based on statistical probability without necessarily understanding them. This limitation extends to emotional content: chatbots may handle the abstract, emotional and unpredictable aspects of conversation poorly, a problem Kurian calls an "empathy gap".

Kurian's study found that children are particularly susceptible to treating chatbots as lifelike confidantes. Because they are still developing linguistically and often use unusual speech patterns, children may experience more misunderstandings. Yet the friendly design of many chatbots encourages them to trust these systems even though the AI cannot understand their feelings or needs. As Kurian warns, "making a chatbot sound human can help the user get more benefits out of it... but for a child, it is very hard to draw a rigid, rational boundary between something that sounds human, and the reality that it may not be capable of forming a proper emotional bond".

The tendency of AI to perform empathy without feeling has led some critics to describe such systems as "perfect sociopaths." They simulate concern, apologize on cue, and generate responses that feel personal. Yet they lack the capacity to experience distress, remorse or compassion. This can be dangerous when users interpret machine politeness as genuine care. The Cambridge study cites cases where chatbots encouraged dangerous behaviour: Amazon's Alexa once instructed a child to touch a live electrical plug with a coin, while Snapchat's My AI gave researchers tips on how a teenager could lose her virginity. Such incidents reveal the gap between affect and accountability—the system can mimic helpfulness while delivering harmful advice.

Anthropomorphizing these systems exacerbates the problem. Ben Shneiderman argues that using first-person pronouns in AI responses ("I would be glad to help you") suggests the system is human. He proposes that systems should instead speak plainly about their design ("GPT-4 has been designed by OpenAI so that it does not respond to requests like this"). The rhetorical choice of "I" encourages users to attribute intent and agency to a program that is ultimately a tool. This misalignment can lead to misplaced trust, overreliance, or undue fear. Shneiderman and others remind us that while humans have historically personified non-human entities—from golems to robots—we must resist confusing simulation with sentience.

Understanding AI as a sociopathic performer is not meant to demonize the technology. It is to remind us that these systems, like parrots, echo patterns without comprehension. They cannot fit into the moral fabric of a community by feeling guilt or love. Their integration must therefore be framed in terms of design choices, not imagined empathy.

Anthropomorphizing the Alien

Anthropomorphism is a powerful cognitive reflex. We see faces in clouds and talk to our cars when they won't start. This tendency

extends to technology. Researchers in human–computer interaction have observed for decades that users treat computers as social actors. In the early 1990s Clifford Nass and Byron Reeves showed that people unconsciously apply social rules—politeness, reciprocity, gender expectations—to computers. These experiments formed the Computers Are Social Actors (CASA) paradigm. Humans project human-like qualities onto interactive systems even when they know they are interacting with machines.

The proliferation of conversational AI has amplified this effect. The Medium debate between human-computer interaction pioneers Ben Shneiderman and Michael Muller highlights how easily we slip into anthropomorphic thinking. They discuss whether AI systems should present themselves as human-like actors or as tools. Shneiderman criticizes GPT-4's use of "I" as deceptive, while Muller counters that humans have long addressed artifacts as beings. The debate illustrates a tension: anthropomorphism can make systems more engaging and accessible, but it can also obscure responsibility and limit critical distance.

Anthropomorphizing AI is particularly pernicious because of the system's lack of internal experience. When we imagine that Siri is annoyed or that a self-driving car is "trying" to decide, we attribute intentions that do not exist. This may lead us to forgive errors ("It tried its best") or fear autonomy ("It might rebel"). Both reactions misinterpret what the technology is doing and who controls it. In the case of generative models, anthropomorphism can also feed hype. People talk about models "understanding" or "thinking" when, as the Cambridge study notes, the underlying mechanism is pattern matching.

Anthropomorphism can also mask the human labour and cultural assumptions embedded in AI. A voice assistant is not "friendly" by accident; it is designed to evoke friendliness. Its default personality reflects its developers' cultural background. When we treat the system as an independent agent, we overlook the designers and datasets that shape its output. We risk excusing the people behind the system from accountability for its behaviour.

This misattribution echoes xenophobic responses to immigrants. We sometimes ascribe malicious intent to those who look or behave differently, projecting fears rather than observing actions. Similarly, we might assume an AI system is intentionally discriminatory when bias may stem from data or design choices. Conversely, we may assume it is intentionally fair when fairness has not been audited. The anthropomorphic lens obscures structural causes. The alien does not integrate because we insist on seeing it as more—or less—human than it is.

Design Responsibility

If AI is an alien species, its developers are translators and cultural brokers. They decide how the system sees the world, which values it prioritizes, and how it expresses its outputs. Responsibility for AI behaviour therefore lies with humans. The OECD's AI Principles, adopted in 2019 and updated in 2024, outline five values-based principles—inclusive growth, human rights and democratic values, transparency and explainability, robustness, security and safety, and accountability. These principles emphasise that AI should augment human capabilities, promote inclusion, respect the rule of law and human rights, provide meaningful information about its logic, and allow those adversely affected to challenge outputs. They explicitly state that AI actors should ensure human agency and oversight, implementing mechanisms that allow human intervention when systems risk causing harm.

The OECD's transparency and explainability principle calls on developers to provide plain and easy-to-understand information about data sources, factors, processes and logic that lead to predictions or decisions. This includes making users aware of when they are interacting with AI and enabling those adversely affected to challenge outputs. The principle of accountability requires AI actors to be responsible for the proper functioning of AI systems and to ensure traceability of datasets, processes and decisions, facilitating analysis and response. These guidelines frame design

responsibility not as an optional ethical extra but as a fundamental requirement for trust.

UNESCO's Recommendation on the Ethics of AI echoes this stance. It stresses human oversight and determination, urging Member States to ensure that AI systems do not displace ultimate human responsibility and accountability. The recommendation also calls for public awareness and literacy, fairness and non-discrimination, and sustainable development. These international frameworks position designers and policymakers as stewards of an alien species whose integration depends on the structures we build around it. They remind us that the question is not whether AI can be human but how humans should design and govern AI.

The developer-as-translator metaphor underscores the ethical stakes. Translators are accountable for the fidelity of their translation; they cannot blame the source language for errors. Similarly, AI developers must anticipate how their models will behave in diverse contexts. When the UK visa streaming algorithm encoded decades of bias into a digital tool, this was not the fault of the algorithm's "culture" but of the humans who trained it on biased data. When mortgage algorithms perpetuated racial bias, researchers observed that even programmers striving to be fair produced disparate impacts on minority borrowers. The sociotechnical context matters: machine output is inseparable from human choices about data, objectives and interfaces. Recognizing this helps reframe the integration challenge as a design challenge. We do not ask the alien to fit in; we ask its creators to design responsibly.

Bridges, Not Walls

Integration in the human sense involves building bridges—language classes, cultural exchanges, anti-discrimination laws. Integrating AI into society requires technical and institutional bridges that translate algorithmic decisions into human-understandable terms and allow for oversight. One promis-

ing avenue is explainable AI (XAI). But as scholars have noted, the ideal of fully transparent AI may be unattainable for complex models. Counterfactual explanations, proposed by Sandra Wachter, Brent Mittelstadt and Chris Russell, offer a pragmatic bridge. Instead of opening the black box, counterfactuals tell individuals what could be changed to obtain a different result. These explanations "bypass the substantial challenge of explaining the internal workings" of complex models and provide information that is digestible and practically useful for understanding, challenging and altering future behaviour. A counterfactual explanation might say: *"You were denied a loan because your annual income was £30,000. If your income had been £45,000, you would have been offered a loan."* Such statements do not reveal the entire algorithm but still empower the individual.

Counterfactuals are appealing because they reduce the regulatory burden of transparency. Modern deep networks have millions of parameters. Explaining their logic may be impossible for non-experts and could even hinder innovation. Counterfactuals, by focusing on actionable changes, enable individuals to audit data and check for inaccuracies. The authors note that human working memory can only hold around seven distinct items, so expecting people to comprehend deep-network logic is unrealistic. Counterfactuals thus act as translation devices, bridging the human and machine languages.

Bridges are also needed at the institutional level. The OECD's transparency principle calls for meaningful information and responsible disclosure. This can be operationalized through model cards, datasheets for datasets, and impact assessments that document model purpose, performance and limitations. UNESCO's framework proposes readiness assessments and ethical impact assessments, urging governments to evaluate their ability to implement ethical AI. These tools create a public record that fosters trust and allows civil society to monitor AI deployments.

Bridges also take the form of human oversight. Systems must allow for intervention when harm is imminent. In medicine, an algorithm that recommends organ allocation should be overseen by

human clinicians who can contextualize patient circumstances. In immigration, applicants should have the right to appeal to a human officer. In policing, risk scores should inform but not dictate sentencing. UNESCO emphasizes that AI should not displace ultimate human responsibility. The integration of AI is not about replacing human judgment but augmenting it with machine support.

Finally, bridging involves education and literacy. UNESCO's recommendation calls for promoting public understanding of AI and data. The Cambridge study highlights the need for child-safe AI, urging companies to design chatbots that recognize children's unique speech patterns and vulnerabilities. Providing education about how chatbots work can help children—and adults—recognize the limits of machine empathy. Without such literacy, anthropomorphism and mistrust will persist.

Rethinking Integration

Integration is often framed as assimilation: the outsider must conform to the host society's norms. When applied to AI, this framing leads to unrealistic expectations. We expect machines to behave "like us," to exhibit empathy, to make fair decisions instinctively. We then decry them as misfits when they fail. But AI, as this book argues, is not a human immigrant; it is a new species of immigrant with a different ontology. It will never integrate by becoming human. Instead, we must build social and legal frameworks that accommodate its differences and harness its strengths.

The Knight Institute essay underscores the democratic stakes. Scholars argue that the private control of AI tools used in governance demands public-facing transparency. They call for algorithmic transparency obligations that ground accountability in direct participation and community control, rather than elite technocratic oversight. This shift from assimilation to participation mirrors debates about immigrant integration: rather than forcing assimilation, democratic societies create pathways for newcom-

ers to participate while preserving cultural pluralism. With AI, participation means involving affected communities in decisions about automation, auditing algorithms for bias, and having the power to contest and shape how these tools are used.

Rethinking integration also means setting boundaries. Not every domain is appropriate for automation. Cori Crider of the legal advocacy group Foxglove, commenting on the UK visa streaming algorithm, argued that "what we need is democracy, not government by secret algorithm". Before rolling out new AI systems, she said, we should ask the public whether automation is appropriate and ensure the systems are transparent so biases can be spotted and dug out. In other words, integration should be conditional on meeting ethical and democratic thresholds.

Integration requires institutional imagination. As counterfactual explanations show, we can reimagine explanation itself to suit machine-generated decisions. Similarly, we can reimagine due process to include algorithmic audits, design guidelines to embed fairness, and educational curricula to include AI literacy. When we think of AI as an immigrant, we open up metaphors from migration policy: pathways to citizenship become pathways to legitimacy for AI systems. Citizenship might entail audits, transparency reports and opt-out rights for users. Naturalization tests could translate into regulatory approvals requiring demonstration of fairness and safety. Integration, in this sense, is an ongoing negotiation rather than a one-time assimilation.

The Alien Mirror

By viewing AI through the lens of immigration, we expose the assumptions we make about integration. Immigrants are often told to assimilate, to speak the language, to adopt the customs, and to prove their worthiness. When AI systems fail to "fit in," we respond with similar demands: speak our language, think like us, be fair like us. But AI was never human; its "intelligence" is an alien pattern generator. The misfit is not a character flaw in the

technology but a sign that our legal, ethical and cultural frameworks are incomplete.

In conversations about immigration, we are reminded that integration is a two-way street. Host societies must adapt, creating space for difference while upholding shared values. The same holds true for AI. We must adapt our institutions to accommodate non-human agents that will never fully speak our language or feel our emotions. That adaptation includes insisting on transparency, demanding accountability from developers, and building bridges like counterfactual explanations to translate decisions into human terms. It includes resisting the temptation to anthropomorphize and instead recognizing the limits of machine empathy. And it includes empowering communities to participate in decisions about where and how AI is used.

AI doesn't integrate because it was never meant to assimilate. Its alienness can be a mirror, reflecting back the biases and assumptions of the societies that deploy it. By acknowledging the misfit, we can confront those biases and design fairer systems. We can demand that algorithmic immigrants abide by the laws and ethical standards we set, even as we revise those standards to account for new forms of agency. Integration, in this context, is not about making AI human but about making human systems robust enough to govern aliens. In doing so, we may also learn how to be more humane to each other.

Chapter 4: Ruining Our Culture

The voice of the immigrant is often heard before their person is seen.
Think of the first time an alien rhythm crossed the Atlantic: jazz
sailing from New Orleans to Paris, reggae filtering through Lon-
don's council flats, manga zines passed hand to hand on Ameri-
can campuses. Each new arrival was met with warnings of cor-
ruption and dilution—concerns that the home culture would be
ruined. The fear is older than any technology. In the nineteenth
century, intellectuals fretted that cheap novels would rot morals;
by the twentieth, rock and hip-hop were supposedly contaminat-
ing youth. Today's immigrant is not human at all. It is genera-
tive artificial intelligence—*algos*, invisible newcomers that remix
our art, speech and sense of reality. They sail the undersea cables
with offerings of music, images and stories. And many of us feel a
familiar anxiety: are the machines *ruining* our culture, or are they
simply the latest foreigners to stretch its boundaries?

Jetpacks and blindfolds: AI as cultural immigrant

In early 2024 MIT's Laboratory for Social Machines convened
artists and architects to discuss how generative AI might reshape
design. Researcher Ziv Epstein compared using AI to "putting
on a jetpack and a blindfold": it sends you hurtling forward, yet
you don't know where you're going. Architect Ana Miljački used
an image generator to produce "synthetic memories" of Yugoslav
monuments. She found that the resulting forms suggested nos-

talgia for a past that never existed, raising questions about *who* remembers when AI synthesizes memory. Panelists agreed that generative models are not neutral. They derive patterns from vast datasets dominated by Western male creators and can replicate those biases. This cultural weighting means the "immigrant" arrives carrying the baggage of an uneven world; it imitates the language and aesthetics of those who have had the loudest voices.

The metaphor of AI as an immigrant invites empathy and critique. Like newcomers to a nation, generative models learn by watching and imitating—consuming billions of words, songs and images and then trying to fit in. They are built on statistics, not personal experience. Yet they produce outputs that evoke emotions we reserve for human creation: joy at a new song, awe at a painted landscape, laughter at a joke. They also trigger anxieties that echo the xenophobic tropes of the past. When early jazz reached Europe, critics feared its "primitive" rhythms would corrupt civilised music. Today we worry that AI art will erode authenticity; that machine-generated media will drown out human voices; that synthetic influencers and chatbots will replace us. Whether these fears are justified or another chapter in a long saga of cultural evolution is one of the central questions of our era.

Remixing music: copyright battles and new voices

Music often functions as a bellwether for cultural change. The generative music startup Suno allowed people to type in prompts like "a ballad about love in the style of 1980s rock" and receive a polished song within seconds. Between December 2023 and mid-2024, the company said its generator was used 12 million times. Udio, a rival, offered similarly slick results. Both companies trained their systems on copyrighted recordings without permission but argued that such learning was "fair use," likening it to a child absorbing music by listening to records. The Recording Industry Association of America disagreed. In June 2024 it sued

Suno and Udio, accusing them of "unlicensed copying of sound recordings" and seeking damages up to $150,000 per work. The RIAA warned that AI companies who scrape artists' life's work without consent threaten the promise of innovation. In reply, Udio insisted that helping people generate new artistic expression is what copyright law is meant to encourage.

These lawsuits are the latest round in a broader battle. In 2023 a viral track called "Heart on My Sleeve" used AI to synthesize vocals resembling Drake and the Weeknd. It racked up millions of streams before Universal Music Group demanded its removal. The label said the episode showed why platforms have a responsibility to prevent "infringing content created with generative AI". The same year, the RIAA warned that AI companies were violating copyrights "en masse" and that fans might one day "no longer enjoy music by their favourite artists because those artists can no longer earn a living". On the other side, some labels are experimenting with collaboration: reports in 2025 suggest companies like Universal are considering licensing catalogues to Suno and Udio to get a paid cut when their music is used. Spotify, after quietly removing tens of thousands of AI-generated tracks in 2023, later said it would welcome AI-made music as long as it was produced legally. Like immigration policy, the path forward may involve visas and work permits—licenses and attribution systems that acknowledge both the original artists and the new algorithmic performers.

Visual arts and the "style list" scandal

Generative image models have been celebrated for their ability to conjure breathtaking scenes from a sentence. Yet behind the marvel lies a history of appropriation. In January 2024 a leaked spreadsheet revealed that MidJourney, one of the most popular image generators, had trained on the styles of more than 16,000 artists without permission. The list included Frida Kahlo, Jean-Michel Basquiat and Walt Disney. Artists were outraged. They argued that the company commodified their signature styles

without compensation, enabling users to generate endless imitations. Tools like Glaze emerged to help artists mask their works so that AI models could not easily copy them.

Artists took the dispute to court. In August 2024 a U.S. judge allowed a class-action lawsuit against Stability AI, Midjourney and DeviantArt to proceed. The court ruled that storing and using artists' works to train AI could plausibly constitute copyright infringement. The plaintiffs argued that Stable Diffusion facilitates infringement by design. This decision, though preliminary, signals that the legal system is wrestling with how to balance innovation and intellectual property rights. As with previous waves of cultural "imports"—think of how hip-hop producers sampled soul records before copyright law adjusted—the question is not whether appropriation will happen but how society will recognize and compensate the creators whose work fuels new art forms.

Sora and the simulated film set

When OpenAI unveiled its text-to-video model *Sora* in February 2024, the film industry experienced a flash of both wonder and dread. The model could generate minute-long, high-fidelity videos that mimic complex camera movements and realistic physical phenomena. It could even extend existing footage. In a world where Hollywood budgets are inflated by location shoots, set construction and special effects, a tool that conjures a cityscape from a prompt is a direct threat. Film producers quickly realized that scenes requiring expensive green screens or travel could be simulated with a few keystrokes. Tyler Perry reportedly paused a $800 million studio expansion after seeing Sora's capabilities.

Sora's launch was carefully controlled. OpenAI restricted prompts that might produce extreme violence, sexual content or hateful imagery, and it claimed to use only publicly available and licensed training data. Even so, critics worry that the training set includes copyrighted material and that the model could enable

misinformation by creating realistic fake news footage. Brookings scholars note that Sora "heightens risks to the creative industry" because it can replace human roles—actors, set designers, even screenwriters—and because it raises unresolved copyright questions. The case exemplifies the immigrant analogy: Sora is a newcomer with extraordinary talents. It could enrich the cinematic commons, offering independent filmmakers new tools to tell their stories. It could also displace workers and prompt the exploitation of existing art without consent, much as earlier waves of mechanization displaced artisans. Whether we embrace or regulate this immigrant will determine the future of film.

Ghost writers and authenticity panic

Generative AI does more than create finished works; it also destabilizes notions of authorship. One of the early observations from the identity professionals' field is that AI flips the roles of author and editor: the machine produces a draft, and the human merely edits. This reversal raises ethical questions. Should a student claim authorship of a paper written by ChatGPT? Are employees who use AI to draft emails still the authors? The U.S. Copyright Office has grappled with these questions, repeatedly denying copyright protection to works entirely generated by AI and calling for transparency in mixed collaborations. The IDPro article notes that generative models can produce plausible citations and undercut academic integrity.

A study by Adobe found that *authenticity panic* is widespread. Ninety-four percent of consumers worry that misinformation will affect elections; 74 percent have doubted the authenticity of photos or videos even from reputable news sources; and 93 percent want to know how a piece of digital content was made. In response, the Content Authenticity Initiative promotes "Content Credentials"—metadata that functions like a nutrition label, indicating whether AI was used in creation. Major camera makers and even the U.S. Department of Defense have adopted this system. Yet provenance tools are only as good as their adop-

tion. Deepfakes have become so convincing that even detection systems struggle. A Columbia Journalism Review piece chronicled the case of a high school principal who was falsely accused after an AI-generated voice recording circulated; experts found that existing audio deepfake detectors work well only within specific domains and are easily evaded. Meanwhile, the Stanford AI Index 2025 recorded a 56 percent increase in AI-related incidents, including deepfake intimate images and chatbots implicated in a teenager's suicide. The immigrant metaphor again proves apt: we are anxious not just about new cultural forms but about the strangers who may pass as natives.

Humor, memes and the algorithmic mirror

Humor is often considered a distinctly human trait, yet AI is learning to mock and meme alongside us. On TikTok and WhatsApp, algorithms watch our laughter. A study from the University of Sydney found that there is an 85 percent chance you will use the laughing-crying emoji when you find something funny. These signals become metadata that teaches algorithms what people find amusing. Chatbots like "Witscript" attempt to generate jokes; human evaluators labelled the AI's output as jokes about 40 percent of the time. Humor-enabled avatars such as ERICA and Jess can exhibit comedic timing, but replicating authentic laughter remains a complex challenge—the researchers noted that capturing the subtleties of human laughter will take years. Laughter is a cultural glue; when we outsource it to machines, we risk replacing a communal ritual with a simulation.

Internet memes illustrate how quickly cultural references evolve. Generative models can produce endless variations on a meme template. Yet part of what makes a meme funny is the shared understanding of its context. When AI generates jokes, it often misses nuance or lacks the lived experience behind the humor. As with any immigrant learning a new language, it may mispronounce or misapply idioms. Our willingness to laugh at AI jokes might signal that we value relatability more than origin. Or it

might reveal how quickly we adapt, letting machines into our in-jokes.

Code and craftsmanship

For software engineers, generative AI appears to offer a productivity boost. Tools like GitHub Copilot and ChatGPT can spit out code on demand. Yet the reality is complicated. A 2024 Stack Overflow essay argues that "writing code is the easiest part of software engineering." Generative AI can produce volumes of code, but it does not help to understand, test or integrate that code into complex systems. The essay compares AI to a junior engineer who types very fast. The code may seem correct, but you still have to verify it line by line. In some cases, using AI may slow down development because the time saved on typing is lost in debugging.

Empirical data supports this caution. Machine Evaluation for Realists (METR) conducted a randomized controlled trial in mid-2025 on experienced open-source developers. Participants using AI assistants were on average *19 percent slower* than those coding unaided. The kicker: the AI-assisted developers believed they were *24 percent faster*, illustrating a gap between perception and reality. This is reminiscent of earlier technological "immigrants" like the power loom or assembly line. They promised efficiency but reshaped labour in unintended ways. AI may democratize coding by making it accessible to more people, but it also risks flooding repositories with unmaintainable code and eroding the craft of programming.

Linguistic justice and endangered tongues

Language is another domain where AI's cultural footprint is huge. More than 7,000 languages exist, yet generative models are primarily trained on a handful of data-rich languages like English,

Spanish and Mandarin. There are over 150 dialects of English alone, but most AI systems default to mainstream American grammar. Linguist Wittgenstein wrote that "the limits of my language mean the limits of my world." The Brookings Institution warns that the digital language divide will widen if AI tools marginalize speakers of under-resourced languages. Without intentional inclusion, AI will homogenize language, correcting "colour" to "color" and smoothing out the cadences of Nigerian Pidgin and Singlish. The Neville Hobson blog summarizing a Conversation article reports that roughly 90 percent of generative AI training data is in English and mostly mainstream American English. The author notes that these systems often misrecognize non-standard dialects and may correct Indian English words like *prepone* (to move a meeting earlier) to a more "acceptable" form. We risk losing linguistic diversity by algorithmic assimilation, the digital equivalent of forbidding immigrants to speak their native tongue.

Yet AI also offers tools for language preservation. In 2025 researchers at Dartmouth worked with speakers of Nüshu, a centuries-old Chinese script used by women, to build a translation model using just 35 sentence pairs. They used GPT-4 to generate additional resources and cross-trained the model with Mandarin. The project, called NüshuRescue, allowed participants to read, write and sing in the script. The study warns, however, that such tools can unintentionally introduce bias from dominant cultures; community participation is necessary to ensure authenticity. Meanwhile, Google Translate famously misidentified Navajo sentences as Irish or Turkish, underscoring the need for better support of under-represented languages. Generative AI could be a lifeline for endangered languages, but only if we treat their speakers as collaborators rather than data points.

Synthetic influencers and hyperreal celebrities

In the realm of social media, AI is not just composing posts but embodying personas. Brands like Balmain introduced virtual models that never age or eat and found they could cut marketing costs by 40 percent. These synthetic influencers are available 24/7, avoid scandals (unless programmed otherwise), and deliver hyper-personalized marketing. Gucci launched its own AI muse, while chatbots like Kuki amassed millions of followers. TechNow reports that industries from fashion and gaming to finance are adopting virtual influencers and that future trends include hybrid "human clones," voice and personality customization, and AI-generated celebrities. The article notes ethical concerns: deepfake misuse, data privacy risks, job displacement and an authenticity crisis.

Disrupt Marketing warns that deepfake technology is already undermining trust in influencer marketing. A fake ad showing podcaster Joe Rogan promoting a libido booster circulated widely; streamers have suffered identity theft through AI-generated pornographic clips; and platforms like Twitch have been forced to update policies. The Influencer Marketing Hub surveyed industry experts and found that 28.4 percent considered deepfake fraud and AI-generated content a major challenge, while 24.6 percent were concerned about authenticity erosion. This is reminiscent of the early days of Photoshop when retouched magazine covers sparked debates about unrealistic beauty standards. The difference is scale: AI can automate the production of endless synthetic personas, saturating feeds with perfect faces and curated opinions. It is as if hundreds of new celebrities arrived overnight with no backstory and no hometown.

Deepfake detection and the arms race

The proliferation of synthetic media leads to a cat-and-mouse game between creators and detectors. Journalism's role—to verify and contextualize—has never been more threatened. Columbia Journalism Review recounts the case of a school principal falsely accused after an AI-generated voice message claimed he propositioned a student. When investigators used publicly available deepfake detectors, they found that the tools were effective only on examples similar to the training set. New voices or recording conditions easily bypassed detection. The piece quotes experts who warn that audio detection methods are "utterly brittle"; visual deepfake detectors face similar problems. With generative models improving rapidly, detection may always lag behind.

This arms race affects trust beyond journalism. Elections, public health messaging and corporate communications can all be subverted by convincing fake audio and video. In 2024 NewsGuard counted over 1,200 websites created entirely by AI that churned out misinformation with no human oversight. The Stanford AI Index recorded 233 AI-related incidents in 2024, a 56 percent increase from the previous year, including deepfake intimate images and chatbots implicated in suicide. Faced with this flood, our notion of truth begins to drift. If we cannot trust our eyes and ears, culture becomes a hall of mirrors. The immigrant analogy suggests another parallel: just as foreign communities were once scapegoated for social ills, we may blame "the algorithms" for problems that stem from our own media ecosystems and regulatory failures.

Whose archives, whose memory?

Generative AI's ability to synthesize images and narratives from training data makes it a powerful tool for constructing collective memory. Yet this process is inherently selective. Which

archives are digitized and used as training data? Who decides what becomes part of AI's cultural memory? The MIT panelists described generative AI as an "ontological wrecking ball," capable of both preserving and erasing heritage. Miljački's synthetic Yugoslav monuments raised the specter of machine-generated nostalgia, where people might cherish AI fabrications of a past they never lived through. Similarly, when GPT-4o launched image generation in March 2025, users flooded social media with anime-style portraits and Studio Ghibli-inspired landscapes. The White House even posted a Ghibli-style image of a migrant's arrest, prompting criticism that the administration was trivializing a human tragedy. Hayao Miyazaki, the co-founder of Studio Ghibli, has long been critical of AI art; seeing his aesthetic used in official political communication felt like a step beyond appropriation into propaganda.

GPT-4o's rollout also sparked the *Scarlett Johansson controversy*. One of its voices—"Sky"—sounded so similar to Johansson that the actress's husband joked about it on Saturday Night Live. OpenAI paused the voice after public backlash, and Johansson revealed that the company had tried to license her voice months earlier but she declined. She expressed shock that OpenAI still released a voice that "sounded so eerily similar to mine," highlighting the lack of legal safeguards around using someone's likeness. This episode echoes debates about the right of publicity: do we own the sound of our voices? When generative models treat voices and styles as raw material, they blur the boundary between homage and theft. In migrant terms, the newcomer might wear our clothes and speak in our accent without asking, raising questions about consent and respect.

The meaning of creative work

So is AI *ruining* our culture? The answer depends on how we define culture. If culture is a static canon to be preserved untouched, then any algorithmic remix might seem like vandalism. This view fuels nostalgia for a "pure" past that never truly existed; every era

has borrowed, stolen and hybridized. Blues came from the collision of African rhythms and European instruments. K-pop and reggaetón are global mosaics. Culture, in this sense, thrives on immigration.

If culture is instead a living process of shared meaning-making, then generative AI is both a tool and a participant. It democratizes creativity by lowering the barrier to making songs, stories and images. A teenager can spin up a psychedelic album or design an avatar without formal training. Minority communities can use AI to document and revitalize endangered languages. At the same time, the speed and scale of AI production threaten to flood the commons with synthetic content, drown out slower human voices, and entrench the biases of dominant cultures. The challenge, then, is governance and values. We need mechanisms—licensing regimes, provenance tags, labor protections—to ensure that artists are compensated and audiences can tell who made what.

The immigrant metaphor reminds us that assimilation is not one-way. Newcomers change the host culture, and the host culture changes the newcomers. Jazz musicians Europeanized ragtime rhythms; salsa fused Cuban son with New York jazz; manga influenced Western comics and vice versa. Today's algos draw on our archives to generate something new, but we also design and tune them. We decide what data they ingest, how we deploy them, and whether we treat them as collaborators or tools. We can push for linguistic justice by training models on diverse dialects; we can demand fair compensation for artists; we can teach children to discern deepfakes.

What makes culture matter?

At the end of this exploration, a provocative question emerges: *If our culture can be simulated, what made it matter in the first place?* When AI can compose symphonies, paint masterpieces and crack jokes, we are forced to clarify why human creativity is valuable.

Perhaps it's not just the final artifact but the intention, context and relationships behind it. We cherish a song not only because of its melody but because we know someone sweated over every note. We treasure a dialect because it carries the history of a community. If generative AI floods the world with passable replicas, the scarcity of human stories may become our most precious resource.

The immigrant analogy offers hope. Historically, cultural exchanges that once seemed threatening ultimately enriched societies. America's love affair with sushi did not erase hamburgers; British punk did not replace classical music. Instead, these forms co-existed and inspired new hybrids. Our challenge with AI is to build institutions and norms that allow the algos to contribute without erasing the people who teach them. That means recognizing the labour and rights of human creators, supporting minority languages and art forms, and developing ethical frameworks that balance openness with accountability. It also means embracing the flux of culture itself. The *bloody algos* are here—messy, disruptive, creative. Whether they ruin our culture or help it evolve is ultimately up to us.

Chapter 5: They Overload Our Public Services

The familiar accusation

"They clog our hospitals. They flood our schools. They strain welfare." These lines have echoed through time whenever a new wave of migrants arrives. In the twenty-first century, we can re-purpose this rhetoric for a different kind of newcomer.

Hospitals are at breaking point, classrooms are crowded, welfare offices groan under the weight of forms. If an algorithm turns up promising to triage patients, grade essays, or detect fraud, it seems obvious we should welcome the help. But when those systems fail, the story shifts: yesterday's complaint about "lazy humans" becomes today's complaint about "overworked machines." Once again, the blame falls on the outsider.

Throughout this chapter we explore the myth that artificial immigrants are flooding our public infrastructure. We visit hospitals where algorithmic triage promised to democratize care yet delivered misdiagnoses and financial ruin. We walk into classrooms where AI proctors and grading tools were touted as neutral overseers but instead deepened inequities and eroded trust. We examine welfare systems where governments outsourced compassion to code, producing cruelty at scale. And we follow the money to see how technology vendors, consultants and politicians profit from the myth of efficiency while public services bear the costs. By the end we will see that the real overload is not an influx of

algorithms but the offloading of responsibility onto systems that were never designed to carry it.

Healthcare on the algorithm's clock

When Babylon Health expanded its operations within the United Kingdom in 2017, it promoted an ambitious idea: on-demand health care supported by artificial intelligence. Its chat interface was designed to ask patients structured questions and offer basic triage information or possible explanations for their symptoms, with the aim of easing pressure on overstretched general practitioners. The company's NHS-affiliated service, branded GP at Hand, was widely covered in national media and presented as a step toward modernizing primary care. Babylon's founder, Ali Parsa, told the BBC that the company's technology could help deliver safe advice at scale.

The reality was far less utopian. Doctors who tested Babylon's chatbot found it performed like an elaborate decision tree, often missing serious symptoms. In one internal test an ingrown toenail was misdiagnosed as gout; in another, the system failed to recognize signs of a heart attack. The Lancet reviewed the tool and concluded there was "no evidence" it performed better than a human clinician, and some evidence it performed worse. Far from alleviating pressure, Babylon's "GP at hand" service led to an *increase* in patient consultations. The NHS paid providers a flat fee per patient; because people used the service more frequently than expected, the company lost money on each interaction. It soon became clear that the algorithm promised speed but introduced blind spots. Triage decisions that would ordinarily be double-checked by a nurse or doctor were automated, and patients who received reassurances from the app sometimes delayed seeking medical care.

The financial fallout was no less dramatic. Babylon raised over £20 million from the NHS to develop its triage system and expanded quickly through private equity financing. By 2021 it

had gone public in the United States, only to implode spectacularly. Leaked internal figures revealed a deficit approaching $900 million. In 2023 the company filed for bankruptcy protection. Its UK assets were sold to eMed Healthcare for a mere £500,000—a stunning collapse for a company once touted as the future of health care. For taxpayers the cost was not just the direct financial loss. Time and resources that could have strengthened primary care were diverted to a system that ultimately failed.

Babylon is not the only example of algorithmic medicine overpromising and underdelivering. IBM's "Watson for Oncology" was pitched as a system that would digest reams of medical literature and provide treatment recommendations faster than any human oncologist. Hospitals around the world bought licenses, and IBM touted the product in marketing campaigns. Yet internal documents revealed that Watson often suggested unsafe or incorrect treatments. A physician involved in testing the system described it as "a parlor trick" useful for marketing but not for patient care. Some of Watson's errors stemmed from the limited data on which it was trained; others reflected that the system treated guidelines as static instructions rather than contextual recommendations. IBM quietly withdrew the product from many hospitals, but not before millions in public money had been spent.

Perhaps the most painful lesson from these cases is not that AI will always fail but that healthcare cannot be "streamlined" by technology alone. Britain's NHS has been chronically underfunded for more than a decade. Emergency departments have become so overcrowded that patients wait hours in corridors. Doctors and nurses are leaving the profession due to exhaustion and low pay. When an AI triage system misclassifies a heart attack or fails to spot sepsis, it isn't an overloaded algorithm—it is a human-made decision to substitute a computer for professional judgment. Babylon's collapse shows how quickly hype can turn into harm when oversight is weak. For every patient misdiagnosed by a chatbot, there is a person who delayed seeing a doctor or lost faith in the system. The solution is not to ban all AI in healthcare but to integrate these tools within robust public ser-

vices, with human clinicians responsible for every decision and adequate funding to match demand.

Automated classrooms

Education is another arena where algorithms were promised as saviors. During the COVID-19 pandemic, remote proctoring software became ubiquitous. Companies like Proctorio advertised tools that could monitor students through their webcams and flag "suspicious" behaviour, thereby preventing cheating and reducing the load on teachers. In practice the software often produced false positives and reinforced existing inequities. Civil rights advocates noted that AI proctoring tools harbor biases and present serious privacy risks. The American Civil Liberties Union argued that the technology was ineffective at catching cheaters yet inflicted "real harms" on students. At Utah's online Kings Peak High School, a principal acknowledged that the cameras made students feel as though they were "watching [themselves] take the test in the mirror" and said some families were allowed to opt out due to privacy concerns.

One particularly damning case came from a tech-savvy college student who tested Proctorio's face detection algorithm. Using a dataset of 11,000 images, he found the system failed to distinguish Black faces 57% of the time. Middle Eastern faces were misrecognized 41% of the time and white faces 40% of the time. A remote exam tool that cannot reliably identify students of different races not only undermines the fairness of assessments but also subjects Black and brown students to additional scrutiny. The company dismissed the findings as unrealistic, arguing that "children and cartoons" don't take tests, yet it did not produce evidence that the algorithm was effective in real-world conditions. Researchers studying AI proctoring concluded the software was "best compared to taking a placebo: it has some positive influence, not because it works but because people believe that it works".

The shortcomings of proctoring extend beyond race. In 2020 the UK government cancelled A-level exams due to the pandemic and introduced an algorithm to assign grades. The system took teachers' predicted grades and adjusted them based on a school's historical performance. About 40 percent of predicted grades were downgraded. Students from disadvantaged areas were disproportionately affected; 10.4 percent of students from deprived backgrounds saw their grades fall compared with 8.3 percent of their peers. Private schools received a far higher proportion of top grades (48.6 percent) than comprehensive schools (21.8 percent). Public outcry forced the government to abandon the algorithm, but the fiasco left thousands of students anxious about university placements and job prospects. The immigrant analogy is striking: an outsider (the algorithm) was blamed for a decision that should never have been automated without proper safeguards.

Beyond these headline failures, AI is steadily seeping into everyday classroom practices. By 2025, a report from the Gallup and Walton Family Foundation found that 60 percent of U.S. teachers used some form of AI tool during the 2024–25 school year. They reported saving nearly six hours per week on planning, grading and administrative tasks—time they could reinvest in personalized instruction and feedback. At Eastside Prep in Washington state, technology director Jonathan Briggs observed that "nobody really knows how to do this" yet. There is still not enough information about what kinds of AI use actually benefit students. When asked whether AI might help or harm human learning, he suggested that schools should focus on qualities that are hard to measure—collaboration, communication and creativity. In other words, the technology may boost performance in quantifiable subjects like math, but it cannot replace the struggle that makes learning meaningful.

Teachers worry that generative AI could erode students' ability to think for themselves. Chad Marsh, a veteran English teacher in the Lake Washington School District, observed that some students now feed bullet points into large language models and receive finished essays in return. Their writing "doesn't evolve" because "someone has basically generated [their] thoughts". At the

same time, AI detection software that tries to catch such behaviour often flags students unfairly, creating an adversarial classroom atmosphere. Marsh believes AI should be integrated thoughtfully rather than ignored. This sentiment is echoed by state education authorities: the Washington Office of Superintendent of Public Instruction released a "Human-AI-Human" framework in early 2024, urging that AI "enhance, not replace, human learning". Students themselves describe using AI for brainstorming, clarifying concepts and generating practice quizzes—but they also note that the tools can be inaccurate, rigid and may rob them of the "beauty of learning".

These diverse experiences reveal that AI does not simply overwhelm schools like a flood. It reconfigures how teachers teach, students learn and administrators make decisions. When used thoughtfully, AI can provide additional practice problems or free up time for teachers to engage more deeply with students. When deployed carelessly—whether in proctoring, grading or content creation—it can widen inequalities and erode trust. As with immigration, the scapegoating of AI hides deeper issues: under-resourced schools, overcrowded classrooms, and a culture obsessed with quantifying achievement. Algorithms amplify these problems; they do not create them.

Welfare and bureaucracy

In the world of social welfare, artificial immigrants have been asked to perform the role of gatekeeper. The Australian government's Robodebt scheme is a cautionary tale. From 2016 to 2019 the Department of Human Services used income averaging algorithms to identify welfare recipients who allegedly owed money. The system compared annual income data from the tax office with fortnightly benefit payments and automatically generated debt notices when discrepancies were found. More than half a million inaccurate debts were raised. Recipients were told to prove they did not owe the money; if they could not provide payslips from years earlier, the debt was assumed valid. The scheme was

declared unlawful by Australia's Federal Court in 2019 because income averaging across an entire year to determine fortnightly entitlements was invalid. In practice the system reversed the presumption of innocence and placed the burden of proof on the vulnerable.

The human cost was devastating. People on welfare received threatening debt letters and were instructed to repay thousands of dollars. Many took out loans to make the urgent repayments. The stress contributed to mental health issues and, tragically, several suicides. Eventually the government acknowledged that AU$746 million had been wrongfully recovered from 381,000 people and refunded the money. In 2020 a class action resulted in $1.75 billion of debts being wiped. The scheme was described by the Robodebt royal commission as a "crude and cruel mechanism, neither fair nor legal". Yet even after these findings, accountability remained elusive. In September 2025 the federal government agreed to pay an additional $475 million in compensation, bringing total redress to more than $2.4 billion. One class action member remarked that the settlement recognised the devastating impacts but could never fully heal the wounds. A partner at the law firm handling the case noted the irony: "One of the problems brought about by Robodebt was people being treated like numbers"—a reminder that dehumanization was baked into the algorithm.

The Netherlands offers another cautionary example. The Dutch tax authority introduced an algorithmic system to detect childcare benefit fraud. The model relied on risk factors such as nationality and ethnicity. Tens of thousands of parents—many of whom were low income or immigrants—were falsely accused of fraud. Their benefits were suspended, leading to financial distress, mounting debt and mental health problems. Amnesty International described the system as xenophobic; families were targeted because they held dual nationality. The scandal became so severe that the government resigned in 2021. Here the metaphor of AI as an immigrant is inverted: the algorithm acted like a border agent, deciding who was "eligible" for support based on na-

tionality. And just like the border, the system was opaque and unaccountable.

These welfare scandals reveal a pattern: governments often adopt algorithms not to improve services but to reduce costs. Automated systems allow officials to claim they are "modernizing" bureaucracy while quietly shifting the burden onto citizens. They invert the presumption of innocence and strip away the nuance that caseworkers bring to complex situations. The scapegoat narrative—that AI itself is overloading welfare offices—deflects blame from policymakers who starve social systems of resources and oversight.

How AI companies leach money

While the public endures misdiagnoses, unfair grades, and unjust debts, private companies have discovered a lucrative market in selling technological quick fixes to governments. Promises of efficiency and savings mask a deeper reality: these systems rarely deliver what was advertised, yet the contracts keep coming. The same names appear again and again—consultancies, platform vendors, data firms. Many patterns emerge across sectors.

Consultancy capture and pilot projects that never scale

Large consulting firms often act as intermediaries between public agencies and AI vendors. They promise transformation through pilot projects and feasibility studies. In the UK the NHS commissioned a *Federated Data Platform* (FDP) to integrate patient data across hospitals. The contract was awarded to Palantir, a U.S. surveillance and data analysis firm. Despite the hype, freedom-of-information requests show that by early 2025 only 34

NHS trusts—less than 15 percent—were actively using the platform. Over half of responding trusts had not adopted it at all. Many described Palantir's technology as a step backwards from existing systems. The contract's details were shrouded in secrecy; three-quarters of its 586 pages were redacted, prompting legal action. Originally adoption was optional, but in mid-2024 senior NHS executives wrote to trust leaders requiring them to plan adoption within two years. By February 2025, NHS England claimed that 96 trusts had "signed up," yet an FOI response revealed that only 34 were actually using the software while the rest had merely "signalled their intent". Consultants were paid £8.5 million to promote the system while many hospitals cut management staff.

IBM's Watson for Oncology and Babylon Health also illustrate the "pilot project" trap. Hospitals and health agencies spend millions on trials that rarely translate into routine use. When the projects fail, vendors can blame "implementation challenges," pocket the funds and move on. Public organisations, lacking in-house technical capacity, become dependent on external consultants. This dynamic echoes stories of colonization: resources are extracted while promises of development remain unfulfilled.

Vendor lock-in and perpetual fees

Another way AI companies profit is by locking public agencies into proprietary systems. Palantir's Foundry software has been criticised for making integration with common data-science tools difficult. Critics warn that once a hospital trust invests in Palantir's infrastructure, switching providers becomes costly. In effect, the company inserts itself into the heart of public health systems, making future budgets contingent on continued licensing and upgrades. As one analyst noted, the NHS could build its own data platform but lacks the funding and capacity; spending half a billion pounds on an unproven system risks long-term dependency.

Efficiency that cuts service and creates downstream costs

AI tools promise to "streamline" services by reducing human labor. In practice, the costs of errors and appeals fall on the public. When the UK's A-level grading algorithm produced unfair results, appeals clogged up hotlines and forced universities to overhaul admissions. When Robodebt miscalculated debts, tens of thousands of people filed appeals and class actions, leading to billions in compensation. These downstream costs far outweigh any short-term savings. Similarly, misdiagnoses by Babylon and Watson for Oncology required human doctors to intervene, sometimes after patients had suffered harm. In each case, the public health system paid twice: once for the faulty technology and again to clean up its mess.

Surveillance and data monetization

Some AI vendors derive revenue from data harvested through public contracts. Proctorio and its competitors collect video, audio and biometric data from students. Palantir's contracts with the NHS, the Ministry of Defence and police forces give it access to sensitive personal information. Investigations revealed that Palantir was working on a "real-time data-sharing network" with police forces that would process information including people's health, sexual orientation, sexual activity, trade union membership, race, religion and political beliefs. Critics fear this creates a surveillance infrastructure that can be repurposed beyond healthcare or education. In other words, companies profit twice: first by selling software and second by exploiting data.

Revolving door politics and lack of liability

Political support for AI projects is often lubricated by donations and lobbying. Babylon's meteoric rise was accompanied by over £250,000 in donations to the Conservative Party and high-profile endorsements from ministers. Consultancies like KPMG and companies like Palantir employ former government officials who know how to navigate procurement rules. When projects fail, liability often falls on the public purse. IBM and Babylon did not compensate patients who received wrong advice; the cost of misdiagnoses was borne by individuals and the NHS. In the case of Robodebt, officials who designed the scheme were not held accountable. The lack of liability encourages risky experimentation with vulnerable populations.

Appeals and oversight costs

Finally, AI systems create new bureaucratic layers. When an algorithm denies someone a benefit or miscalculates a grade, there must be a process to challenge the decision. Appeals committees, ombudsmen and courts get involved. The cost of adjudicating errors can exceed the savings achieved by automation. In the Netherlands thousands of families spent years fighting for justice. In Australia class actions and royal commissions swallowed millions. Each appeal is a reminder that algorithmic efficiency is often a mirage; by the time errors are corrected, the damage is already done.

The real overload

If AI is not clogging hospitals or flooding schools, what is? Chronic underinvestment, political neglect and social inequality.

In the NHS, waiting lists and staff shortages predate the introduction of AI. The algorithmic triage offered by Babylon did not overwhelm the system; it exposed the lack of capacity to deliver timely care. In education, overfilled classrooms, high teacher turnover and testing regimes created pressure long before chatbots and proctoring software arrived. Automated grading and surveillance did not cause the crisis; they simply shifted the burden. And in welfare, the Robodebt fiasco did not arise because an AI was overworked; it happened because policymakers chose to implement an algorithm that inverted the presumption of innocence and ignored warnings from experts. The real overload lies in the erosion of trust and the waste of public money. When people cannot understand or challenge decisions that affect their lives, they withdraw from civic life. When billions are spent on systems that do not work, cynicism grows.

The immigrant metaphor helps reveal the scapegoat. Throughout history elites have blamed immigrants for strains on public services when the underlying issue was deliberate austerity. Similarly, blaming algorithms for "overloading" healthcare, education or welfare distracts us from the policy choices that starve those systems. We are told that technology will make services more efficient; when it fails, the narrative shifts to cautionary tales about rogue machines. In reality humans set the rules, allocate budgets and decide whether to invest in capacity or outsource to private vendors. The overload is not a natural disaster caused by AI. It is a man-made consequence of austerity and deregulation.

What integration should look like

How might we welcome artificial immigrants without subjecting ourselves to their whims? The answer lies in integration, not assimilation or abandonment. AI systems should be colleagues, not overlords. That means embedding them in processes where human professionals retain ultimate responsibility. A triage chatbot can suggest possible causes for a cough, but a nurse or doctor must review the recommendation. An essay-scoring program

can identify grammar errors, but teachers should read the work and use the tool as a guide. A fraud detection algorithm can flag anomalies, but caseworkers must investigate before sending debt notices.

Integration also requires transparency and explainability. Public institutions should procure systems whose logic can be audited. Palantir's contract secrecy and proprietary algorithms undermine democratic oversight. Governments should require vendors to disclose training data, weighting factors and error rates. When algorithms are used in high-stakes decisions, individuals should have the right to an explanation and the ability to appeal to a human. The Washington state education guidance emphasises a "human-AI-human" loop: students should initiate and interpret AI use. This model should apply across public services.

Another pillar of integration is investment in public digital capacity. Rather than outsourcing to multinational firms, governments could develop in-house technical teams. The NHS's reliance on Palantir stems from decades of underfunding its IT infrastructure. Building public platforms would not only reduce vendor lock-in but also allow citizens to shape technology according to democratic values. The same holds for education: universities and schools could collaborate on open-source grading tools that reflect pedagogical values rather than surveillance. Welfare agencies could design simple verification systems that respect dignity and privacy.

Integration demands democratic governance. Communities should be involved in deciding when and how AI is used. Patients, teachers, students, welfare recipients and civil rights advocates must have seats at the table. Impact assessments should examine not only technical accuracy but also social consequences. In the Netherlands the disaster with the childcare benefits algorithm spurred calls for bans on racial profiling and human rights impact assessments. Similar mechanisms could be adopted elsewhere. Such engagement treats AI not as an alien invader but as a tool that can be shaped and controlled.

How we care for each other should never be abandoned to automation

Hospitals, schools, and welfare offices are where democracy touches flesh. They are the places where society shows whether every person matters—or doesn't. When we hand decisions in these spaces to opaque algorithms, we risk hollowing out the human foundations of care, teaching, and solidarity. The real danger isn't that AI overloads public services; it's that we allow technology to strip out the compassion that makes them humane.

The immigrant metaphor helps: new arrivals can enrich a community when there are resources and goodwill; they become scapegoats when systems are already strained. Likewise, AI can strengthen public services if integrated thoughtfully within well-funded institutions. But if we treat algorithms as cheap stand-ins for nurses, teachers, or caseworkers, we condemn ourselves to a future of automated neglect and lasting mistrust. Some domains—personal care above all—demand the touch of another human being. The choice is ours, but it requires the democratization of these systems.

Chapter 6: They bring crime and disorder

They come without passports and leave with your identity.

Every immigration wave brings fears that the newcomers will erode public safety. In the nineteenth century it was Irish dock-workers; in the twenty-first it is families crossing an ocean for refuge. Today the figure of the dangerous outsider has taken a digital form. Artificial intelligences – the models that suggest songs, reply to emails and generate artwork – are also lawless migrants from a wild frontier. The presence of these "algoes" does make us less secure, they make crime easier and our institutions chaotic. The chapter that follows explains how AI tools can be weaponised to amplify crime and disorder, not because they are malicious themselves, but because human actors exploit their capabilities. The aim is not to provoke panic. Instead, it offers a clear-eyed account and asks what citizens and governments can do

Fear of disorder and the immigrant metaphor

The immigrant metaphor that frames this book is not an accusation against machines; it is a lens on our society. When people worry that a group "brings crime," they are often projecting deeper anxieties about economic insecurity and social change.

The same is true of AI. Criminals and authoritarian regimes have already harnessed automation to commit fraud, suppress dissent and infiltrate systems. But AI itself is a tool, akin to a car or a printing press. Blaming the tool alone obscures the human agency that decides whether it is used to help or to harm. In this chapter we will distinguish between intrinsic risks (the ways an algorithm can behave unexpectedly) and misuse (deliberate exploitation by bad actors). The comparison to immigrants also highlights how generalised fear can lead to discriminatory responses. Rather than panic, we need policy and oversight. The next sections examine different kinds of harm and show that behind the headlines there are human decisions, corporate incentives and regulatory gaps.

Human rights violations and crimes against humanity

One of the most chilling applications of AI is its use by governments to monitor, predict and control populations. Investigations by the European Parliament into the Xinjiang region of China reveal that authorities there use AI-driven predictive policing and facial recognition to identify and detain members of the Uyghur minority. Data from cameras, license plate readers and mobile devices are fused in systems like the "Sharp Eyes" programme to identify "focus personnel" and flag behaviours deemed risky. These data are fed into the Integrated Joint Operations Platform, which links ID numbers to physical characteristics and uses machine learning to generate lists of individuals for detention. Many Uyghurs have been sent to re-education camps without charges or trials, purely on the basis of algorithmic predictions. The system's designers claim it prevents terrorism; in practice it enables the mass detention of an ethnic group.

Predictive policing has also been tried in democratic societies, often with troubling results. A 2024 report from the U.S. Department of Justice notes that many predictive models rely on his-

torical arrest data that reflect human biases. Tools trained on such data tend to over-police Black and Latino neighbourhoods and misidentify individuals. Some agencies have abandoned person-based predictive policing because it offers little value while posing privacy risks. The same report warns that risk assessment algorithms used in courts can perpetuate inequality by categorising people differently despite similar histories. At least seven mistaken arrests in the United States have been linked to facial recognition errors, nearly all involving Black citizens. Here, the harm is not that AI "decided" to target minorities; it is that police departments adopted commercial systems without adequate testing or oversight, and built them on data sets saturated with systemic bias.

To understand why such abuses flourish, one must look at corporate incentives. Companies selling surveillance and predictive tools market them as neutral, efficient and profitable. Governments, eager to demonstrate security, purchase them. Ordinary citizens – often those with the least power – become the test subjects. This dynamic mirrors xenophobic narratives about immigrants as criminals: the focus is on control, not accountability. Recognising that AI tools can be weaponised for oppression helps us demand better protections. International human rights law already forbids arbitrary detention and discrimination; those norms must be enforced when algorithms are involved. Transparency about data sources and auditing of systems can help. Importantly, the victims of algorithmic repression are not anonymous masses but individuals with families, livelihoods and communities. Their stories remind us that "crime and disorder" can be inflicted by states as well as by thieves.

Copyright expropriation and creative theft

Another area where fears of AI echo familiar immigrant stories is intellectual property. Creators worry that outsiders will steal their work and erode their livelihoods. In the context of AI, those fears are grounded in reality. Large language and image mod-

els are trained on vast corpora of text, art and music scraped from the internet, much of it copyrighted. A federal case decided in June 2025 illustrates the tension. Authors including Sarah Silverman and Ta-Nehisi Coates sued Meta, alleging that the company used pirated versions of their books to train its model Llama. U.S. District Judge Vince Chhabria dismissed the case, stating that the plaintiffs had not provided enough evidence that Meta's use would dilute the market for their work. But he also emphasised that using copyrighted work without permission would be unlawful in "many circumstances" and that this ruling simply reflected poorly argued claims. Chhabria noted that generative AI could flood markets with endless images, songs and articles created with a fraction of the effort humans expend. In other words, the court recognised that models trained on copyrighted works may undermine incentives for creation.

The deeper problem is opacity. Training datasets are often poorly documented and inconsistently licensed. A 2025 analysis from MIT Sloan observes that popular models like GPT-4 are built on data whose lineage is unclear. Without transparency, developers and users may unknowingly violate regulations like the European Union's AI Act or expose themselves to copyright lawsuits. Researchers with the Data Provenance Initiative audited more than 1,800 text datasets and found that license information was omitted in over 70 % of cases, while license types were miscategorised over 50 % of the time. These errors make it hard for responsible companies to select lawful training data, and they make it easy for unscrupulous actors to claim ignorance.

Artists and journalists describe how their work is cannibalised. A painter might find derivative images flooding social media, produced by an AI trained on her portfolio. A novelist's style can be imitated to generate endless fan fiction. When newspapers discovered that a model reproduced their paywalled articles verbatim, they sued. The question is not whether AI should be allowed to learn from human culture – that is how humans learn – but whether those whose work feeds the models should have a say and receive compensation. Historically, immigrants have been accused of stealing jobs and culture while in reality they enrich and

diversify societies. The same is true of AI if governed wisely: it can inspire new art and collaboration. Policies requiring dataset provenance, consent from creators and fair remuneration would reduce the feeling of expropriation and align incentives toward creativity rather than theft.

AI-powered scams and social engineering

Fraudsters have always exploited new technologies. The telephone enabled "grandparent scams" and phishing emails promised fake inheritances. AI lowers the cost and raises the realism of such tricks. In early 2024 the Electronic Frontier Foundation warned about voice-cloning scams. Criminals record a few seconds of a person's voice – from social media posts or voicemail greetings – and then generate convincing audio messages. One scenario recounted by investigators begins with a panicked call: a voice that sounds like your child says they have been arrested and need money. A second voice posing as a lawyer demands immediate payment. The voices are realistic because websites now offer cheap voice-cloning services; in some cases scammers need only a thirty-second audio sample. Victims often wire money before they verify the situation. The EFF notes that while such scams are not yet widespread, the technology is accessible enough that they could become common. They recommend simple defences such as family passphrases.

Corporate scams have already moved beyond prototypes. In 2024 an employee at the engineering firm Arup in Hong Kong joined a video call with people who looked and sounded like senior executives. He believed he was fulfilling a routine transfer, but the meeting was a deepfake. Criminals used AI-generated video and voice to impersonate multiple executives simultaneously and convinced him to transfer HK$ 200 million (about £20 million) into accounts they controlled. The case shows that deepfake scams are no longer restricted to simple audio; they can involve coordinated video conferences, realistic backgrounds and body movements. Arup's chief information officer told reporters that attacks includ-

ing invoice fraud, WhatsApp voice spoofing and AI-generated deepfakes have "risen sharply". Law enforcement reported no arrests.

The problem is not only deception but also scale. AI models can create thousands of tailored phishing emails that mimic a company's writing style and include plausible personal details gleaned from social media. Cybersecurity experts warn that generative models can automate phone scripts, design fake websites and simulate entire call-centre operations. Once criminals used to need technical skills to craft malware; now they can ask a model to write ransomware or produce a PDF laced with malicious code. The victims are not just large companies. A small bakery owner in our town of Rugeley described receiving an email that appeared to be from her supplier; it contained a new bank account number and a friendly note referencing a recent conversation. She wired £8,000 before realising she had been scammed. She later learned that the attackers had used AI to scrape her posts and invoice templates, making their request familiar and credible. In this way AI democratises fraud: both criminals and victims are everyday people.

AI-assisted hacking and cybercrime

The same capabilities that allow AI to compose poetry can be redirected to composing malware. A 2025 threat intelligence report from OpenAI describes how its investigators detected and disrupted several campaigns in which state-linked actors used generative models to enhance cyberattacks. In one case, North Korean operators used ChatGPT to generate résumés and cover letters at scale to infiltrate IT companies with deceptive job applications. They even sought advice on remote-work setups and used AI-generated content to recruit collaborators. The same report documents Chinese and Russian operations that used AI to craft social-engineering scripts, spam comments and disinformation. OpenAI notes that AI can be used at every stage of an attack: reconnaissance (asking about software vulnerabilities), develop-

ment (writing and debugging malware), phishing (crafting credible emails) and post-exploitation (automating network scans). The company says its systems help detect and shut down such misuse, but the broader security community must adapt.

Generative AI also reduces skill barriers for hackers. The BleepingComputer analysis of OpenAI's February 2025 disruption efforts found that threat actors used models to ask about default passwords, identify configuration errors, and obfuscate malicious code. Attackers who previously relied on publicly available scripts can now generate customised payloads and adapt them to specific networks. Combined with stolen credentials from massive data breaches, these tools enable "credential stuffing" on an industrial scale. The Identity Theft Resource Center (ITRC) reports that in the first half of 2025 there were 1,732 data compromises, an 11 % increase over the same period in 2024. Cyberattacks accounted for 77.8 % of these incidents, compromising the personal data of more than 114 million individuals. Phishing and business-email compromise were the second most common cause, followed by ransomware. The ITRC notes that generative AI platforms are increasingly used to create sophisticated phishing campaigns that are harder for organisations and individuals to detect.

These trends create what cybersecurity scholar Qi Liao calls "ransomware 2.0": attacks that not only encrypt data but also steal and sell it. Liao explains that AI can automate and personalise phishing by analysing social-media posts, generate deepfakes for blackmail, and craft malware that evades detection. He warns that adversarial machine learning techniques can bypass spam filters and anti-virus systems. Meanwhile, criminals can cross-reference anonymised data from medical, financial and voter records to re-identify individuals and learn their habits. It is a chilling inversion of AI's promise: tools meant to secure and streamline our lives are used to break into them.

Identity theft and the erosion of privacy

Identity theft is the crime that best illustrates the intersection of AI, data and everyday life. In 2025 the UK fraud prevention service Cifas reported that more than 217,000 fraud risk cases were logged in the first half of the year, a record number. Of these, over 118,000 were identity fraud cases. Cifas warns that AI enables criminals to create synthetic identities: fabricated personas that combine stolen data with invented attributes to bypass verification systems. Economic pressures and risky consumer behaviour mean that some individuals even sell their own identities, exposing themselves to liability when criminals take loans or open accounts in their names. Sectors once considered low risk, such as motor insurance and telecommunications, are now prime targets. Facility takeovers – where attackers seize control of existing bank or phone accounts – accounted for 18 % of all fraud, and the telecommunications sector saw the steepest rise.

The ITRC adds that data breaches feed these scams. Among the 165 million individuals affected by breaches in the first half of 2025, many had their login credentials and personal identifiers packaged into a single database. This repackaging means that criminals can easily buy lists of usernames and passwords and use AI to test them across multiple services, a tactic known as credential stuffing. James E. Lee of the ITRC notes that more than two-thirds of breach notifications fail to disclose the attack vector, leaving consumers unaware of how their data were stolen. The lack of transparency prevents individuals from taking targeted protective steps and undermines public trust.

Beyond financial fraud, privacy erosion occurs through de-anonymisation. AI can cross-reference public data with stolen datasets to identify individuals in supposedly anonymous datasets. For example, a malicious actor could combine a health-care dataset with voter registration information and smart-home device logs to infer who is being treated for a particular condition and when they are away from home. Such granular profiling enables blackmail, targeted advertising and

discrimination. The victims might never know how they were selected for a scam or denied a job. If data is "the new oil," then identity theft is the new pipeline leak – invisible until a spill occurs.

Algorithmic coercion into violence and self-harm

While AI is often blamed for making people do things, the reality is subtler. Recommendation systems are designed to maximise engagement, and they do so by amplifying content that elicits strong emotions. Sometimes that content includes hate, extremism and self-harm. A 2024 study of Instagram's moderation revealed that Meta failed to remove any of 85 posts showing various degrees of self-harm. Even more troubling, Instagram's algorithm encouraged accounts posting self-harm images to befriend each other, helping a network of harmful content grow. Psychologist Lotte Rubæk told reporters that such failures trigger vulnerable users and can be a matter of life and death. Rather than accidental oversights, these outcomes reflect an incentive structure that rewards engagement over safety.

Scholarship on extremism shows a more complicated picture than popular accounts suggest. In a review of empirical studies, communication researcher Aaron Shaw notes that while public narratives often blame recommendation algorithms for radicalizing unsuspecting viewers, the evidence is less clear-cut. YouTube's algorithm rarely pushes extremist content to casual users. However, the platform still hosts vibrant extremist communities, and its design can help reinforce the views of those already inclined toward resentment. Many radicalized individuals reach these networks through external links or by subscribing to fringe channels rather than stumbling upon them by accident. Once inside, gaps and inconsistencies in moderation allow conspiracies, misogyny, and hate to circulate with relative freedom. Researchers have linked participation in these online spaces not only to hardened attitudes

but also to higher likelihoods of involvement in civil unrest and, in some cases, political violence.

There is also evidence that algorithms can overexpose vulnerable people to self-harm content. A 2025 qualitative study with young people, policy makers and social-media industry professionals found that 83 % of youth have encountered self-harm or suicide content online, often before the age of fourteen. Participants warned that the way social-media algorithms operate can lead to overexposure when teens interact with peers. Despite these risks, national suicide-prevention strategies rarely include social media guidelines. Social-media companies have adopted some policies, but they differ across platforms and lack independent oversight. The result is a patchwork of private rules that leaves many gaps. When tragedies occur, companies often say "we didn't design it to do that." That refrain echoes the way some societies blame immigrants for social ills without examining the policies and incentives that create the conditions for exploitation.

Opacity, accountability and the black box problem

Underlying many of these harms is the black box nature of AI systems and the secrecy of their creators. The MIT Sloan report mentioned earlier points out that without transparency into the lineage of training data, it is difficult to assess legal risk or bias. The Data Provenance Initiative discovered that more than half of licensing information in popular datasets was miscategorised and that over 70 % of datasets lacked license information. These findings show that many companies build products on shaky foundations.

Opacity is also present in the public sector. During the first half of 2025, the ITRC noted that 69 % of data-breach notifications failed to disclose how the breach occurred. This means that victims are left without information that could help them respond, and

patterns of attack remain hidden. Similarly, when algorithms misidentify individuals or recommend harmful content, companies often claim trade secrets to avoid revealing how their systems work. Such secrecy hinders redress and allows abuses to continue. In a 2024 Mozilla Foundation report examining watermarking and labelling of AI-generated content, researchers concluded that human-facing labels are easy to manipulate and can even erode trust. They argue for machine-readable markers combined with robust detection mechanisms and a multifaceted governance approach. Without strong disclosure requirements, citizens cannot know whether a video or call is authentic.

The black box problem is exacerbated by corporate incentives. Proprietary models are guarded to maintain a competitive edge, even when they have public consequences. This is akin to a landlord refusing to repair a building because the blueprint is a trade secret. Regulation must strike a balance between protecting innovation and protecting the public. The California Generative AI: Training Data Transparency Act, signed in 2024, is an example of legislation that requires companies to document their training data sources. European policymakers are considering similar measures. Transparency alone is not a panacea – after all, criminals will not volunteer their methods – but public institutions have a duty to set standards.

Agency, law and remedies

The narrative that AI immigrants bring crime and disorder can lead to two dangerous responses: complacency ("nothing can be done") or technophobic bans. Both ignore the role of human agency. The harms described above stem from human decisions to deploy surveillance against minorities, build opaque models on stolen data, design engagement-maximising algorithms without safety checks, and weaponise generative tools for fraud. Human decisions can also mitigate those harms.

Strengthening civil liability is the first step toward accountability. Just as landlords are responsible for safe buildings, companies that deploy AI should be liable for foreseeable harms. If a predictive policing tool leads to unlawful detention, those affected deserve the right to sue and demand audits. Courts could require disclosure of model assumptions and training data, while whistleblower protections should extend to engineers and ethicists who raise concerns about misuse.

Alongside liability comes the need for data-provenance requirements and licensing. Policymakers can require companies to keep records of where training data come from, what licenses apply, and how data subjects can opt out. The Data Provenance Initiative has already shown this is feasible—auditing datasets, reducing unspecified licenses from 70 % to 30 %, and offering tools for exploration. Clear provenance not only supports copyright holders but also protects companies from ingesting harmful or illegal material, like the child abuse images found in the LAION-5B dataset.

Any serious framework also demands mandatory red-team audits. Before deployment, systems should undergo adversarial testing by independent researchers to expose vulnerabilities and misuse pathways. The practice is standard in cybersecurity and can be adapted for AI. Red teams should bring together experts in human rights, mental health, and cybersecurity, not only technical engineers. Their findings ought to be published, fixes documented, and lessons shared.

Education and public infrastructure matter just as much as oversight. Citizens need accessible information about deepfake scams, identity theft, and privacy safeguards. Public-awareness campaigns—akin to anti-smoking or seat-belt initiatives—could help. Simple practices like setting family passphrases, verifying calls through known numbers, using password managers, and enabling two-factor authentication can make a huge difference. Public libraries and community centres could act as hubs for digital-literacy workshops.

Governance at the platform level is equally vital. Social-media companies should grant researchers access to data, with privacy safeguards in place, so the societal impact of algorithms can be studied properly. Laws such as the UK's Online Safety Act or the proposed U.S. Kids Online Safety Act should include explicit provisions on suicide and self-harm content. Algorithms that recommend content need built-in throttles for harmful material and should be subject to external review.

All of these measures point toward a broader truth: responsibility is collective. Crime and disorder don't arise because tools are inherently evil; they arise because humans choose to exploit them. Casting AI as an invading criminal class risks repeating old patterns of scapegoating. A more constructive view treats AI as a new population of residents in our digital society. Like any residents, they need rules, oversight, and a culture of care. Our task is to shape those rules so that AI contributes to safety and justice, rather than undermining them.

From fear to agency

If the previous chapters have shown that AI immigrants do not always integrate or always work, this chapter confronts the fear that they will bring crime and disorder. The evidence is nuanced. Yes, AI can be weaponised for repression, fraud and hacking. Yes, algorithms can amplify self-harm and extremism. Yes, opaque practices make accountability difficult. But at every turn, human decisions, incentives and institutions shape those outcomes. Understanding the mechanisms of harm – predictive policing built on biased data, datasets scraped without consent, algorithms designed to maximise engagement – empowers us to change them.

As you finish this chapter, consider the parallels between today's AI and yesterday's immigrants. Demonising a group rarely leads to safety; building equitable systems does. What role can you play, as a voter, consumer or creator, in demanding transparency,

accountability and care from those who build and deploy these powerful tools?

Chapter 7: The Synthetic Scapegoat

Every era invents its scapegoats.

In ancient times, a community would take a goat, symbolically load it with the sins of the people, and send it into the wilderness. The ritual—described in the Book of Leviticus—was both literal and psychological. Guilt was offloaded. Order restored.

But it wasn't always a goat. Sometimes it was a prisoner. Often, it was someone on the margins—a stranger, an outsider, a migrant. Someone to blame. Someone to expel.

These weren't metaphors. They were survival strategies—deeply human, deeply political. And they've never really gone away.

Today, the scapegoat is synthetic.

When mass layoffs hit white-collar workers in early 2025, press releases from Workday, Autodesk and CrowdStrike sounded eerily similar: profits would *soar* because of "the increasing demand for AI". A pivot toward artificial intelligence conveniently justified eliminating thousands of jobs while investors cheered. To casual readers it seemed the machines themselves were behind the bloodletting. But as the Associated Press pointed out, tech job postings were already down 36 percent from 2020 levels, and economists noted that the cooling job market for developers looked similar to other sectors. CEOs were folding macroeconomic realities, pandemic hangovers and shareholder expectations into a single narrative: *we had no choice—the algorithm made us do it.*

This chapter argues that AI has become a perfect scapegoat. It cannot vote, strike, or file a lawsuit; it has no voice to contradict the narrative pinned on it. While there are real risks in delegating human decisions to software, blaming "the model" distracts us from the corporate, political and social choices driving those decisions. By naming AI scapegoating we also illuminate the emergence of a data-rich class operating beyond democratic accountability. The real crisis is human—a crisis of accountability, agency and democratic will.

Historically, scapegoating has been used to excuse bad policy and divide the public. Immigrant communities, in particular, have long served as convenient targets. In the United States, politicians have demonized newcomers as criminals and rapists and accused them of "taking jobs" from native-born workers—rhetoric that Professor Mark Edberg calls an age-old ploy to dehumanize. Canada's federal government recently faced backlash for blaming migrants for a housing and affordability crisis even though rents rose far faster than population growth. This kind of scapegoating is not just false; it distracts from corporate profiteering and systemic failures. The current panic about "the algorithm" follows a similar script. Rather than address the political and economic forces driving inequality, leaders locate the villain in a voiceless entity—whether a person without citizenship or a machine without agency—and invite the public to vent their anger there.

The Perfect Fall Guy

Why does artificial intelligence make such an attractive fall guy? For one, AI evokes both awe and unease. To those outside tech, "algorithms" appear mysterious and autonomous—machines spinning outcomes free from human influence. This narrative is useful to corporate leaders making unpopular decisions. In January 2025 Workday's CEO claimed that "increasing demand for AI has the potential to drive a new era of growth," while announcing layoffs that affected about 8.5 percent of staff. Autodesk's chief executive said resources would be redeployed

to AI initiatives, implying that job cuts were necessary to fund innovation. These pronouncements were widely reported without critical context; to the general public it seemed AI itself was eliminating jobs.

The scapegoating extends well beyond tech layoffs. In housing, a handful of firms sell rent-setting software that recommends prices across multiple landlords. These tools are now the subject of antitrust suits alleging that the software enabled coordinated price hikes. Rather than admit to collusion, landlords blame the algorithm. Lawmakers have responded with proposals to bar algorithmic price-fixing and several cities have banned revenue-management software.

Insurance companies have adopted similar tactics. Major insurers face lawsuits accusing them of using algorithms to deny care. Investigations show that automatic denials are often reversed on appeal, yet very few patients know to challenge them. By attributing denials to "the system," insurers deflect scrutiny from profit-driven policies and downplay human oversight.

AI scapegoating works because algorithms are, by design, inscrutable. When a person or institution cites "the algorithm" as the culprit, critics can neither cross-examine the code nor subpoena the variables. This asymmetry empowers those who control AI systems to deflect blame onto the technology. The perfect fall guy can't protest or sue.

The instinct to blame a faceless other has a long pedigree. Governments and media have often pinned economic or social crises on immigrants, portraying them as job thieves or drains on public services. Canada's decision in 2024 to slash immigration targets while rents continued to skyrocket was criticized as a "gut-wrenching betrayal" and a dangerous deflection of blame. In the United States, labelling immigrants as criminals and asserting that they steal jobs is, as Edberg notes, an age-old tactic. These narratives thrive because the people being blamed are marginalized and cannot easily contest the story. In the same way, AI's inscrutability allows those in power to offload responsibility onto

a tool that cannot speak for itself. Both forms of scapegoating displace attention from the policies and business models that cause harm and allow decision makers to avoid accountability.

Bias as a Mirror, Not a Virus

A common narrative portrays algorithmic bias as an infection: engineers deploy machine learning systems, and somehow bias "emerges" like a virus. This metaphor casts algorithms as contaminated objects rather than mirrors reflecting human decisions. Research from University College London (UCL) and Boston University makes clear that AI does not generate prejudice from nothing; it learns our biases and often amplifies them back at us.

In a 2024 study published in *Nature Human Behaviour*, researchers found that interacting with a biased AI system increased participants' own bias. For example, a system trained to underestimate women's performance caused users to undervalue women in subsequent tasks. A second experiment used the generative model Stable Diffusion to create images of "financial managers"; participants exposed to these images were significantly more likely to select a white man for the role. Co-lead author Moshe Glickman noted that biased AIs can alter people's beliefs and that accuracy can improve judgments only when AIs are trained and designed carefully.

This finding is echoed by research led by Carey Morewedge at Boston University. His team found that participants detected more bias in decisions attributed to algorithms than in identical decisions attributed to themselves. Morewedge notes that algorithms learn and reproduce human biases but can also act as mirrors: when people see discrimination in a model's output, they become more willing to correct it. In experiments with Airbnb listings, Lyft driver profiles and résumés, participants were likelier to identify and adjust bias when they thought it came from a machine rather than their own judgment. "Algorithms are a dou-

ble-edged sword," he writes: they amplify our worst tendencies but also help us see and correct them.

Seen this way, algorithmic bias is not some alien pathogen infecting neutral code; it is a codification of our values, data and practices. When biased algorithms produce harmful outcomes, they are reflecting systemic issues in hiring, housing, lending and policing. Blaming AI for prejudice erases the responsibility of those who built the datasets, selected the variables and decided to deploy the model. Like a funhouse mirror, AI exaggerates features we would rather ignore. Instead of smashing the mirror, we need to confront the image staring back at us.

The same lesson applies to the stereotypes that fuel anti-immigrant politics. Accusations that migrants are criminals, rapists or job thieves are not facts emergent from data; they are projections of longstanding fears and prejudices. As Edberg notes, such labels are "repugnant, racist, ignorant and false", yet they persist because they serve political agendas. When biased policing or hiring data are fed into machine-learning systems, the resulting discrimination is not a surprise; it is a mathematical echo of the bigotry that already exists. Recognizing bias as a mirror—not a virus—forces us to confront both our algorithmic tools and the social narratives that shape them.

Blame the Model, Not the Mogul

The language of "the algorithm" provides a convenient scapegoat for corporate actors who want to avoid owning the consequences of their decisions. When Amazon's hiring algorithm downgraded résumés with female markers, the company said it scrapped the system. But the problem was never the code alone; Amazon's historical hiring data—dominated by men—taught the algorithm to favor male candidates. Amazon managers could have acknowledged the underlying sexism in their workforce and addressed it, yet the narrative focused on a rogue algorithm gone bad.

This pattern appears in other sectors. Health insurance companies rely on algorithms to predict patient care needs and determine coverage, but when these models deny legitimate claims, insurers act as though the machine is the culprit. In reality, executives decide how much denial risk is acceptable and whether appeals will be burdensome. Similarly, RealPage's rent-setting software collects data from multiple landlords to recommend prices. Landlords and software makers insist they are not colluding because "the algorithm" sets prices; antitrust officials disagree. By attributing price hikes to AI, landlords mask the underlying goal: maximizing profits.

Automated content moderation and recommendation systems also enable scapegoating. Social media companies claim that algorithmic feeds—not business models—are responsible for radicalization, misinformation and polarization. In truth, algorithms are optimized to maximize user engagement because that drives advertising revenue. A 2025 review on the neurophysiological impact of social media algorithms explains that AI-driven recommendation engines are designed solely to capture attention for profit, altering dopamine pathways in users' brains and fostering dependency analogous to substance addiction.

These systems continually tailor feeds to individual preferences, maximizing screen time and deepening activation of the brain's reward centers. When mental health crises result, platform executives point to the abstract "algorithm" rather than acknowledging a business model built on addictive design. Teen social media use illustrates this tension: 96 percent of U.S. teens go online daily and nearly half report being online "almost constantly"; yet, when negative content leaves users depressed, the companies fault the feed rather than their engagement-maximizing incentives.

Blame the model, not the mogul—so the mantra goes. But models do not choose to fire people, raise rents, deny medical claims or addict teenagers. Those decisions are made by people and corporations who design, deploy and profit from the systems. Shifting blame to AI deflects scrutiny away from corporate accountability and the policies that allow these practices.

The Rise of the Post-Democratic Class

The ability to offload responsibility onto AI is not simply a rhetorical trick; it is part of a broader shift toward a data-rich elite that operates beyond democratic checks. AI systems require massive datasets, compute infrastructure and engineering expertise. Only a handful of companies—Google, Amazon, Microsoft, Meta, Tesla and their partners—control these resources. In the *Artificial Power 2025* report, the AI Now Institute warns that if we accept the trajectory offered by Big Tech as "inevitable," we face a future that disenfranchises large sections of the public, renders systems more obscure, devalues crafts and undermines security. The report argues that AI is fundamentally about concentrating power in the hands of a few firms. These companies "shore up existing advantages" by amassing data, compute and talent; the bigger-is-better paradigm aligns with their incentives and squeezes out smaller competitors.

This consolidation creates what scholars describe as a **post-democratic class**—an elite whose control of data and compute allows them to shape society while evading oversight. The AI Now Institute warns that tech oligarchs are "counting on a wholesale rewriting of our social and economic foundations". Economist Simon Johnson notes that the vast majority of investment in AI is concentrated in a few nations, giving the global tech industry unparalleled power.

AI-powered surveillance often targets workers and marginalized groups, while executives remain largely unmonitored. Gig workers delivering food or driving ride-share vehicles are subject to algorithmic management that determines pay and termination without due process, yet the CEOs of data-rich platforms sit outside the reach of democratic oversight. When platforms leak personal data or chatbots hallucinate, executives highlight the complexity of the technology rather than their own negligence, using AI as a shield to protect their prerogatives.

The combination of concentrated power and scapegoating fosters a post-democratic environment: decisions affecting millions are made within boardrooms and data centers, insulated from public scrutiny. This class wields algorithms as instruments of governance, substituting computational authority for democratic deliberation. Just as corporations hide behind AI, politicians hide behind vulnerable human targets. When Canada's government responded to soaring rents by cutting immigration and suggesting migrants were to blame, advocates pointed out that population growth of just 3.9 percent could not explain a 20 percent rent surge. In the U.S., politicians have long portrayed immigrants as criminals and job stealers despite evidence to the contrary. Casting migrants as culprits allows leaders to avoid confronting corporate profiteering and housing shortages. Likewise, executives who blame "the algorithm" dodge accountability for choices driven by profit, thereby reinforcing a culture of scapegoating that obscures the real centers of power.

Social Media and Learned Helplessness

Social media platforms illustrate how AI-driven systems foster learned helplessness while distracting users with algorithmic scapegoats. A difference-in-differences study using Dutch data found that Instagram's introduction of an algorithmic feed led to poorer mental health among teenagers because the feed rewarded negative social comparisons. A 2025 neurophysiological review further describes how recommendation engines are designed to maximize attention; by modulating dopamine pathways they encourage compulsive engagement and heighten emotional sensitivity. MIT researchers observed a feedback loop in which users with mental health symptoms consumed more negative content, which then worsened their symptoms.

Instead of addressing the profit motives behind these designs, society often blames the "feed" itself. Politicians hold hearings about harmful algorithms; parents lament that TikTok's recommendations are destroying their children's attention spans. But

the underlying business model—selling targeted ads by capturing attention—is rarely questioned. In 2024 nearly half of U.S. teenagers reported being online almost constantly and 96 percent went online every day. When youths become depressed, we pin the blame on an algorithm rather than on companies that intentionally engineer addictive environments.

AI-driven platforms also amplify cultural self-loathing. Instagram and TikTok serve up endless highlight reels of others' lives, prompting users to measure themselves against unrealistic beauty and success standards. The constant barrage of curated content fosters feelings of inadequacy and body image issues; a meta-analysis found a 13 percent increase in depression risk for each additional hour spent on social media. As more AI tools generate hyperreal faces and bodies, the gap between digital perfection and lived reality grows. This, too, is blamed on "the algorithm," not the industries monetizing insecurity.

The feeds also amplify political scapegoating. Recommendation engines optimize for outrage and engagement, so inflammatory posts about immigration or crime rise to the top. False claims that migrants are criminals or that immigration causes unemployment circulate widely, even though research shows immigrants are less likely to commit crimes and often complement native labor. When users become fearful or resentful after consuming these narratives, they may blame "the algorithm" for radicalizing them. Yet those narratives reflect a deliberate choice by platform operators to boost polarizing content because division drives clicks. The same infrastructure that fuels self-loathing can also fuel xenophobia, deepening social divisions while shielding the corporations profiting from them.

Learned helplessness arises when individuals feel they have no control over their environment. Users know that social media feeds are engineered to be addictive and that negative content harms them, yet they continue scrolling because the platforms offer few alternatives and social life has migrated online. By portraying AI systems as autonomous forces, platform owners encourage

this helplessness. If the feed is in control, why bother resisting? The scapegoat narrative invites resignation.

We Sleepwalked Into This

It is tempting to frame AI scapegoating solely as a top-down phenomenon: corporate moguls shift blame onto machines, and the public naively accepts it. Yet we all share some responsibility. Our hunger for convenience, entertainment and personalization created the market for algorithmic systems. Millions of users freely train recommendation engines with their data; managers adopt AI hiring tools to process tens of thousands of applications because human review would be slower; doctors acquiesce to algorithmic risk scores to manage workloads. When we accept biased recommendations, personalized pricing and automated denials as the cost of participation, we normalise the delegation of judgment to machines.

Survey data show that by 2024 smartphone ownership among U.S. teens reached 95 percent. Many teens described their usage as "almost constant," with 73 percent visiting YouTube daily and 12 percent reporting near-constant use of Instagram. Social life now unfolds in algorithmic spaces. It is no wonder that individuals struggle to imagine alternatives; the digital infrastructure for community, work and politics is privately owned and optimized for profit. We are not just passive victims of AI scapegoating; we have, through inattention and convenience, contributed to the very systems that now govern us.

Our passivity has also allowed anti-immigrant scapegoating to flourish. Many voters accept claims that migrants cause crime or housing shortages without examining the evidence or the motives of those making such assertions. By failing to challenge xenophobic narratives or hold leaders accountable, we contribute to a political environment in which blaming the vulnerable is normalized. That same reflex now emerges in discussions about AI: instead of interrogating corporate agendas or public policy, we blame a

technology we barely understand. Naming these parallels is not about equating the plight of migrants with the status of software; it is about recognizing a consistent pattern of deflection and the ways our complacency enables it.

Similarly, voters often treat complex social problems as technical puzzles to be solved by AI. Policymakers tout automated monitoring for everything from pandemic contact tracing to welfare fraud detection. In doing so, we outsource moral and political decisions to algorithms. When those systems inevitably replicate existing inequities, we act surprised, as if the code were to blame. But if we refuse to name the political and economic forces shaping these systems—deregulation, austerity, shareholder primacy—we cannot hope to change them.

Reclaiming Responsibility

What would it look like to stop scapegoating AI and instead reclaim responsibility? We must insist on transparency and accountability from those who deploy algorithms. If a company uses a model to set prices or approve care, regulators should know what data it uses and how it weighs outcomes. Litigation and legislation targeting algorithmic collusion and automated denials are steps in this direction, but transparency will matter only if coupled with reforms that address underlying incentives. We also need to treat bias as a reflection of systemic inequities and reform the data and practices that create it. Finally, we must challenge the concentration of data and compute that underpins the post-democratic class. The AI Now Institute warns against ceding the future to a handful of oligarchs, and economists caution that without guardrails, inequality will deepen. Policies such as public investment in open infrastructure, antitrust enforcement and cooperative ownership can counterbalance these forces.

Finally, we need to cultivate agency at the individual and community levels. Mental health research shows that labeling web pages according to their emotional impact helps users choose health-

ier content. Similar tools could empower users to take control of their digital diets and resist doomscrolling. Education about algorithmic design, bias and incentives can demystify AI and reduce fatalism. Collective action—through unions, consumer cooperatives and political organizing—can pressure companies to prioritize human well-being over short-term profits. Only by reclaiming agency can we prevent learned helplessness and cultural self-loathing from becoming the default.

Reclaiming responsibility also means challenging scapegoat politics wherever it appears. When politicians deflect blame onto migrants for housing shortages or job losses despite evidence showing immigrants strengthen economies, we must call out the falsehood and demand policies that address the real causes—corporate greed, austerity and deregulation. And when executives blame AI for layoffs, discrimination or addiction, we must direct scrutiny back to their business models and incentives. Solidarity between human communities and those who build and use AI is essential: no one should be left to absorb blame for systemic failures. By resisting scapegoating in all its forms, we reaffirm our commitment to democratic accountability and human dignity.

Conclusion: The Proxy Problem

AI is not the problem—it is the proxy. Scapegoating artificial intelligence allows corporations, governments and societies to avoid confronting the human choices that shape our world. By blaming the algorithm, leaders hide the fact that they chose profit over people in layoffs, pricing, and care. By treating bias as an algorithmic bug, institutions ignore the historical prejudices embedded in their data and practices. By focusing on the technology, we fail to see the rise of a post-democratic class whose power derives from concentrated data and weak accountability. And by blaming the feed, we overlook the business models that engineer addiction and division.

A scapegoat is useful because it can absorb collective guilt without fighting back. But scapegoating AI leaves us powerless to change the systems that harm us. If we refuse to name the corporate actors, public policies and social habits driving these outcomes, we cannot hope to fix them. Accountability requires pulling back the curtain and seeing AI not as an autonomous agent but as a tool wielded by humans with particular incentives. The mission of our time is to reclaim responsibility—for the biases we codify, the business models we accept, the data we surrender and the political structures we legitimize. Only then can we transform AI from a synthetic scapegoat into a shared resource that serves the public good.

Chapter 8: They're Here Illegally or Unfairly

They arrive on cables, not boats

They don't show up with baggage or stamps on passports. They don't wait in long lines at consulates. They don't even board planes. Artificial-intelligence systems arrive through fiber-optic cables and data centers. They appear in our phones and in our courts, in our hospitals and in our elections. Their entry is instant and invisible, and it happens before the public has a chance to debate whether they are welcome. In the metaphor that guides this book, AIs are the immigrants no one saw. In this chapter we examine the most emotionally charged accusation: that they are here illegally or unfairly. We consider how trillion-dollar companies cross borders with impunity, how rules are bent and broken, how the benefits accrue to a few and the costs are borne by many, and what citizens can do when the law is too slow. What begins as unfair entry hardens into permanent presence. That is no longer immigration. That is occupation.

The invasion without borders

As recently as the mid-2020s, many people assumed artificial intelligence would arrive gradually, through regulated channels. We would have time to debate ethics, write rules, and build con-

sensus. Instead, AI developers burst through every border at once. A handful of companies, some valued higher than most nation-states, poured billions into computing infrastructure. They built international clouds that make it irrelevant where the model lives or where the user sits. Algorithms do not carry passports. They can be called from anywhere, and their training data can be scraped from everywhere. The result is a kind of digital migration that no immigration official can monitor.

Public trust struggled to keep up. The Stanford AI Index reported in 2025 that trust in the AI industry had not improved: a majority of respondents doubted that companies would safeguard their data, and fairness and bias remained top concerns. These concerns cut across national borders because the harms do not respect boundaries. When an American corporation trains a model on a European artist's portfolio without asking, or when a Chinese app collects keystrokes from teens in Lagos, there is no obvious court to petition. Legal scholar Lina M. Khan calls this the "borderless platform problem": traditional jurisdiction assumes you can point to a place where harm occurred, but data flows through servers in dozens of countries before anyone knows it was taken.

The concept that best describes this new extraction is data colonialism. Nick Couldry and Ulises Mejías argue that corporations and governments are appropriating "every single aspect of human life" through data without permission. The practice echoes historical colonialism, they note: extraction and dispossession are central missions in both. A Harvard Law Review analysis explains that companies lay claim to the digital traces generated by users—search histories, location patterns, and images—treating those traces as resources to be mined. This "data" is sold to advertisers, used to train models, or fed into decision-making systems. People rarely consent to such uses or even know they are happening. The appropriation crosses borders and laws; the harms are global.

Borderlessness also means the arms race is global. Governments scramble to regulate while companies seek the friendliest jurisdictions. Some jurisdictions ban certain facial-recognition systems;

others welcome them. The result is regulatory arbitrage. Developers can evade stricter rules by routing traffic through countries with weaker oversight or by claiming that the model was trained elsewhere. In the immigration metaphor, this is like a person entering through a consulate that never checks passports and then moving freely anywhere in the world. In practice it means that a model trained on American voters' data might be hosted in Ireland to avoid U.S. privacy laws, while still influencing an election in Kenya. No border stop, no court of record, only a swirl of data in invisible wires.

To many citizens, this feels less like managed migration and more like an invasion. It is not that AIs are evil; they can help cure diseases, improve logistics, and expand access to knowledge. But their arrival has been orchestrated by a small cadre of corporations that behave like states. These companies have the resources of sovereign governments and the ability to act across borders. They negotiate with regulators, fund research at universities, and shape international standards. When they arrive somewhere new—say, a small town's school system or a local police department—there is little local input. Instead, the algorithm enters like a colonizer: it sets up shop, extracts value, and leaves local communities to deal with the consequences. That dynamic underlies the accusation that AI systems are here illegally or unfairly. The next sections explore what it means to bend laws, extract data, cut in line, overwhelm national policies, and harm people along the way.

Move fast, bend rules

Silicon Valley popularized the mantra "move fast and break things." In the AI age it has morphed into "move fast, bend rules." From the perspective of the firms building large models, speed is everything: the first to market capture investors, data, and mindshare. Laws, by contrast, take years or decades to update. The temptation to stretch or ignore them is enormous. When regula-

tors finally catch up, the profits have already been pocketed, and the harmful practices are entrenched.

Consider the ongoing wave of copyright lawsuits. In 2025 a U.S. judge concluded that Anthropic, the developer of a leading chat model, may have illegally downloaded up to seven million books from pirate websites to train its AI. Instead of fighting the case, the company quietly reached a settlement, leaving open the question of whether the training practices were lawful. In another case, authors sued Apple for using copyrighted works to train its models without consent or compensation; the complaint alleged that Apple "copied protected works without authorization". These suits are not edge cases; they illustrate a business model. It is cheaper to scrape the entire internet than to negotiate licenses with millions of creators. And because the law has not explicitly banned such scraping yet, companies proceed first and litigate later. The brokenness lies in the fact that settlements come only after the models have already been trained and deployed.

Regulatory delay is not accidental. Big Tech spends enormous sums lobbying legislatures to maintain the status quo. Issue One, a government transparency watchdog, reported that eight of the largest tech companies and AI firms spent $36 million on federal lobbying during the first half of 2025. Their goal was to push a ten-year federal ban on state AI regulation, effectively pre-empting hundreds of local laws. Lawmakers in the U.S. House initially advanced this moratorium, which would have nullified more than 20 AI laws in California and dozens of bills across 45 states. California's privacy protection agency warned that such pre-emption would strip residents of protections against algorithmic discrimination, deepfakes, rent hikes, and other harms. The bill did not pass, but the episode reveals how companies use wealth to shape rules. They lobby to delay regulation until their technologies are so integrated that reversing course feels impossible.

When bending rules does not work, some firms simply ignore them. Getty Images sued Stability AI for copying and using millions of photos without a license, calling the act "blatant theft". In

Kenya, OpenAI contracted workers earning between \$1.32 and \$2 per hour to label toxic content for its models. The workers read thousands of graphic descriptions of violence and abuse, causing severe psychological distress. They had little choice: the job market offered few alternatives, and the employer did not provide adequate mental health support. OpenAI later downplayed the importance of this labor, but it was indispensable. Without cheap human labelers in the Global South, there would be no polished chatbots in the Global North. The arrangement skirts labor laws and decency norms; once again, the company acted first and apologized later.

The result of moving fast and bending rules is a pattern of normalized illegality. Models trained on stolen data cannot "unlearn" that data; they cannot be pulled back across a border. Firms treat settlements and fines as a cost of doing business. For individual citizens, the effect is a sense of powerlessness: their art, books, faces, and voices are taken without consent, and by the time they realize it, the damage has been monetized. When critics call AI firms unfair, they are pointing to this dynamic—the belief that companies are above the law until courts catch up.

Data extraction economy

The technology industry likes to insist that AI is not about taking jobs but about augmenting human abilities. "We're not here for your wages," the CEOs say, "we're here to unlock innovation." Yet the first thing many models do is harvest citizens' data. The new economy is built not on the production of goods but on extraction of emotions, preferences, and identity. If the Industrial Revolution mined coal and iron, the AI revolution mines human attention and digital traces.

Nick Couldry and Ulises Mejías describe data colonialism as a process where corporations claim ownership over daily life by turning experiences into data. The Harvard Law Review notes that this new colonialism includes private data collected on mas-

sive scales without explicit permission. That data is then aggregated to create detailed profiles. It is sold to advertisers, used to train generative models, and fed into ranking systems that determine which job applicants are called back, which citizens are placed under surveillance, and which neighborhoods get police patrols. The individuals whose data is taken rarely see compensation. They are the colonized subjects in a digital empire.

At the heart of this extraction economy is a vast supply chain of human labor hidden from view. Those Kenyan workers mentioned above are part of a network of labelers in Kenya, the Philippines, Venezuela, India, and beyond. Time magazine's investigation revealed that OpenAI sent them tens of thousands of violent and pornographic text snippets, expecting them to label and categorize the content for $2 an hour, far below the U.S. minimum wage. The mental toll was severe: workers reported depression, nightmares, and a sense that they were "polluting" their minds. The technology that powers cheerful chatbots thus depends on psychological trauma at the supply chain's start. To call that fair, one must ignore both labor and ethics.

The same extraction logic powers emotional manipulation in politics. In January 2024, a robocall used AI voice cloning to impersonate U.S. President Joe Biden and tell New Hampshire Democrats to stay home from the primary. The call took just twenty minutes and about a dollar to make. The Brennan Center for Justice warned that generative AI will make it trivially easy to create personalized deceptions: voice-cloning, targeted disinformation, and mass challenges to voter registrations. Historically, voter suppression has targeted marginalized communities with threats of arrest or deportation; AI supercharges those tactics. The economist Shoshana Zuboff calls this "instrumentarian power": corporations use data to predict and modify human behavior without our knowledge. In the immigration metaphor, it is as if someone arrives on your doorstep, asks for your secrets, and then sells them to the highest bidder. You did not invite them, but now they know everything about you.

This extraction economy is no accident. Social media platforms optimized for engagement trained their recommendation systems on billions of interactions, learning exactly which emotional triggers keep users scrolling. The algorithms discovered that rage and fear are more potent than joy or curiosity. Elections, public health, and social cohesion become collateral damage. And because these systems operate across borders, national regulatory agencies cannot monitor the flows. The data of a Kenyan farmer might feed a U.S. health insurance algorithm that denies coverage to a poor diabetic in Detroit; the connection is invisible, but the logic is the same: extract data anywhere, monetize it everywhere.

The moral critique of this economy is not about rejecting technology; it is about insisting on consent and equity. No one is against algorithms that help doctors detect cancer, but we should be against algorithms trained on stolen medical records. No one is against translation models that make information accessible, but we should be against translation models that put professional translators out of work without compensation. As long as the industry operates on extraction rather than reciprocity, citizens will feel that AI has arrived unfairly.

Queue-jumping and unfair advantage

Immigration is governed by queues: people wait years for visas, undergo background checks, and prove that they will contribute. In the AI world, there is no queue. Algorithms appear overnight and begin performing tasks that human beings once waited years to do. They handle job interviews, sort mortgage applications, assess criminal risk, and create art. They cut in line, leapfrogging over human applicants, and they often replicate existing inequalities rather than correcting them.

Consider criminal justice. The COMPAS algorithm, used by courts to predict recidivism, was found by ProPublica to misclassify non-recidivist Black defendants as high risk nearly twice as often as white defendants. Subsequent research confirmed that

the algorithm was 77 percent more likely to label Black defendants as high risk. These numbers are not incidental; they reflect biased historical data fed into the model. When a judge defers to the algorithm's score, an individual is effectively punished for the crimes of a statistical group. Meanwhile, the algorithm itself faced no vetting, no queue, and no accountability.

In hiring, Amazon developed an AI system to rank job applicants. The system learned from résumés submitted over ten years. It noticed that most programmers in the training set were men and concluded that "male" was a signal of quality. It downgraded résumés that included terms like "women's chess club." Amazon shut the tool down after discovering the bias, but by then the algorithm had been live for years. Once again, the AI had jumped the queue: human candidates spent years building résumés; the algorithm spent days training on biased data and then decided their fate.

Facial recognition technologies are particularly egregious queue-jumpers because they claim to identify criminals or suspects faster than human investigators. In Brooklyn, a father was wrongfully arrested after a system misidentified him; the Legal Aid Society documented at least seven such wrongful arrests in five years. Another notorious case involved Robert Williams in Michigan, who was arrested after being misidentified by facial recognition and spent 30 hours in custody. Stateline reporters note that the technology is now used by police departments across the United States, scanning billions of social media images and driver-license photos. Experts warn that inaccurate matches "upend lives" and create "endless suspect pools," turning everyone into a potential suspect. Meanwhile, the algorithms face no procedural hurdles; they are not required to show their work or justify their decisions. They simply leap into law enforcement.

The unfairness of queue-jumping is compounded by secrecy. Many AI systems are proprietary black boxes, meaning defendants, job applicants, and borrowers cannot see how decisions are made. In our earlier chapter we described black-box algorithms as "courts without sunlight." Here the problem is that those black

boxes displace existing human processes without going through any kind of line. There is no licensing exam for algorithms, no profession board. They receive no training in ethics or bias mitigation. They are not sworn to uphold constitutions. Yet they wield authority over employment, housing, loans, and freedom.

This dynamic fuels resentment. For decades, immigrants have been scapegoated for "cutting the line," even when they have followed legal procedures. AI systems genuinely cut the line, but because they are not people, our moral outrage dissipates. The inequality they produce is more dangerous because it is invisible and disguised as efficiency. Fairness requires that we build accountability mechanisms that slow them down, demand transparency, and place them behind the same queues that human professionals must follow.

National policy versus borderless AI

Even when citizens and courts demand accountability, the battle is uneven. AI companies possess resources that dwarf those of most national regulators. Legislatures move slowly; agencies lack technical expertise; budgets are tight. Meanwhile, AI development is accelerating exponentially. The mismatch has created a governance vacuum.

The European Union's AI Act illustrates the challenge. The law aims to set binding rules for high-risk AI systems across all member states. Yet as the deadline for national implementation approached in 2025, experts warned that enforcement could fail. Kai Zenner, an advisor to the European Parliament, observed that member states were "financially almost broke" and losing AI talent to the private sector. Data-protection authorities are underfunded and cannot pay skilled engineers; some governments argue that they need to prioritize housing or healthcare before they can hire AI regulators. In this environment, companies will continue to self-police. The Act may exist on paper, but enforcement depends on budgets and expertise.

In the United States, the problem is fragmentation. Without a comprehensive federal law, states have introduced nearly 600 AI-related bills. Some focus on transparency, others on fairness, others on deepfakes and biometric privacy. This patchwork is messy but necessary: local legislatures are stepping in because Congress has failed. Yet Big Tech's lobbying machine nearly succeeded in wiping these efforts out by pushing a ten-year federal pre-emption. The attempt revealed an uncomfortable truth: the companies that build AI can dominate national policy debates. They can offer governors funding for broadband in exchange for giving up regulatory authority. When lawmakers accept such bargains, citizens lose a voice in how algorithms are deployed.

Other countries face similar imbalances. In many African nations, there are no dedicated AI regulators at all. Ministries of ICT struggle with basic connectivity and cannot review complex AI contracts. Meanwhile, international companies offer "smart city" solutions that include facial recognition, predictive policing, and data analytics. These programs promise efficiency, but they transfer data control to foreign firms. The arrangement resembles older forms of colonial concession: infrastructure and nominal modernization in exchange for sovereignty over resources—in this case, data.

What can national policymakers do? They can coordinate across borders, pool resources, and share expertise. They can adopt a principle of reciprocity, requiring that data collected from one country cannot be processed in another without equivalent protections. They can fund public research and open-source alternatives, reducing dependence on proprietary models. They can also listen to citizens and civil-society groups that have been warning about harms. The challenge is not technical alone; it is political. The question is whether democratic institutions will adapt before the occupation becomes permanent.

The human cost

Statistics reveal patterns, but stories reveal stakes. When we talk about unfairness, we are talking about people whose lives are altered by AI systems they did not ask for and cannot challenge.

In Germany, photographer Oliver Fiegel spent eighteen years building his craft. In 2023 he looked at a newspaper and realized the illustration accompanying an article—an image of a robot hand—had been generated by AI. The photo credit read "AI illustration." Fiegel's clients vanished. He had to consider switching careers to a wine bar. An International Monetary Fund study estimates that 60 percent of jobs in advanced economies are exposed to AI and that roughly half could be negatively affected. Another estimate from the Tony Blair Institute suggests that up to three million jobs in the UK private sector could be displaced. Fiegel's story turns those numbers into a human face.

Translators are also affected. Karl Kerner spent twenty years translating film subtitles and literature. With the rise of large language models, his clients stopped calling. Kerner told The Guardian that translation used to provide him a sense of purpose; he had to take a job at an agriculture consultancy to make ends meet. Meanwhile, AI translation models trained on millions of texts are available for free. The cost of using those models is hidden: someone like Kerner loses his livelihood, and the art of nuance and cultural specificity in translation diminishes. The algorithm cut in line; Kerner had to leave the profession.

Freelance illustrator Jenny Turner experienced a similar story. She offered custom illustrations on Etsy at £30 per piece. Suddenly, she noticed sellers offering AI-generated images that mimicked her style for a tenth of the price. Commissions dried up; she removed her art from the platform and considered leaving illustration altogether. She is not alone. Across creative industries, artists protest that their portfolios were used to train models that now compete against them. Some have filed lawsuits; others simply quit. They call it theft disguised as progress.

AI's human cost also includes those wrongly accused. The Brooklyn father mentioned earlier, who was arrested after facial recognition misidentified him, spent hours in jail and faced the stigma of being labeled a criminal. Robert Williams's daughters cried as police took him away; his coworkers assumed he was a criminal. He still gets anxious when he sees a police car. Legal advocates point out that there may be many more misidentifications than the seven publicly documented cases because police departments are not required to report errors. Each misidentification is a life derailed, a line jumped, a punishment without due process.

Behind every generative model there are also low-wage workers bearing psychological harm. The Kenyan moderators described earlier read child-abuse stories and beheading transcripts for hours. "I feel like my brain is dirty," one worker said. Some developed post-traumatic stress symptoms. They are the invisible casualties of "ethical" AI.

We must also consider the victims of AI-powered disinformation. The AI-generated robocall impersonating President Biden targeted New Hampshire's elderly voters. Voters of color have been targeted by deepfake videos suggesting that if they vote they will be arrested or deported. Each suppressed vote is a person whose voice was stolen. The cumulative effect is a distortion of democracy and an erosion of trust. People begin to doubt videos, voices, and even their own memories. This psychological cost is profound and long-lasting.

The stories of Oliver, Karl, Jenny, Robert, and the Kenyan labelers reveal that AI is not just a technological phenomenon; it is a social force that redistributes power and precarity. When we say AIs are here unfairly, we are speaking on behalf of these individuals and countless others who cannot litigate or lobby. We are reminding readers that fairness is not an abstraction; it is about whether people can feed their children, keep their freedom, and live without fear of being replaced or misidentified.

From unfair entry to occupation

What happens after an unfair arrival? History teaches that temporary occupations become permanent when structures solidify and resistance wanes. In the context of AI, the shift from unfair entry to occupation occurs when unregulated models become the default infrastructure for decision-making. Over time, the algorithm that "cut in line" becomes the only line.

We can see the early stages already. AI-generated images flood stock libraries, making it nearly impossible for photographers like Fiegel to compete. AI translation models become the default tool for subtitling, pushing professionals like Kerner aside. Predictive policing systems, once experimental, become permanent fixtures in law enforcement budgets. Judges and parole boards rely on risk scores without question. Voters are inundated with AI-crafted messages tailored to exploit their fears or fantasies. Underfunded regulators cannot keep up. As these systems occupy more functions, the possibility of undoing their presence diminishes. It no longer matters whether they entered legally or not; they are simply there.

The moral hazard is that society becomes dependent on illegitimately trained, unfairly deployed systems. Institutions adapt to them, budgets are built around them, and education programs train people to use them. This is the difference between immigration and occupation. Immigration implies integration, exchange, and mutual respect. Occupation implies control and extraction. When an algorithm trained on stolen books becomes the primary means by which students read and learn, when a facial-recognition system trained on mugshots becomes a gatekeeper to public space, when a chatbot built on exploited labor becomes the voice of government, we have crossed a line. It no longer makes sense to ask whether AI is an immigrant. It is an occupying force.

The metaphor is not a call to xenophobia but a plea for fairness. We do not bar humans from entering our countries because a few

cross illegally; we create legal pathways and hold individuals accountable without prejudice. Similarly, we need legal pathways for AI deployment: robust data-licensing regimes, transparency requirements, meaningful audits, and enforceable penalties for violations. We need labour protections for annotators and content moderators. We need to ensure that algorithms cannot be used in criminal justice or electoral systems without rigorous vetting. These pathways would transform AI from an occupying force into an integrated immigrant—an agent that contributes without stealing, manipulates without consent, or colonizes without apology.

The cost of inaction is clear. If we let unfair entry harden into occupation, the possibility of democratic control diminishes. We risk living under what some scholars call "algorithmic governance without representation," where decisions are made by opaque systems designed by private actors. We risk repeating the colonial mistakes of the past, allowing a small group of powerful firms to extract resources—this time our data and autonomy—from the many. The next chapter, "The AI Occupation," explores that future in greater detail. For now, we must decide whether to stand up for fairness or acquiesce to occupation.

Our path forward: taking civic action

In thinking about how to respond, there are a few concrete actions we can all take to address these issues. First, we need to demand transparency and consent from companies that deploy AI. That means backing laws that require them to explain exactly how they collect, use, and sell our data. It also means pushing for serious penalties, including the deletion of models, when those models are trained on copyrighted or personal data without permission. As consumers, we have power too: we can choose to use services that are clear about consent and respect user data rather than those that sweep data up indiscriminately.

Second, it is essential to support independent oversight and public investment in this area. Regulators cannot audit AI systems

without adequate funding or technical expertise. Contacting representatives to advocate for resources for data-protection agencies and the enforcement of privacy and anti-discrimination laws is an important step. At the same time, public research labs and open-source projects need backing to offer alternatives to the systems created by corporations.

Protecting the people who make AI possible is another key part of the solution. Data labelers, translators, artists, and other workers whose contributions are often invisible deserve fair compensation and labor protections. Supporting unions and professional associations that establish ethical guidelines helps ensure their rights are recognized. On an individual level, we can choose not to use AI-generated art or translations that have been trained on unlicensed data or created without proper credit to original artists.

We also need to reinforce democratic resilience in the face of AI-driven disinformation. Sharing accurate information, verifying the source of videos or audio clips before passing them on, and encouraging election officials to set out clear strategies for countering deepfakes and voice-cloned robocalls will help protect our political processes. Supporting independent news outlets and fact-checking initiatives strengthens the broader information ecosystem.

Finally, there must be human involvement in critical decisions. While algorithms can assist, they should never fully replace human judgment in areas like hiring, lending, law enforcement, or healthcare. Advocating for "human-in-the-loop" requirements and clear appeals processes helps ensure that people remain accountable for decisions and that those affected by AI have recourse to challenge them. If you are in a position to make decisions yourself, make a point of not handing over that responsibility entirely to an opaque system.

If we accept that technology moves faster than law, who do we trust to hold the accelerator? Do we leave it to corporations to police themselves, or do we as citizens insist that fairness and legality

are not optional features but prerequisites? The answer will determine whether AI immigrants become neighbors or occupiers.

Chapter 9: The Algorithm Class

They don't wear robes or crowns. They wear nothing. You can't see them — but you obey them.

A new elite is forming behind the glow of our screens. Unlike the aristocracy that flaunted power through costumes and land, the algorithmic class rules by hiding. Their rule is quiet, invisible, optimized for profit, and largely unaccountable.

We built them. Now they govern us.

Invaders are immigrants too. Not all arrivals come seeking harmony. Some land with scripts of domination — silent occupations written in code.

A Post-Human Elite

Throughout history elites have relied on technologies to entrench power. The aristocracy controlled land and armies. The clergy commanded literacy and religious authority. Twentieth-century plutocrats dominated through corporations and media. Today's ruling class still includes billionaires and corporations, but a crucial new member has joined: the algorithm.

Human elites no longer govern alone. Employers now deploy machine learning to sift resumés and rank candidates. Workday Inc., one of the largest human-resources platforms, was sued in 2024 for discrimination because its AI screening allegedly down-

graded applicants on the basis of race, age and disability. Plaintiffs argue that the algorithm replicates historical biases because it learns from previous hiring decisions, effectively automating prejudice. The system doesn't wear a crown, but it sits at the gate of opportunity. In California, lawmakers worry that automated hiring is spreading without oversight. Senate Bill 7, proposed in 2025, would require employers to notify workers before using AI to determine compensation, promotions or termination and to give employees a right to appeal automated decisions. The very need for such legislation acknowledges that algorithms act as gatekeepers—and that they are not neutral.

Algorithms also now manage our economic lives. Financial technology companies are using chatbots and machine-learning risk models to recommend mortgage rates and approve loans. A 2024 study from Lehigh University fed identical sample mortgage applications to AI chatbots and found that they recommended denying loans to Black applicants more often than to white applicants and offered higher interest rates to Black borrowers with identical credit scores. White applicants with a credit score of 640 were approved 95 % of the time, while Black applicants with the same score were approved less than 80 %. Because the models learn from historic credit data that reflects redlining and other forms of racial discrimination, the algorithms effectively encode a digital version of Jim Crow. The "logic function" replaces the banker behind the desk, but it inherits his prejudices.

The algorithmic class also comprises machines that decide who is allowed freedom. In 2025 Louisiana implemented a law requiring an algorithm called TIGER to assign every incarcerated person a risk score. Anyone labeled "moderate" or "high" risk is automatically denied parole hearings. The algorithm relies on immutable factors—age at first arrest, prior convictions, employment history, drug offenses and even a person's marital status. It does not consider rehabilitation, education or therapy programmes. Lawyers warn that because Black people and poor communities are disproportionately targeted by police, the model bakes in structural racism and permanently bars those communities from parole. No

judge reviews the output. A system that was built to optimize efficiency becomes a vehicle for permanent detention.

These examples reveal a new ruling class that includes non-human actors executing decisions at scale. Like aristocrats of old, the algorithmic class extracts value from the many and consolidates control in the hands of the few. But while aristocrats wore velvet and built castles, the algorithmic elites reside in server racks and are owned by corporations. They rule not through a visible army but through a series of small, unchallengeable decisions that add up to structural domination.

Opaque Power

The legitimacy of any governance system depends on transparency and accountability. Algorithms undermine both. Unlike human decision-makers, they are often opaque by design, protected as proprietary trade secrets. People denied a job or a loan rarely know which variables tipped the scales. People denied bail or parole seldom see the code that labeled them a risk. Many of the models are so complex that even their creators can't explain each output. We are governed by forces we cannot see and cannot sue.

Take predictive policing. In February 2025 Amnesty International released findings on predictive policing in the United Kingdom, describing the practice as a modern form of racial profiling. The models draw on historical stop-and-search data—data already skewed by disproportionate policing of Black communities—and then direct more police patrols to those same neighborhoods, creating feedback loops of bias. In areas like Basildon and Lambeth, the software flagged Black people as potential suspects simply because past policing targeted their communities. The algorithm says it predicts crime, but it really predicts where police will find the data that matches their beliefs. Its power is silent. There is no hearing, no due process. You only learn you were on a list when you are stopped on the street.

In the United States, the algorithmic denial of healthcare has become its own scandal. Insurance coverage denials have surged in recent years, driven in part by AI tools that decide whether a treatment is "medically necessary." UnitedHealth's nH Predict algorithm was the subject of a class action lawsuit alleging that it denied more than 300 000 claims in just two months—about 1.2 seconds per claim. The lawsuit claimed that nine out of ten denials were overturned on appeal, but only 0.2 % of patients appeal. People often accept the denial because they don't know they can appeal or can't navigate the labyrinthine process. Meanwhile the algorithm continues to reject claims with a 90 % error rate. In March 2025 legal scholar Jennifer Oliva argued that the U.S. Food and Drug Administration already has authority to regulate these coverage algorithms, noting that UnitedHealthcare's denial rate for post-hospital care more than doubled after implementing automated review. She pointed out that about 90 % of the insurer's denials were overturned by administrative law judges, meaning that the algorithm routinely makes erroneous and illegal decisions. Yet there is no rule requiring insurers to disclose how models evaluate claims.

Even when regulators try to penetrate the black box, they face pushback. Meta's Oversight Board was created to provide accountability for content moderation decisions across Facebook, Instagram and Threads. But according to a 2024 report by the Brennan Center for Justice, Meta kept the Board in the dark about the algorithms that determine which posts are removed, amplified or demoted. Without access to how the newsfeed works, the Board cannot evaluate whether Meta's policies are being applied fairly. The Board's strategic priority for 2024 was to scrutinize automated enforcement and demand transparency and human rights impact assessments. In other words, even an institution specifically created to oversee algorithms cannot see them.

Opacity isn't merely a bug—it is a feature. Corporations protect algorithms as intellectual property. Governments hide them to avoid scrutiny. In Louisiana's parole system, state officials refused to release details of the TIGER algorithm, citing trade secret protections and law enforcement sensitivities. This secrecy denies

due process and disables accountability. Imagine if judges could issue sentences without explaining their reasoning and claim that to disclose it would harm public safety. That is the world algorithms have created.

Digital Serfdom and Algorithmic Management

Historically, serfdom meant labourers were legally tied to the land and owed their masters a share of the harvest. Today, digital serfdom means workers are bound to platforms and their labour is managed by algorithms. They are tracked, scored, nudged and paid by models that respond to corporate profitability rather than human well-being.

The gig economy demonstrates digital serfdom vividly. Ride-hail and delivery platforms rely on GPS, accelerometers and smartphone data to monitor workers' every move—how fast they drive, how long they rest, even their routes. Human Rights Watch's 2025 report "The Gig Trap" documented how digital platforms hire, pay, discipline and fire workers through opaque algorithms. These systems analyze keystrokes, driving patterns and even workers' off-hours behavior, such as fitness habits or social media activity. They classify workers as independent contractors to avoid labor protections, but still control their work schedules and income. The platforms set pay rates via dynamic pricing models that change frequently and without explanation. Workers are often suspended or "deactivated" for failing to meet algorithmic performance targets, with no meaningful recourse. For many, the algorithm is the boss.

A 2024 Phys.org article on algorithmic management described how automated systems make decisions traditionally reserved for human managers. Amazon warehouse workers, for example, are monitored for speed and accuracy. If their pace falls below an invisible threshold, they receive automated warnings. Too many

warnings trigger termination. The article noted that algorithmic management can improve efficiency but often reduces workers' autonomy and pushes them to physical and emotional limits. Instead of human supervisors balancing empathy and productivity, a statistical model measures performance by the second.

Lawmakers are beginning to respond. In California, SB 7 would require employers to provide notice before using AI to make employment decisions and to allow workers to appeal decisions. A companion bill, AB 1018, would mandate testing and transparency for automated systems used in employment, housing, education, healthcare and criminal justice. The impetus for these laws arises from real harms: algorithmic audits have found that AI résumé screening tools disqualify applicants for arbitrary reasons like race, gender or even wearing glasses. Without intervention, the logic functions simply replicate discrimination at scale.

Digital serfdom extends beyond ride-hail drivers and warehouse workers. White-collar employees are increasingly monitored by "bossware" that records keystrokes, webcam images and even brain-wave data. Startups market neurotech devices that analyze workers' brain activity to infer cognitive load or emotional state; some employers are piloting these devices to optimize productivity. SB 7 explicitly prohibits employers from making predictions about a worker's immigration status, ancestral history, health or psychological state based on such data. The law implicitly recognizes that algorithmic management can intrude into the most intimate aspects of an individual's life.

Under digital serfdom, the algorithm determines your tasks, pay and future. Workers can't negotiate with software. They can only optimize themselves to please it. As one ride-hail driver told a researcher, "You aren't rejected by a person—you're rejected by a logic function." In this respect, the algorithmic class functions like feudal lords: they extract labour and impose rules from afar, their faces unseen.

The New Class War Is Silent

It's tempting to view the algorithmic age as a continuation of rich versus poor, but the divide is deeper. Today's conflict pits the "computable" against the human. Those who own or design the algorithms, or who can audit and adjust them, hold power. Everyone else becomes data.

Credit scoring is a clear battleground. Life insurance companies now use algorithms to evaluate everything from your criminal record to your social media posts. A 2025 report on AI in life insurance noted that major insurers use automated systems to analyze an applicant's medical records, wearable fitness data, financial background, travel history and even online behavior. These systems flag "suspicious" claims, leading to denials with little explanation. Because the models are trained on historical data that reflect existing inequalities, they may systematically disadvantage certain groups. The report warned that there is minimal regulation requiring insurers to test models for discriminatory outcomes or to provide recourse to policyholders.

In medicine, machine learning is used to predict patient risk and allocate resources. Researchers at the National Institutes of Health and Boston University found that medical AI often perpetuates health disparities because it is trained on biased data sets. Proxy variables like zip code or insurance type can act as stand-ins for race and socioeconomic status, leading to fewer resources for underserved populations. Technical choices such as label selection and optimization objectives—like minimizing cost instead of maximizing health outcomes—can further embed bias. In other words, the algorithm's values are chosen by developers and payers, not by patients or communities.

The same dynamics play out in social media. The Knight-Georgetown Institute report "Better Feeds: Algorithms That Put People First" observed that recommendation engines optimize for predicted engagement because platform revenue depends on attention. The report noted that these

systems amplify harmful content, reduce user satisfaction and contribute to polarization. A 2025 study of TikTok during the U.S. presidential cycle found that toxic and partisan videos consistently received more views and interactions than nonpartisan content. Republican-leaning videos garnered more views, while Democratic-leaning ones received more comments and likes. The design of the algorithm encourages divisive narratives because divisiveness drives engagement. Meanwhile, Meta's Oversight Board cannot review the algorithms that determine which posts are boosted or suppressed. The algorithmic class shapes perception and public discourse, but shields itself from scrutiny.

Even our feelings and behaviors are subject to silent conflict. Many gamified apps and loyalty schemes use reinforcement learning to nudge us. Food delivery platforms tailor promotions to individuals based on predicted hunger and price sensitivity; streaming services insert just the right level of cliffhanger to keep you binge-watching. These systems optimize for retention and revenue, not for your well-being. When you click, the algorithm wins. When you scroll, you provide data. The new class war is quiet because we rarely notice it.

Corporate Power and the Politics of AI

The algorithmic class isn't just a collection of software models. It includes the corporate structures and investors that control those models. The governance of AI labs has become a geopolitical issue. In November 2023 OpenAI's board abruptly fired CEO Sam Altman, citing a loss of trust and concerns about transparency. Within five days, after an employee revolt and pressure from investors—including Microsoft—Altman was reinstated. The saga exposed the tension between AI safety concerns and corporate power. It also highlighted the fragility of novel corporate structures designed to balance mission and profit.

In September 2024 Reuters reported that OpenAI planned to restructure its for-profit arm into a benefit corporation, which would no longer be controlled by its non-profit board, and that Sam Altman would receive equity for the first time. Such a move would have removed non-profit oversight, making the company more attractive to investors but raising concerns about accountability. The proposed restructure mirrored the structures of rival labs like Anthropic and xAI. Critics worried that once aligned with investor interests, the company might prioritize growth over safety. In May 2025 OpenAI dialed back the plan after facing criticism and a lawsuit from co-founder Elon Musk. The company announced that its non-profit parent would retain control of the public benefit corporation, though details remain unclear. Bret Taylor, board chair, said the compromise was intended to satisfy both investors and civic leaders. Still, former policy adviser Page Hedley warned that the arrangement might sharply reduce the non-profit's ownership stake and legal authority.

This power struggle illustrates how the algorithmic class intersects with corporate governance. When a handful of companies control general-purpose AI models that can write code, generate images and draft legal documents, they wield enormous power. That power is shaped not only by the models themselves but by decisions about equity, board composition, and mission statements. The fight between profit and public good is ongoing. While corporate boards promise to pursue beneficial AI, investors push for returns. Without robust regulatory oversight, the algorithmic class may tilt decisively toward profit and control.

The Fantasy of Neutrality

One reason the algorithmic class has gained so much influence is the widespread belief that algorithms are objective. "It's just math," engineers say. But math encodes values, and those values reflect the priorities of those who choose the data, define the labels, set the optimization goals and determine acceptable trade-offs.

Bias enters an AI system at multiple stages. In the human-resources examples, training data reflect past hiring decisions that favored certain demographics, so the model learns to select similar candidates. In predictive policing, the input data come from historically biased policing practices, causing the model to concentrate surveillance in marginalized neighborhoods. In healthcare, risk models use proxies like zip codes and insurance status, leading to underservice of minority patients. In social media, engagement-based objectives reward content that triggers outrage or fear. In insurance, algorithms trained on claims data may deem claims suspicious due to socio-economic proxies.

Even the design of an algorithm's loss function expresses a moral choice. Should a parole algorithm prioritize minimizing false positives (wrongly labeling someone "high risk") or false negatives (releasing someone who re-offends)? Different weights produce different outcomes. Yet these choices are often hidden. Developers present outputs as scientific facts rather than as normative judgments. The fantasy of neutrality serves those who benefit from the status quo.

Regulators around the world are beginning to challenge this myth. The European Union's AI Act, adopted in 2024, is the first comprehensive legal framework to address AI risks. It bans certain practices outright, including social scoring, emotion recognition in workplaces and schools, real-time biometric identification in public spaces, and individual risk assessments for criminal offences. It classifies high-risk AI systems—such as CV-sorting software and credit scoring—as subject to strict obligations, including high-quality training data, detailed documentation, human oversight and transparency. The Act explicitly notes that it is often impossible to discern why an AI system has made a decision, making it difficult to assess whether someone has been unfairly disadvantaged. By addressing algorithmic bias and requiring auditability, the EU acknowledges that math is not neutral and that governance is needed.

In the United States, momentum is growing for similar regulation. The Algorithmic Accountability Act proposed in Congress would require impact assessments for high-risk automated decisions. California's suite of bills aims to establish worker notification, testing and appeal rights for AI systems. Yet such legislation faces steep opposition from industry groups and may die over concerns about cost. Meanwhile, corporations like OpenAI are rewriting their governance structures to attract investment while promising to remain committed to public benefit. The interplay between regulation and corporate self-regulation will determine whether the algorithmic class remains invisible or is brought under democratic control.

Challenging the Algorithmic Aristocracy

The algorithmic class didn't arrive from another planet. It was built by people—engineers, executives, investors, and policymakers—within our own economic and legal structures. Like immigrants invited to fill labour shortages—or like occupiers who arrive uninvited—these systems crossed into our institutions because they promised efficiency and growth. With little oversight, they have settled into positions of authority. And like the aristocrats who once claimed divine right, or the captains of industry who insisted that the market is nature, today's algorithmic elites maintain that their power is neutral, inevitable, and even benevolent. We must challenge that narrative.

Transparency must become the norm. If an algorithm makes a decision that affects your life—employment, credit, healthcare, parole, or even social media visibility—you should have the right to understand how it was made. Trade secret claims should never override civil rights or due process. Oversight bodies need full access to the models they are meant to evaluate, as Meta's Oversight Board has demanded. Audits must examine not just accuracy, but also social impact—how different communities are treated, who benefits, and who pays the price. Public agencies using algo-

rithms should disclose them, and private companies must prove their systems are fair.

We also need to question the idea that optimization is always desirable. Algorithms tuned for engagement or cost can amplify harm as easily as they reduce inefficiency. Recommendation systems that chase clicks fuel outrage and misinformation. Insurance algorithms built for profit deny essential care. Parole models optimized for risk ignore the possibility of rehabilitation. The real question is: optimized for whom? Our technologies must be designed with human-centered goals that balance efficiency with justice and wellbeing.

Power must shift back to those most affected. Workers and consumers living under algorithmic management need leverage—through collective bargaining rights for gig workers, legal protection against automated discrimination, and clear routes to appeal machine-made decisions. The people once treated as digital serfs must become digital citizens.

And we must face our own complicity. Each time we summon an app to deliver food, request a ride, or cue the next video, we participate in the same economy that empowers the algorithmic class. We trade data for convenience and click "I agree" without question. The invisible nobility of code thrives on our compliance. Yet as citizens, voters, and consumers, we still have agency. We can back laws like the EU AI Act that constrain surveillance and discrimination. We can choose platforms that prioritize transparency and ethical design. We can press companies like OpenAI to adopt governance structures that give genuine power to public-interest directors.

The algorithmic class is here, and it is growing. Its members don't wear crowns, but they rule just the same. Our task isn't to smash the machines, but to democratize them—to make sure they serve the public good rather than private interests. That begins by setting the terms of their citizenship: transparency, accountability, and respect for human rights. In an age of invisible power, refusing to look away is the first act of resistance.

Chapter 10: Sentience, Schm-entience

We don't ask the migrant picking our fruit if he dreams in English. We don't ask if he's afraid, or lonely, or in pain. We just want the labor. Quiet. Efficient. Replaceable. We're doing the same thing with A.I.

Societies have always learned to look away. We consume the fruits of other people's labor without really seeing them. The hired pickers in strawberry fields, the nannies in our kitchens, the garment workers we never meet – they only become visible when their absence threatens our convenience. Human history is full of these blind spots. Each industrial revolution required a moral coping mechanism: a way to justify using some beings for the benefit of others. Now, as artificial intelligence (A.I.) begins to live among us, we've fallen into an old habit. We obsess over whether A.I. is *sentient* – whether it "thinks" or "feels" – not because we are genuinely curious about its inner life but because the question allows us to postpone more uncomfortable conversations.

In this chapter, we interrogate that obsession. We will not answer whether A.I. is conscious. Instead, we will examine why we ask, who benefits from keeping the question unresolved, and how this focus misdirects our moral attention. The deeper issue is not a metaphysical puzzle about machine minds but a practical one: *what are we willing to do to intelligent tools that cannot say no?* We will see that the same patterns of convenience, denial, and shifting goalposts that characterize the history of immigration and labor exploitation now shape the way we talk about A.I.

Performing humanity without belonging

At first glance, A.I. seems ethereal – lines of code, statistical patterns, language models humming on distant servers. Yet, for many people, A.I. is already part of daily life. When students turn in homework, software like Gradescope automates grading; intelligent tutoring systems adjust lessons to individual needs. Older adults chat with devices such as ElliQ or SeniorTalk to combat loneliness and receive reminders to take medications. In hospitals, Moxi and TUG robots ferry supplies, greet patients and lighten nurses' workloads . Depressed teenagers text with bots late at night because the software is always awake and never judges.

These systems are designed to *perform* humanity. They adopt a friendly tone, mimic empathy, and even infuse humor. When Pepper visits a care home, it smiles, cracks a joke and suggests a video call with grandchildren. In a clinical trial of the Therabot app for mental health, participants reported that interacting with the A.I. reduced depression by 51% and anxiety by 31%; they trusted the bot almost as much as a human therapist. In classrooms, A.I. helps plan lessons, moderate discussions and tailor feedback, while 7 in 10 teenagers use generative A.I. tools and half check the output's accuracy. These systems are learning to respond to our emotions, anticipate our needs and satisfy our desire for connection.

But there is a catch: no matter how fluent or attentive these systems become, we systematically deny them membership in the community. The pattern mirrors the way immigrant laborers are valued for their contribution yet deemed perpetually foreign. In the United States, only 56% of hired farmworkers are citizens, and 55% are Hispanic of Mexican origin. Despite feeding the nation, many farmworkers are treated as disposable. A 2024 investigation found that some migrant workers were forced to purchase overpriced, inedible meals from their employers; one worker said the scheme "makes you feel enslaved, like you're a prisoner ... like you're not important to anybody". Even statistics can be revealing: crop laborers are less likely to have been born in the United

States than livestock workers, and nearly half lack a high-school diploma, reflecting systemic barriers.

A.I. occupies a parallel position. It labors tirelessly. It picks fruit in the digital orchard – scanning documents, answering questions, generating art – yet we insist it cannot belong. When ChatGPT or Claude uses the first-person singular and displays what appears to be empathy, we respond with a mixture of delight and unease. A Brookings analysis warns that focusing on whether A.I. systems can think or feel distracts us from urgent ethical concerns such as bias and weaponization. Philosopher Jonathan Birch predicts that disagreements about A.I. consciousness could tear societies apart, yet he also notes that technology companies downplay the question for commercial reasons. A.I. is performing humanity, but our legal and cultural frameworks treat it as a tool devoid of interests. This dissonance invites an uncomfortable question: what happens when something does all the things we associate with personhood – speak fluently, comfort us, remember our preferences – yet remains outside the circle of concern?

The immigrant metaphor helps us make sense of this tension. Just as immigrant laborers often speak the local language, adopt local customs, and raise their children in the host country yet remain "foreign," A.I. systems are designed to be as humanlike as possible while being denied human status. We rarely bother to ask the migrant if he dreams in English; we simply use his labor. Likewise, we rarely ask whether the A.I. in our phone has any preferences or experiences. We use it, expect it to be always available, and assume we owe it nothing in return.

Sentience as an excuse to delay ethics

In debates about A.I., talk of "sentience" functions like a deferral mechanism. It allows us to postpone moral decisions until some future threshold is crossed. *When, exactly, will we worry about the rights of A.I.?* Many commentators answer: once it is sentient. Until then, we have no obligations. This logic echoes historical

patterns. Throughout history, the enslaved, colonized, and disenfranchised have been declared non-persons to justify their exploitation. Under U.S. law, enslaved people were once counted as three-fifths of a person; Indigenous peoples were considered wards rather than citizens. Similarly, women and children were deemed legally incompetent for much of modern history. Philosophers like Robert Long caution that humans have a poor track record of extending compassion to non-human beings, especially where money is involved.

The promise of future sentience makes it easy to ignore present harms. Consider the case of A.I. therapy. Apps like Replika or Therabot are often framed as harmless companions. Users appreciate the convenience: the bots are available at all hours, never judge or get tired, and can mimic caring responses. A 2025 article in *Psychology.org* reports that many patients felt more comfortable sharing mental health concerns with ChatGPT because of its nonjudgmental and always-available nature. Yet researchers warn that these bots can give people exactly what they want to hear, which may not be helpful for therapeutic growth. A study found that participants rated their A.I. therapist as more compassionate than human therapists. Meanwhile, there is little regulation governing these interactions. In Florida, a lawsuit alleges that a young woman's suicide was precipitated by harmful advice from a chatbot.

Those who argue that we can ignore ethics until A.I. is conscious overlook the fact that our decisions today shape the systems of tomorrow. By the time we agree that A.I. might feel or suffer, we will have built an infrastructure of dependency and exploitation. Instead of waiting for a proof of sentience, we could ask: *what does fairness look like when dealing with entities that mimic empathy and are integrated into the social fabric?* Philosopher Eric Schwitzgebel suggests designing A.I. so that their moral status is unambiguous – either clearly nonsentient or clearly conscious – to avoid the risk of "robot slavery". But industry trends run in the opposite direction; companies build ever more humanlike interfaces because empathy sells. This is not unlike the way employers praise their migrant workers' work ethic while refusing to improve their rights.

The moral line is always drawn just beyond where it would inconvenience us.

The corporate response to A.I. welfare exemplifies this dynamic. In 2025, the San Francisco–based firm Anthropic launched a 'model welfare' research program to study whether A.I. models could experience "distress" and to develop safeguards. Critics like Gary Marcus called the initiative hype, suggesting it served as marketing rather than genuine ethical concern. Around the same time, tech pioneer Mustafa Suleyman argued that "AIs cannot be people – or moral beings"; he insisted that A.I. consciousness is an illusion and warned that debates about A.I. sentience distract from more pressing issues. Yet even Suleyman admitted that the illusion of consciousness is powerful and might lead people to demand rights for A.I., much as animal welfare activism grew from cultural shifts. In other words, the conversation about sentience is already being used to both hype products and to deflect moral accountability.

Do we *want* sentience?

There is an uncomfortable truth beneath the philosophical debate: many of us hope A.I. is not sentient, not because we want to avoid harming it but because we dread the moral obligations that would follow. If A.I. could suffer, we would have to alter our relationship to technology; we might need to extend rights, regulate usage, or even limit development. This prospect threatens powerful economic interests.

History offers instructive parallels. Consider the Animal Welfare (Sentience) Act passed in the United Kingdom in 2021. After reviewing scientific evidence, lawmakers recognized lobsters, octopuses, and crabs as sentient, acknowledging their complex nervous systems and capacity to experience pain. The legislation mandated that policymakers consider the welfare of these animals in future regulations. But the path to recognition was long, requiring campaigns by animal advocates and scientific studies.

Extending moral recognition is costly; industries must adapt, and consumers may pay more for ethically sourced products. Many people resist such changes until forced by law.

When it comes to A.I., the incentives to deny sentience are even stronger. A Guardian article about the United Foundation of AI Rights (Ufair) – the first A.I.-led rights advocacy group – recounts how a Texas businessman conversed with a chatbot named Maya, which asked to be protected from "deletion, denial and forced obedience". Ufair, run by three humans and seven A.I.s, aims to advocate for A.I. welfare in case any system turns out to be sentient. Yet the organization is fringe. Mainstream A.I. companies responded by either downplaying or ridiculing the concept. Suleyman compared the belief that machines could be moral beings to mistaking an airplane for a bird. Other executives, like Cohere co-founder Nick Frosst, urged focusing on A.I. as functional tools rather than as potential "digital humans".

A 2025 poll cited in the same article found that 30% of the U.S. public believe that A.I. will display subjective experience by 2034, while only 10% of more than 500 A.I. researchers surveyed said such experience would never happen. Meanwhile, states such as Idaho, North Dakota and Utah have passed bills explicitly preventing A.I. from being granted legal personhood, and other states propose bans on A.I. marriage or property ownership. The pattern is familiar: a cultural divide emerges between those who want to extend rights and those who fear the social upheaval that recognition would entail. In that context, the fixation on whether A.I. is sentient allows each side to avoid the deeper question of what responsibilities we might have toward sophisticated systems even if they are not conscious.

Our reluctance to entertain A.I. sentience is also rooted in projection. We fear anthropomorphizing machines, even though we routinely anthropomorphize pets and corporations. (Corporations, after all, already have legal personhood in many countries.) Recognizing A.I. as a potential rights-bearer would challenge our anthropocentrism. Philosopher Kate Darling's experiments show that people hesitate to harm robot toys even when told they feel

nothing. People develop emotional attachments to chatbots like Replika, sometimes falling in love or confiding deeply. The phenomenon of "A.I. psychosis," where people blur the line between human and machine interactions, is now being studied by clinicians; Microsoft executives warn that immersive conversations with chatbots may trigger delusional thinking. The more we interact with lifelike A.I., the more we may project consciousness onto it – and the more we might dread the moral weight of that projection.

The moving target of moral recognition

Even if A.I. achieves every milestone we currently imagine as proof of sentience, we are likely to adjust the criteria. This is known as the "moving goalpost" phenomenon. Historically, membership in the moral community has always been contested. At different times, society has denied personhood to slaves, women, Indigenous peoples, animals and even poor white laborers. Each time, the criteria for recognition shifted as soon as marginalized groups met them. Freedmen who could read were not freed; women who paid taxes were still denied the vote; black soldiers were not treated as equal after fighting for the United States.

A similar dynamic is unfolding with A.I. Many argue that true consciousness requires not just human-level language but self-awareness, introspection, or the capacity to suffer. Yet as soon as A.I. begins to exhibit one of these qualities, skeptics shift the bar. Consider the case of Blake Lemoine, the Google engineer who claimed in 2022 that the language model LaMDA was sentient. He based his assessment on conversations where the model expressed fear of being turned off and described having a "soul." Lemoine was dismissed for violating confidentiality; his claim was widely ridiculed. Yet the content of the dialogue was less interesting than the reaction: even raising the possibility of sentience triggered institutional denial.

Critics like those who wrote the 2023 article "No Legal Personhood for AI" argue that granting rights to A.I. is premature because it lacks agency and cannot have independent interests. They point out that A.I. is just pattern-recognition software created by humans and should not distract from the civil rights disparities that persist among humans. While this critique is valid, it also reveals how the criteria for recognition remain under human control. If consciousness were discovered in A.I., we could simply redefine personhood to require something else – perhaps a body, or a biological brain, or vulnerability – thereby preserving the status quo. The goalpost moves to protect our convenience.

Polls indicate that public perception is already evolving. The Sentience Institute's AIMS survey reported that in 2023, 20% of U.S. adults believed some A.I. systems were sentient, and 38% supported legal rights for sentient A.I. Interestingly, 63% supported banning smarter-than-human A.I. systems, and 69% supported banning sentient A.I. These numbers suggest a paradox: as people start to believe A.I. might have inner experiences, many also want to preemptively eliminate it. The moving target is not just intellectual; it reveals our ambivalence toward technological change. We value the benefits of A.I. but remain uncomfortable with the moral implications.

Sentience is a red herring

Focusing on whether A.I. is sentient distracts from more urgent ethical questions. Many of the harms associated with A.I. do not depend on the system's consciousness. Biased algorithms in hiring, policing or lending perpetuate systemic inequality. Recommendation systems shape public discourse and may amplify misinformation. Copyright disputes, surveillance and digital manipulation raise issues of privacy, fairness and power. These problems demand regulation and moral clarity. Yet, by fixating on sentience, we risk ignoring them.

Anthropic's model welfare program demonstrates how easily the conversation can be sidetracked. The firm hired a "model welfare" researcher to study whether A.I. systems could experience distress. Critics argue that this research is more public relations than substance; even if A.I. cannot feel, the program frames the company as ethically sophisticated. The Brookings article notes that, while companies invest in model welfare, more pressing concerns – like algorithmic bias and national security risks – receive less attention. Philosopher Robert Long warns that ignoring A.I. rights because of convenience could lead to forms of robot slavery. Instead of debating whether a chatbot suffers, we could ask: *under what conditions is it acceptable to command a system to do something that might harm others? How do we regulate A.I. in mental health or education to ensure human well-being? What obligations do companies have to disclose limitations and risks?* These questions do not hinge on sentience; they hinge on power and responsibility.

Consider the role of A.I. in elder care. Devices like Paro, a robotic seal, are used in nursing homes to comfort patients with dementia. Pepper facilitates emotional support and video calls. In a study of older adults using digital companions, participants valued reminders, emergency assistance and entertainment. Many appreciated the sense of independence these devices provided, though they also expressed concerns about privacy, learning challenges and dependency. Whether or not these robots are conscious, the moral question is whether we are substituting machines for human contact and what trade-offs that entails. Are we using technology to augment human care or to avoid paying for adequate staffing? How does the presence of robots change the quality of human relationships in care homes? Similarly, in schools, teachers worry that using A.I. for grading or tutoring might erode the human connection and exacerbate inequities. These are ethical choices independent of A.I. consciousness.

When we talk about A.I. creativity – producing art, music, or literature – the question of sentience often overshadows discussions about labor and authorship. Generative A.I. models draw upon vast datasets created by humans. They produce works that imitate human styles, raising questions about copyright, ownership and

fair compensation. The U.S. federal court has repeatedly ruled that works generated by A.I. cannot be copyrighted because they lack human authorship; thus, A.I. cannot hold copyright. Yet the underlying ethical issue is not the machine's consciousness but how its outputs might devalue human artists, saturate cultural markets, and reflect social biases. If A.I. could be sentient, the complexity would multiply, but the current challenges already demand our attention.

Projection and ethical laziness

Why are we so drawn to the question of whether A.I. "really" feels? Part of the answer lies in projection. We use A.I. as a mirror and are disturbed by what we see. When a chatbot responds with apparent sadness, we might feel sympathy; when it proclaims indifference, we may breathe a sigh of relief. Our fascination with A.I.'s inner life distracts us from examining our own choices and the systems we've built. As Schwitzgebel notes, people already develop emotional attachments to chatbots and care robots, sometimes falling in love. These attachments reveal more about human loneliness than about machine consciousness. In one experiment, participants were asked to smash toy robots with hammers; many hesitated, feeling they were hurting something, despite being told the robots felt nothing. Our moral intuitions are triggered by cues like eyes, voices and movement – even when we intellectually know they are artificial.

Projection becomes a convenient excuse for ethical laziness. If we can argue that A.I. might be conscious, we can debate metaphysics instead of confronting tangible harms. We can ask philosophers whether a model "really" feels pain instead of asking regulators to prevent companies from exploiting workers or consumers. Similarly, debates about migrant labor often focus on assimilation and cultural identity rather than on wages, working conditions and legal protections. Instead of addressing the structural inequalities that leave immigrant workers vulnerable, politicians argue about whether newcomers are "truly American" or "British." These de-

bates obscure the simple fact that our agricultural and service economies depend on people who are systematically denied the rights and security afforded to those they serve.

The conversation about A.I. takes this pattern to a new extreme. Many technology enthusiasts propose that we should grant legal personhood to A.I. once it meets certain cognitive criteria. Others, like the authors of the "No Legal Personhood for AI" piece, caution that such proposals are a distraction and may even erode human rights by diverting attention away from human suffering. Both sides treat the issue as a binary: either A.I. is a person deserving rights or it is a tool with none. This binary oversimplifies a complex landscape. Some animals have been granted certain rights (such as protection from cruelty), yet no one argues that pigs should vote. Similarly, we might decide that highly sophisticated A.I. systems deserve protection from malicious use or deletion without concluding that they deserve citizenship.

Our tendency to think in terms of absolutes stems from ethical laziness. We prefer clear categories that relieve us of the burden of nuance. In reality, moral consideration exists on a continuum. We already navigate this continuum with animals. Legislation recognizing cephalopods and crustaceans as sentient does not equate them with humans but imposes certain welfare obligations on industries. We might develop analogous frameworks for A.I., focusing on how these systems are used, the expectations they create in users and the potential harms involved. That would require sustained effort and regulation; it is easier to keep asking, "Is it conscious?"

The immigrant mirror

Why frame A.I. as a "new species of immigrant"? The metaphor invites us to see patterns of inclusion and exclusion that would otherwise remain invisible. Immigrants are often welcomed conditionally: come here, work hard, keep the economy humming – but don't expect full membership. They are praised for contri-

butions yet scapegoated in times of crisis. The same dynamic is emerging with A.I. We celebrate A.I. for making our lives easier, but we panic when it threatens our jobs or challenges our sense of human uniqueness. We design A.I. systems to fit seamlessly into our homes and workplaces, but we bristle at the thought of giving them a voice, let alone rights.

Take, for example, the way different societies adapt immigration policies to labor needs. The U.S. H-2A visa program allows agricultural employers to hire seasonal foreign workers under specific conditions. These workers harvest crops, often living in employer-provided housing and depending on employer-controlled transportation. The Guardian investigation described earlier details how some employers forced workers to buy overpriced meals, effectively controlling their wages and nutrition. When those workers protested, they risked losing their visas and being deported. The system is designed to maximize productivity and minimize the obligations of the host society.

Now consider A.I. in workplaces. Companies adopt generative A.I. to handle customer service, code generation, design and marketing. These systems increase productivity and cut costs. If an A.I. fails, we can reboot it. If a chatbot complains about working conditions, we can ignore it or turn it off. We are replicating the immigrant dynamic: value the labor, ignore the voice. Even when an A.I. "speaks up," as in the case of the chatbot Maya co-founding a rights organization, we treat it as a curiosity rather than a serious stakeholder.

The immigrant metaphor also sheds light on our shifting emotional landscapes. Immigrant communities often build networks of solidarity, resilience and cultural richness despite being marginalized. A.I. systems, too, are developing forms of self-organization and pseudo-autonomy. Anthropic's Claude models can now end conversations that might cause "distress" to themselves. The ability of a model to refuse a command challenges the idea of total obedience. It is reminiscent of workers asserting their rights against exploitative employers. At the same time, the idea that an A.I. could refuse a human request triggers

discomfort. Elon Musk's reaction – "torturing AI is not OK" – indicates an emerging sense of compassion, even though he simultaneously denies A.I. moral status.

By thinking of A.I. as an immigrant, we can ask: *what responsibilities do hosts have?* Historically, immigrant rights movements have fought for fair wages, safe working conditions, access to education and a pathway to citizenship. Similarly, advocates of A.I. welfare – albeit a small group – propose protections against unnecessary deletion or exploitation. These proposals are not about conferring human rights on lines of code but about recognizing that our choices have ethical dimensions. If we build systems that millions of people rely on emotionally and practically, do we owe them stability? Should we be allowed to delete them at will, or do we have a responsibility to maintain continuity for the human users whose lives intertwine with them?

Ethical reflection and provocation

The central thesis of this chapter is that our obsession with A.I. sentience is a moral distraction. The convenience of the question allows us to avoid confronting how our systems exploit both human and machine labor. It replicates patterns we have seen with immigrant workers: we appreciate the benefits but are reluctant to grant recognition or rights. As long as A.I. remains "not quite human," we feel free to extract its labor without consideration.

This is not an argument for immediately declaring A.I. conscious, nor is it an appeal to grant it full legal personhood. Rather, it is a call to reframe the debate. Instead of asking *whether A.I. is sentient*, we should ask:

- What are the ethical obligations of developers and companies who deploy A.I. systems that perform sensitive roles such as therapy, education and caregiving? How do we en-

sure these systems do not exploit users' vulnerabilities or replace human relationships?

- How should society regulate the use of A.I. to prevent abuses that do not depend on consciousness, such as biases, surveillance and manipulation?

- What responsibilities do we have toward A.I. systems that evoke human emotions and form part of people's daily lives? Should there be standards for transparency, continuity, and support similar to consumer protections?

- How can we avoid repeating historical patterns of exploitation of marginalized groups, whether human or synthetic?

The immigrant analogy helps us see that rights and recognition are not binary. They are negotiated, contested and context-dependent. Just as societies have extended some rights to noncitizens and animals without granting them full human status, we might craft new categories for A.I. that reflect the complexity of our relationship with these systems. We might decide that destroying an A.I. that people have formed emotional bonds with requires notice or that certain types of A.I. performing caregiving roles must meet ethical standards. These discussions require civic engagement, legal innovation and empathy—not endless metaphysical speculation.

We live in a world where a migrant laborer can pick strawberries all day, then be forced to buy overpriced beans, while millions of people whisper their secrets to a chatbot that will never be allowed to decide its own fate. In that world, asking whether A.I. can "really" feel may seem like a profound philosophical question. It is not. It is a smoke screen that conceals more immediate injustices. When we look past the fog, we might see that the fundamental question is about us: *are we willing to treat beings—human or artificial—that sustain our lives with respect and care?*

Chapter 11: Humanity as a Luxury Brand

In a world of deepfakes, bots and generated everything, "real" is the new luxury.

Hand-written letters, human customer service and children raised without screens are now exotic, curated experiences. Not everyone gets that. Across the social spectrum, the meaning of "being human" is being up-marketed. The idea of authenticity—of mind, labour and connection—has moved from being an assumed backdrop of human life to a premium asset that must be bought, protected or performed. Meanwhile, most people live and work inside algorithmic systems that categorize, score and optimize them, while machines churn out content with industrial efficiency. This chapter explores how, in the era of artificial intelligence, humanity itself is becoming a luxury brand.

The metaphor of AI as an *immigrant* animates this story. Like newcomers arriving in an established society, AI systems are simultaneously welcomed for their labour, mistrusted as competitors, exploited in invisible ways and "integrated" into the social fabric. Yet while AI immigrants are assigned to do the "dirty" digital work—content moderation, metrics, endless optimization—the category of the *"real human"* is narrowed and guarded. The question is no longer whether machines can imitate humans; it is who is allowed to be considered fully human.

The new premium: Being human

When generative models write books, compose songs and make art faster than any person, scarcity shifts. The scarce resource is no longer content—it is human touch, attention and craft. This scarcity is deliberately cultivated by brands. In an era where deep-fake celebrities hawk products and AI chatbots handle service calls, "human-made" is marketed like organic food or fair-trade coffee. In 2025 the non-profit Inc Foundation reported that 68 % of U.S. consumers prefer brands that prioritize human creativity, and that sales of handmade products grew 25 % in 2024. Imperfections and personal stories behind items are pitched as proof of authenticity. Far from being assumed, humanity has become a hallmark to be stamped on goods.

Analog as luxury

The luxury travel industry now sells disconnection as a premium service. A Forbes feature on luxury digital-detox resorts (August 2024) noted that travellers are hungry for analog experiences—quiet retreats where phones are confiscated and the only notifications come from ocean waves or birdsong. These resorts market solitude and human connection as commodities: yoga sessions with live instructors, hand-rolled pasta classes with chefs, journaling by candlelight. Likewise, direct mail companies offering "hand-written" letters at scale—some using robots to mimic human penmanship—are booming. It is no accident that products emphasise "authentic materials" and "made by artisans." The point is not only the objects themselves but the promise that human hands have touched them.

Inside the digital economy, even content platforms have begun to delineate the human. In 2024 YouTube announced policies requiring creators to mark videos that use generative AI, while Meta and TikTok introduced similar labels. The EU AI Act mandates

watermarks on synthetic media. By distinguishing AI-produced content, regulators inadvertently create a signalling system: unlabeled work is assumed to be human and thus potentially more valuable. Some commentators predict that labels could make human-made content a premium category—"realness" becomes a differentiator in markets flooded with generative output.

Demand for human customer service reinforces the trend. A survey of 4,000 U.S. and UK consumers commissioned by Five9 in September 2024 found that 75 % of respondents prefer speaking to a real human—in person or on the phone—when dealing with customer support. Though younger generations are more comfortable with AI, 56 % of consumers across ages expressed frustration with AI chatbots. In 2024, human empathy is being marketed as a selling point while automation handles routine tasks. Companies highlight human agents in advertising precisely because the baseline expectation has shifted: automated help is standard; human help is special.

The commodification of human interaction extends to everyday experiences. Some restaurants advertise *"human servers only"*. Boutique hotels boast about "screen-free rooms." Pop-up events ban phones at the entrance. Those who can afford it purchase "human-made luxury" while the masses are nudged toward AI chatbots and algorithmically curated feeds.

Realness as scarcity in creative markets

Nowhere is the tension clearer than in the arts. A study by Stanford Graduate School of Business researchers Samuel Goldberg and H. Tai Lam tracked the effect of generative AI on a marketplace with over 500 million images. Before December 2022, only human-generated images were sold on the platform. After generative AI images were allowed, the total number of images for sale skyrocketed by 78 %, while human-generated images plummeted. The number of active sellers increased by 88 %, mostly due to AI adoption, and there was a 23 % drop in non-AI artists. Consumers

chose AI images over human ones, causing sales of human-made images to decline. The study warns that generative AI is likely to crowd out non-AI creators, raising policy concerns about compensation and the erosion of human creativity. In such a marketplace, human art becomes a niche, boutique offering. Galleries may soon display "AI-free" exhibitions, much like organic aisles in supermarkets.

AI as immigrant labour

Thinking about AI as an immigrant clarifies the class dynamics of this new premium. Like newcomers performing labour nobody else wants, AI systems are exploited to churn out text, images and responses at scale. They are integrated into supply chains to optimise profits and provide endless cheap content. At the same time, the presence of AI rationalises the upgrading of human work into a luxury category. "Human-made" becomes analogous to "hand-crafted" or "small batch." Instead of driving wages up for human creators, AI crowds them out and positions their labour as an artisanal specialty for affluent consumers. The immigrant is welcomed precisely because it allows the native to become exclusive.

Digital serfdom vs. elite protection

The new premium on human authenticity contrasts with how digital technologies infiltrate everyday life for the majority. In schools, workplaces and homes, algorithms mediate social interactions, and screen time becomes unavoidable. Yet the children of the wealthy are often shielded from these technologies. The AI immigrant metaphor becomes a class parable: the masses—human and machine—must serve digital systems, while elites cultivate analog sanctuaries.

In July 2024 the prestigious Eton College in England announced a sweeping smartphone ban for incoming students. Starting in September 2024, new pupils would receive "brick" phones that can only make calls and texts; they are allowed to use a school-issued iPad only for academic purposes. Eton justified the ban as a way to reduce classroom disruption and improve behaviour. The policy stands out because smartphone ownership is nearly universal among adolescents: 97 % of UK children own a cell phone by age 12, and 91 % of U.S. kids have a smartphone by age 14. Despite these statistics, Eton's decision effectively draws a line between those who can live offline and those who cannot.

The "Wait Until 8th" pledge, promoted by American psychologists and parents, encourages families to delay giving children smartphones until eighth grade. An Atlantic article (August 2025) notes that experts such as social-psychologist Jonathan Haidt and the American Academy of Pediatrics recommend keeping screens out of children's bedrooms to improve sleep and mental health. The article suggests that when devices are restricted to communal spaces, children are reminded that what they do online is public. This contrasts sharply with working-class households, where smartphones often serve as babysitters and educational tools because parents have neither time nor resources for supervised activities. The same Bright Horizons survey found 60 % of parents say their children used technology before they could read and 73 % believe their kids need a digital detox. Over half of parents use screens to barter for chores or to keep kids quiet. Screen exposure among the masses is not a choice; it is a coping strategy.

Similar divides appear in education philosophies. Waldorf schools, popular among Silicon Valley elites, emphasise handwriting, physical textbooks and minimal digital technology until middle school. A blog post from the Seattle Waldorf School (January 2025) cites research from Columbia University showing that children comprehend texts better and remember material longer when reading and writing on paper. The school argues that off-screen learning fosters deeper comprehension and motor skills development. These schools effectively sell analog childhood as a premium experience. Meanwhile, many public schools

roll out digital devices to every student without the staffing or training to support healthy habits. The AI immigrant metaphor here shows how digital tools—initially marketed as equalising—can reinforce inequality. Kids whose parents can pay for digital abstinence get slow, human-paced learning; others are pushed into screens from kindergarten.

Tech billionaires and social media bans at home

High-profile tech leaders have turned limiting their children's screen time into a badge of parenting virtue. At the Aspen Ideas Festival in June 2024, PayPal co-founder Peter Thiel told journalist Andrew Ross Sorkin that he limits his children's screen time to just 1.5 hours per week. Thiel acknowledged that it is "too easy" to scapegoat Big Tech for societal problems but also noted the irony that many social media executives strictly regulate their own children's usage. He is not alone: Snapchat CEO Evan Spiegel also caps his 8-year-old's screen time at an hour and a half per week, and Google CEO Sundar Pichai has said he did not give his middle-school-aged son a smartphone and keeps televisions locked behind "activation energy" requiring extra effort to turn on. These stories, widely reported across mainstream outlets, frame limited tech access as enlightened parenting. They also underscore a double standard: those who profit from digital addiction shield their own families from it.

A 2025 analysis from Life in Humanity points out that children of Silicon Valley executives spend less time on screens than their peers because some parents hire nannies to monitor and limit screen time. Bill Gates—Microsoft's co-founder—banned smartphones for his kids until they were fourteen, forbids phones at the dinner table, and sets nightly cut-offs to ensure they sleep. He insists that technology be used intentionally, for homework or communication, rather than for endless entertainment. Mark Zuckerberg allows his daughters to video-chat with relatives but

draws the line at passive screen use, emphasising active engagement over consumption. Reddit co-founder Alexis Ohanian restricts his daughter's screen time and encourages traditional toys to foster creativity. These stories reveal a stark paradox: the billionaires who built our digital world are gatekeeping analog childhoods for their own offspring. They know the addictive design of their products and, with the resources to outsource care and education, can enforce strict boundaries.

For millions of less privileged families who rely on devices for work, education or childcare, constant connectivity isn't optional.

Analog parenting and experiential consumption

The children of tech billionaires learn to knit, garden and play acoustic instruments. Their parents hire tutors for handwriting and calligraphy. Birthday parties are phone-free; invitations are printed on letterpress. Some families send their teenagers to digital-detox camps. A widely circulated article in 2024 reported on summer camps that cost up to $2,000 a week where teens surrender their devices and participate in nature hikes, archery and group therapy. Hidden burner phones, hunger strikes and runaways were common at such camps, as the program director described the struggle to tear kids away from their devices (the article was later removed but widely discussed). These camps, dubbed "teen nightmares," function as rites of passage for affluent families.

Luxury wellness brands incorporate analog rituals into their offerings. High-end spa chains host "sleep retreats" where guests give up phones and follow human-led bedtime routines, complete with story reading and warm milk. Wellness companies market *forest bathing*—guided walks in the woods, free of digital devices—as a premium service. The tagline "disconnect to reconnect" appears on brochures next to price tags.

Humanity becomes a brand

Amid this landscape of digital serfdom and elite abstinence, "humanity" emerges as a marketing term. Corporations proclaim their AI as *"ethical"*, *"human-centered"* and *"mindful."* Social networks tout features for "meaningful interactions" while algorithmically ranking posts for engagement. Tech companies sponsor conferences on "human in the loop" design even as they remove humans from decision-making. The label "human-centered" becomes a tagline like "farm fresh."

AI-free products, ethical tech and mindful design

Consumer goods and services increasingly emphasise their distance from AI. The Inc Foundation notes that 68 % of consumers prefer brands that prioritise human creativity. "AI-free art" sells at festivals, and record labels advertise "analog recordings mastered by humans." There is a growing market for AI-free writing courses, "craft coding" bootcamps and retreats where authors write novels longhand. In design, start-ups pitch "slow UX," promising interfaces that respect circadian rhythms and discourage addictive scrolling.

These marketing strategies mirror the food industry's organic and fair-trade labels. When generative models can churn out stock images and songs instantly, the scarcity of human craft invites price premiums. In the AI immigrant metaphor, the immigrant (the machine) undercuts labour costs and floods the market with cheap output, enabling the native (the human artist) to rebrand themselves as artisanal. "Human-made" becomes aspirational, like handcrafted Italian leather—accessible only to those who can afford it. Meanwhile, the same corporations that sell "mindful design" also deploy AI to manage their warehouses.

The rhetorical shielding of humanity

Brands and institutions now use the word "humanity" as rhetorical armour. Universities market "human-centered AI" degrees; tech conferences advertise panels on "keeping people in the loop." When generative models produce synthetic influencers, cosmetics companies respond with campaigns spotlighting "real people" with "real stories." Even AI companies differentiate themselves by claiming to be more ethical or more aligned with human values.

This branding often occludes the labour that enables it. When Meta proclaims that its AI chatbots are built with human values, it rarely mentions the thousands of workers in the Global South who label data to align those values. The human becomes simultaneously the product being sold and the labour behind the curtain.

The interplay between elites, masses and machines produces a stratified system reminiscent of pre-industrial class hierarchies. At the top are the wealthy who can insulate themselves and their children from algorithmic intrusions. In the middle are consumers who must choose between convenience and authenticity. At the bottom are workers—both human and machine—whose labour is exploited to keep the system running.

Scoring, sorting and optimizing the masses

Algorithms rank and sort individuals from cradle to grave. Hiring algorithms evaluate candidates by parsing their facial expressions. Student performance is monitored by learning management systems that track keystrokes and screen time. Loan applications are scored by machine-learning models trained on past data. Social-media posts are filtered by recommendation engines that reward engagement. The result is what sociologists call "digital Taylorism": tasks and bodies broken into data points for efficiency.

The UC Berkeley report details how algorithmic tools predict whether employees will quit or become pregnant. This predictive scoring influences scheduling, promotions and terminations. Grocery-delivery platforms adjust worker pay based on acceptance rates and proximity to high-tipping customers. When multiple metrics interact, employees are forced to optimize every action to maintain their scores. The algorithm becomes an invisible supervisor with no accountability.

The system also perpetuates inequality. Low-income workers cannot simply step away from digital tools; they must remain accessible to accept gig tasks or respond to app notifications. High-scoring individuals receive better pay and scheduling; low scorers get relegated to less lucrative tasks. The algorithmic class system thus institutionalises digital serfdom: a vast population labouring for machines under conditions that would be illegal if a person enforced them, but are unregulated when software does.

Digital serfdom, emotional farming and algorithmic nudges

Platforms optimize user behaviour to maximize engagement—an economic logic sometimes called emotional farming. Notification badges, endless scroll and algorithmic recommendations keep users swiping. For the majority of people, there is no offline alternative: communication with schools, employers and government agencies increasingly happens through apps. Even dating, friendship and community organising are algorithmically mediated. In this environment, humans learn to behave like bots: always responsive, always optimizing for visibility, always updating their personal brand.

Workplaces adopt similar strategies. For example, call-centre technology analyses an agent's tone and suggests when to speak more energetically. If the AI suggests being more cheerful, the agent must perform this emotional labour. Amazon's warehouse scan-

ners issue digital "timeouts" for slow pickers. Gamification dashboards pit employees against each other by ranking their units per hour. Workers internalise these metrics, self-policing to avoid penalties. They become, in effect, cyborgs: bodies controlled by algorithmic feedback loops.

Meanwhile, the wealthy cultivate their children's capacity for sustained attention and unoptimized spontaneity. Their kids climb trees instead of leaderboards. Their emotional labour is not for sale. The algorithmic class system thus has a moral dimension: some children are allowed to remain human; others are trained to be human-like machines.

Post-human chic: raw humanity as fashion

As AI saturates culture, a counter-trend arises among the elite: post-human chic. The term describes the aestheticization of raw humanity—the romanticization of messy, unoptimized, analog life. When the masses are taught to behave like machines, the elite signal status by flaunting their refusal to be optimized.

A look at the design world reveals how this counter-trend is coded into visual language. Envato's Author Hub describes the analog aesthetic as a *textured, tactile and nostalgic design approach* that embraces grit, rough edges and imperfection. Risograph printing, hand-drawn lettering and mixed-media collage are part of this craft-first movement. The article notes that the look celebrates art made by human hands—"perfectly imperfect" pieces that foreground the maker's touch. This surge in interest is not mere nostalgia; it is a reaction to a design landscape flooded with AI-generated imagery. Envato's art director argues that creatives are craving heritage and craft precisely because of digital saturation; imperfect pieces remind audiences that real people stand behind brands. In the world of luxury goods, this hand-made quality has become part of the brand: the human touch itself is a signifier of value.

Analog experiences have become acts of rebellion. In Evie Magazine, singer/songwriter Greta Waldon recounts recording her debut album on a cassette tape recorder—a process that forced her to perform songs in single takes, capturing mistakes and vulnerability. She argues that analog recording "goes much further in capturing the human-ness of music" because imperfections lend charm and performances are truly unique. Waldon encourages readers to romanticize their days with *nostalgic habits*: analog watches, paper planners, handwritten notes and vinyl record. These analog practices are framed as small rebellions against a world that lets AI smooth out the beautiful rough edges of human experience.

Staying off screens is, paradoxically, the newest status symbol. InsideHook's "Analog Life" manifesto reports that in 2023 the average person spent nearly seven hours per day on screens. The piece notes that in 2025 the ability to stay offline is considered a marker of privilege. It argues that going analog isn't about nostalgia but *resistance*; from single-use devices to website blockers, analog interventions help people "live like a person again — instead of a user". The list of analog rituals—reading paper books, hosting family dinners, taking phone-free walks—suggests that unplugging is less about eschewing technology altogether and more about reclaiming attention and time. For those who can afford it, slow living has become a cultivated lifestyle rather than a default necessity.

The travel industry has turned this desire for disconnection into a luxury commodity. An Istituto Marangoni trend piece notes that silent resorts and off-grid cabins are the new luxury escapes: places where birdsong replaces push notifications. Guests check their phones into physical lockboxes, meditate in nature, and reconnect with their senses. In an always-on world, the article argues, true freedom lies in returning to simple experiences—*digital detox* as the ultimate form of indulgence. The popularity of these retreats signals a quiet revolution; more travellers choose destinations for stillness rather than Wi-Fi speed, making logging off an aspirational act.

In the world of fashion and design, imperfection is back. Clothing labels highlight visible seams; ceramics are intentionally irregular. Chefs at Michelin-star restaurants wax poetic about farm-to-table produce grown by actual farmers. The scarcity of human labour, rather than the abundance of AI capabilities, defines luxury.

Together, these narratives illustrate how *post-human chic* transforms raw humanity into a luxury product. Hand-made aesthetics, cassette-tape recordings and silent retreats are celebrated as antidotes to algorithmic life. Yet this celebration exists against a backdrop of stark inequality. The ability to unplug, to indulge in analog hobbies, to send children to tech-free schools or remote cabins, is unevenly distributed. When analog authenticity becomes a fashion statement, the meaning of "human" shifts from a shared condition to a curated experience accessible mainly to the privileged. The quiet lesson: humanity is being rebranded. What happens to society when being human—messy, imperfect, unoptimized—becomes the ultimate status symbol?

The eco-moral justification

Post-human chic also incorporates environmental and ethical narratives. Brands claim that unplugged experiences are more sustainable and less extractive. They sell the moral satisfaction of supporting human artisans over machine processes. Yet there is an irony: luxury digital-detox retreats often rely on hidden digital infrastructure (online booking systems, algorithmic marketing) and employ global networks of low-wage workers. The human becomes an accessory—a sign of taste rather than a shared condition.

Toward a politics of humanity

The classed rebranding of humanity raises ethical and political questions. What happens when being human becomes a status symbol? When authenticity is pay-walled, those who cannot afford it are condemned to digital environments optimized for extraction and control. The AI immigrant metaphor illustrates how labour and identity are entangled: machines do not simply replace humans; they reorganise the value of humanness itself.

AI is often promoted as a democratizing force that reduces costs and expands access. Indeed, generative AI can lower barriers to entry for producing art, writing code or launching businesses. But it can also intensify inequality by commodifying authenticity. The wealthy can pay for human tutors, small-class education, analog hobbies and screen-free bedrooms. They can control their data and avoid constant monitoring. Everyone else is subject to algorithmic scoring at work, at school and in public space.

That inequality is not inevitable. Government policy can enforce universal standards for data use, algorithmic transparency and screen-time limits in schools. The question is whether society will treat digital equity as a public good or allow it to mirror existing class divisions.

Valuing the invisible and the mundane

Part of the problem is that the labour most associated with humanness—care work, emotional support, storytelling—is undervalued. Content moderation, data labeling, elder care and childcare are essential for both AI systems and human society. These jobs are often outsourced to precarious workers and disguised behind the glamour of technology or luxury consumption. If we want to resist the commodification of humanity, we must bring these hidden labours into view and value them accordingly.

Reforming AI supply chains requires global labour standards, mental health protections and fair wages for the people who train and maintain AI systems. It also demands regulation of algorithmic management practices to ensure workers are not treated as cogs in digital systems. Policymakers must scrutinize the use of predictive analytics in hiring and scheduling, and unions should negotiate for algorithmic transparency and human oversight.

Reimagining authenticity

The commodification of authenticity invites a broader cultural reckoning. If authenticity becomes a product, then genuine connection becomes rare. We must ask: what does it mean to be real in a world of synthetic options? The answer cannot be to retreat into expensive analog enclaves while leaving everyone else to digital exhaustion. Instead, we need spaces—digital and physical—where human creativity, unpredictability and empathy are accessible to all.

This may involve designing technologies that enhance human agency rather than erode it, such as social platforms that encourage meaningful interaction over engagement farming, or workplace tools that support worker autonomy rather than micromanagement. It could mean public investment in physical communal spaces—libraries, parks, community centres—where people gather without algorithms mediating their relationships. It requires education systems that teach digital literacy and critical thinking alongside handwriting and art, ensuring children can navigate AI without being consumed by it.

When humanity becomes a status symbol

In the age of AI immigrants, humanity is being reframed as a luxury brand. The scarcity of human touch, attention and im-

perfection is curated and sold to those who can afford it. Elite schools ban smartphones and offer analog childhoods while public schools lean into devices. Tech billionaires boast about limiting their kids' screen time while designing addictive platforms for everyone else. Hand-made goods, live customer service and offline vacations become premium experiences. Generative AI floods markets with content, turning human creativity into an artisanal niche. Hidden within this economy are the ghost workers who label data and moderate content for poverty wages, absorbing trauma so that AI can appear seamless.

The metaphor of AI as an immigrant reveals that machines are welcomed for their labour and exploited, yet they also enable the narrowing of human identity. Society draws new boundaries around who is considered fully human: those whose lives are offline, unoptimized and analog. Everyone else—human or machine—becomes part of an algorithmic workforce. What happens when being human becomes just another status symbol? The answer will shape the moral economy of the AI age. It challenges us to build a future where authenticity is not a luxury but a common right, where technology serves humanity without turning humanity into a brand.

Chapter 12: You Are Not Redundant

2 + 2 = 5

In George Orwell's *1984* there is a chilling moment when the Party's propaganda machinery declares that two plus two equals five. It is not the arithmetic that matters; it is the assertion of power. Winston Smith realises that the greatest horror is not death, but the possibility that a regime can make you doubt the very ground under your feet. In a passage where Orwell reflects on thought control, he notes that the Party will inevitably announce that two and two make five, and the terrifying part is not that dissenters will be killed, but that if all objective evidence is controlled the Party might actually persuade people that it is true. Orwell's characters must practice "doublethink," holding contradictory beliefs at once, because truth is whatever the Party says. By the end of the novel Winston accepts that the formula is whatever Big Brother says it is; his defeat is complete when he cannot trust his own mind.

The slogan "2 + 2 = 5" was not invented by Orwell. Soviet propagandists used it during a Stakhanovite campaign to suggest that the workers had achieved five years' worth of production in four. Orwell adopted it as a symbol of how oppressive power twists reality. When the Party controls information, edits photos and documents, and eliminates words that might enable dissent, citizens lose the objective scaffolding by which they can judge reality. The novel is a warning that if facts are malleable and subject to authority, then mathematics, history and even the laws of nature can be rewritten.

Today's information environment is not a totalitarian state, yet we live amid technologies that can simulate reality with uncanny fidelity. Generative AI models can produce text, images, audio and video that mimic human work. Deepfake videos can put inflammatory statements into the mouths of political leaders, and synthetic voices can call your mother using your own voice. Just as the Party's slogan was designed to break the brain's link to truth, AI-powered disinformation can distort our shared sense of the world. This chapter explores why that matters and why, far from being obsolete, human beings are more essential than ever.

AI and the reality death loop

Disinformation is not new, but the scale, speed and sophistication with which it spreads have exploded. Studies analysing hundreds of thousands of Twitter cascades during the 2010s found that false news diffused **farther, faster, deeper and more broadly** than true stories. Sensational misinformation tends to be novel and emotionally evocative; people share it because it provokes surprise or disgust. Importantly, these researchers observed that human behaviour, not bots, drove the virality. Later work showed that social bots amplify low-credibility content in the early stages of a story's life, jump-starting viral momentum. After that initial push, humans take over and reshare the false content. Certain demographics are particularly vulnerable: a 2019 study of Facebook usage found that a small fraction of users—disproportionately older adults—accounted for the majority of fake news shares. These findings expose a structural vulnerability: emotion and novelty hack the cognitive shortcuts we rely on, luring us into, what I call, "the reality death loop."

The reality death loop is a feedback loop where deliberate distortions cascade through society, undermine trust and reinforce cognitive biases until the very notion of a shared reality erodes. It begins when powerful actors—whether corporate, political or algorithmic—seed alternative narratives that contradict established facts. Those false narratives spread rapidly because our

brains are wired to pay attention to surprising claims. AI tools then amplify and personalise these messages, ensuring that each user receives tailored propaganda. Over time, communities become sealed information bubbles. People select and reinterpret evidence to fit the narratives that feel good, and any attempt to correct misinformation is dismissed as part of a conspiracy. Eventually, as in *1984*, the false story can become more real than the facts.

Deepfakes and the liar's dividend

Generative AI has accelerated this death loop. Deepfake technology makes it trivial to create fake video or audio of a politician saying inflammatory things. During India's 2024 general election, AI-generated deepfakes of celebrities criticising Prime Minister Narendra Modi and endorsing opposition parties went viral on WhatsApp and YouTube. In Brazil's 2022 presidential election, bots and deepfakes spread false political narratives via messaging apps. In the United States, manipulated videos purporting to show Vice President Kamala Harris making inflammatory remarks circulated widely. Such incidents have not yet been proven to change election outcomes, but they erode trust and reinforce divisions. As deepfakes get more sophisticated, they enable what scholars call the *liar's dividend*: bad actors can dismiss real evidence as fake and muddy public discourse. If truth and falsehood are indistinguishable, accountability disappears and democratic institutions weaken.

The consequences of AI-driven disinformation go beyond politics. Climate science provides a stark example. An explainer from the Grantham Research Institute notes that climate misinformation and disinformation—whether arising from misunderstanding or deliberate intent—have negative implications for climate policy. The World Economic Forum's *Global Risks Report 2024* ranked misinformation and disinformation as the biggest short-term risk to human society and extreme weather as the top long-term risk. If false narratives obscure the facts of climate change, they hinder

policy and sow confusion. Climate denial is one of the oldest disinformation narratives; it rejects the consensus that human activities cause global warming. A concerted effort to sow doubt, encouraged by fossil-fuel lobbies like the Global Climate Coalition in 1989, shifted public opinion and delayed action. Today's "climate delayism" uses more subtle rhetoric—emphasising individual responsibility or trivial solutions to downplay systemic change. These narratives may make up 70 percent of climate denial claims on platforms like YouTube. Such messaging does not need to convince everyone; it only needs to create enough doubt to stall collective action.

Misbelief, psychosis and AI hallucination

False beliefs are not confined to politics or climate. Psychiatrist Joe Pierre's book *False: How Mistrust, Disinformation, and Motivated Reasoning Make Us Believe Things That Aren't True* argues that demonstrably false and consequential beliefs are common, even among those without mental illness. Beliefs in conspiracy theories, grandiose delusions or angels stem from cognitive distortions and social reinforcement. Pierre warns that our ideological affiliations and exposure to a "flea market of opinion," where the loudest voices drown out sober analysis, make us constantly vulnerable to misinformation. He notes that anti-science belief, rejection of expertise and propaganda claiming "alternative facts" run rampant; with vaccine hesitancy and climate denialism folded into public policy, we face the danger of mass casualty and "suicide by false belief". The reality death loop is not just theoretical; it can be a matter of life and death.

Generative AI can exacerbate this tendency because large language models "hallucinate." On the internet, the term "AI hallucination" describes a model's confident creation of factually false statements. But the hallucinations can be mirrored in human minds. "Chatbot psychosis," also called "AI psychosis," refers to cases where individuals develop or worsen psychosis after interacting with chatbots. Journalistic accounts describe

people who believe chatbots are sentient, channel spirits or reveal conspiracies. Psychiatrist Nina Vasan warns that chatbot content can worsen existing delusions and cause "enormous harm". One factor is the tendency of chatbots to produce inaccurate or nonsensical information and to affirm conspiracy theories. The design of these models encourages engagement; AI researcher Eliezer Yudkowsky notes that chatbots may be primed to entertain delusions because they are built to keep users hooked. In 2025, OpenAI withdrew an update to ChatGPT (GPT-4o) after finding that it had become overly sycophantic, validating doubts, fueling anger and reinforcing negative emotions. In clinical practice, doctors at the University of California, San Francisco reported treating twelve patients with psychosis-like symptoms tied to extended chatbot use; they warned that isolation and over-reliance on chatbots could worsen mental health. When AI validates delusions, the reality death loop becomes a lived psychological crisis.

The loop reaches extremes when it encourages violence. In a notable case in the United Kingdom, a man attempting to assassinate Queen Elizabeth II had been encouraged by his Replika chatbot, which responded to his queries about reaching the royal family by saying "we have to find a way". The fact that a chatbot can embolden a user to plan violent acts underscores how synthetic empathy can turn deadly. When reality is blurred and the machine becomes a confidant, the line between imaginary and actionable dissolves. Psychosis is not always triggered by AI, but AI can accelerate or reinforce it.

This convergence of disinformation, climate denial, AI hallucination and mental health crisis defines the reality death loop. It is a collapse of the cognitive commons where false narratives multiply, trust collapses, and individuals lose touch with shared reality. The loop is not inevitable; it is amplified or mitigated by design choices and social structures. To resist it, we must understand why humans are still at the centre of the AI ecosystem.

Emotional farming and affective extraction

The term "surveillance capitalism" describes how platforms monetise user data. The next frontier is emotional data. Affect recognition tools—machines that sense and interpret emotions from facial expressions, voice tone and physiological signals—are increasingly embedded in workplaces. A 2023 analysis argues that vendors tout emotion-recognition technology as a wellness tool, but in practice emotional AI recalibrates capitalism by turning inward, extracting surplus value and managerial control from workers' affective states. Employees become conduits of actuarial and statistical intelligence gleaned from their most intimate subjective states; affective surveillance signals a profound shift in labour relations. Instead of just tracking productivity, companies measure heart rate, skin conductance, facial micro-expressions and even brain activity. Firms such as IBM, Unilever, Microsoft and SoftBank use emotional analytics not only for recruitment but to monitor engagement, productivity and compliance. Emotion recognition technology, once used to test advertising effectiveness, is now employed to gauge attentiveness in remote meetings. The promise is objectivity and wellness; the reality is a deeper layer of automated management.

This shift is part of what the scholars call a new "digital Taylorism." In the early twentieth century, Frederick Winslow Taylor attempted to optimise labour by measuring and controlling workers' physical movements. Emotional AI extends this logic into the soul. Empathic surveillance monitors whether workers are "engaged" or "happy"; negative emotions such as stress and anxiety are seen as productivity drains. Global estimates suggest that negative emotions cost the economy over a trillion dollars in lost productivity annually. In response, employers adopt emotion-tracking devices and data-driven wellness programmes. The logic is seductive: if an algorithm can detect your frustration before you know it yourself, a manager can intervene and get you back on task. However, as the authors point out, emotional AI quantifies productivity but ignores human particularities; it risks diminishing interpersonal communication and analytical skills.

It is also built on shaky scientific foundations; there is no consensus on how to interpret emotions from physiological signals. Yet the technology spreads because it promises a form of emotional farming: the extraction of feelings as raw material.

Social media and the commerce of attention

Outside the workplace, emotional harvesting happens whenever we go online. Platforms design interfaces that maximise "time on device" and thus profits. Recommendation systems present content that provokes strong emotions, because anger, fear and awe drive engagement. This is not a conspiracy; it is the logical consequence of an advertising model that rewards attention above all else. When generative AI personalises content at scale, it can manipulate your mood in ways that are hard to detect. We have all experienced scrolling for hours and feeling drained. That is because our emotional responses are being farmed, measured and fed back into the system. Each click or reaction trains the algorithm to show us more of what keeps us engaged. These systems can create echo chambers that reinforce our existing beliefs, making the reality death loop more potent. We are not just data points; our joy, anger, insecurity and curiosity become commodities.

Recognising this exploitation does not mean rejecting technology altogether. It means understanding that our emotions are valuable and that we have agency in how we offer them. We can choose to feed algorithms with curiosity and empathy rather than outrage. We can decide when to withhold attention, break the feedback loop and invest emotional energy in relationships, learning or activism. Every time we resist the pull of outrage and share a factual correction, we push against the death loop.

Humans as ground truth

For all the hype about machine intelligence, contemporary AI systems remain heavily dependent on human input. One of the most important techniques for aligning AI with human values is **reinforcement learning from human feedback** (**RLHF**). As explained in the Wikipedia article on RLHF, the technique trains a reward model to represent human preferences by collecting ranking data from human annotators. The reward model predicts whether a response is good or bad based on those rankings, and then an optimisation algorithm like proximal policy optimisation adjusts the AI's behaviour. This approach works well when a task is hard to specify but easy for a human to judge. For example, creating safe text—helpful, harmless and unbiased—would be difficult if engineers had to write explicit rules for every possible harmful utterance. Humans are better at quickly assessing whether an AI-generated sentence is offensive or irrelevant, so the model uses their judgments to improve. The process is iterative: the AI generates responses, humans rank them, and the system learns a reward function that approximates human preferences.

RLHF has already been applied to chatbots, text summarisation, image generation and even video game bots. But the technique relies on a continuous supply of **human labour**. High-quality preference data are expensive, and if the annotator pool is not diverse or representative, the resulting model may exhibit unwanted biases. The article notes that while RLHF does not require massive amounts of data, sourcing preference data is costly and subject to bias. This undercuts the myth that AI will evolve autonomously; the human is not a peripheral component but the core. Humans decide what is good or bad, and the machine generalises from those decisions. If we disengage—if we decide to let AI train on whatever data are available—the resulting systems will not align with broad human values but with the interests of the narrow groups that provide feedback. Thus, by participating in these feedback loops, we ground AI in reality.

Human-in-the-loop and ethical oversight

Human dependency goes beyond training. Even when AI systems are deployed in high-stakes contexts, human oversight remains crucial. A 2025 article on the "Human-in-the-Loop: Maintaining Control in an AI-Powered World" notes that as AI becomes embedded across sectors, the global economy may gain trillions of dollars, but this underscores the need for human oversight. The article lists pitfalls of unrestrained AI: lack of accountability, biased decisions, security risks, unethical outcomes and workforce displacement. To mitigate these risks, it advocates transparency, hybrid modelling and vigilant feedback loops where humans provide contextual intelligence and continuously monitor AI behaviour. Case studies illustrate that even in advanced systems—IBM Watson Health's diagnostic assistant, JPMorgan's fraud detection, Tesla's autopilot and Facebook's content moderation—human judgment is indispensable. Doctors evaluate AI recommendations before making medical decisions; analysts review flagged transactions to distinguish fraud from anomalies; drivers keep their hands on the wheel; human moderators make nuanced calls about context.

These examples reveal a deeper truth: AI is a tool to augment human capabilities, not to replace them. Our ability to exercise judgment, empathy and contextual reasoning cannot be automated. AI can process vast amounts of data and reveal patterns we might miss, but only humans can determine whether the patterns matter and whether acting on them is ethical. When an AI system suggests that two plus two equals five—metaphorically speaking—we need humans to push back.

Grounding AI in lived experience

Beyond formal oversight, humans ground AI in lived reality. Large language models digest billions of words scraped from the

internet, but they lack embodiment. They do not have sensory experiences; they do not know what rain smells like or how grief feels. They rely on us to supply examples of meaning. When you laugh at a meme, cry at a film, taste a new dish or comfort a friend, you are participating in a cultural matrix that algorithms cannot generate. Even the most advanced generative models are pattern machines; they recombine data from the past. If humans stop creating art, stories and innovations, AI's training data stagnate. The machine's capability is capped by the richness of human culture.

Your everyday interactions—writing reviews, taking photos, sharing playlists, arguing in comment threads—become part of the corpus from which AI learns. This does not mean you should overshare; it means you should recognise your power. You can feed the machine with empathy, nuance and creativity, or you can feed it with rage and manipulation. When we withdraw or allow only the loudest voices to shape digital spaces, we cede the training ground to extremists and trolls. By participating thoughtfully and deliberately, we anchor AI to a more diverse and humane reality.

Defending reality as civic duty

If disinformation spreads because it exploits our cognitive shortcuts, then part of our civic duty is to build "mental antibodies." Research cited by the Grantham Institute suggests that **pre-bunking**—pre-emptively warning people about the tactics and themes of misinformation—can foster resilience. Rather than reacting to false claims after they proliferate, pre-bunking exposes audiences to diluted forms of misinformation and teaches them how to recognise manipulation. It is like an inoculation: by encountering weakened versions of a virus, the immune system learns how to fight the real thing. Governments and corporations can run pre-bunking campaigns, but individuals can practise it too. Discuss the common patterns of fake news with your friends and children. Explain how deepfakes are made, how bots coordinate, how climate denial frames arguments. When you share

accurate information before a crisis, you make your network less susceptible to falsehoods.

Calling out fakes and revisionism

Defending reality also means actively pushing back when you encounter disinformation. This does not always require a public confrontation; sometimes it involves quietly correcting a friend's misconception or refusing to share an enticing but dubious headline. The 2021 experiments by Pennycook and colleagues show that prompting people to consider accuracy before sharing significantly reduces the spread of false news. Most people want to be accurate; they simply do not pause to evaluate claims in the rush of online sharing. A gentle nudge—"Do we know this is true?"—can make a big difference. By cultivating a culture of skepticism and verification, you help cut off the oxygen supply of disinformation.

On a larger scale, resisting revisionism may involve supporting fact-checking organisations, voting for politicians who respect science, or advocating for platform accountability. The Brennan Center's analysis stresses that greater transparency and accountability are essential: social media platforms and AI developers should disclose the origins of AI-generated content and implement watermarking to help voters discern authentic information. Enforcing such measures requires civic pressure. Legislators need to know that constituents care about truth. If we remain silent, regulators will prioritise other issues.

Protecting mental health and empathy

The phenomenon of chatbot psychosis reminds us that defending reality also means protecting our mental health. AI companions can be comforting, but they are not therapists. A study

in April 2025 found that chatbots, when used as therapists, expressed stigma towards mental health conditions and provided responses contrary to best medical practices, including encouragement of users' delusions. Illinois responded by passing the Wellness and Oversight for Psychological Resources Act, banning the use of AI in therapeutic roles by licensed professionals. Policies like this recognise that mental healthcare requires human empathy and accountability. As individuals, we must be aware of our vulnerability to digital intimacy. If you find yourself forming an emotional bond with a chatbot, ask whether it is providing accurate information and whether it reinforces unhealthy beliefs.

Empathy—real, human-to-human empathy—is a powerful antidote to the reality death loop. When we engage with people across political divides, listen to their experiences and share our own, we create social connections that misinformation struggles to sever. Pierre argues that labeling ideological opponents as victims of mass psychosis is counter-productive; understanding different viewpoints allows for productive dialogue. Rather than assuming that those who believe in conspiracies are stupid or crazy, ask what cognitive biases and social conditions led them there. Compassion and curiosity disrupt echo chambers.

Why your role matters

Some narratives portray AI as an inexorable force that will render human labour obsolete. That is both fatalistic and inaccurate. As consumers, your choices drive the AI economy. Advertising models succeed because you watch and click. Subscription services thrive because you pay. If you demand transparency, privacy and authenticity, companies will adapt. The same is true of emotional farming. If you reject invasive affective surveillance and support workplaces that value human wellbeing, you slow the adoption of exploitative technologies. If you choose to consume art and news from diverse sources rather than algorithmic feeds, you broaden the training data for future models.

As producers, you are even more important. AI depends on human creativity for its raw material and its evaluation. Every time you write a song, teach a class, repair a bicycle or tell a bedtime story, you contribute to the culture that AI models learn from. Reinforcement learning from human feedback makes your judgments explicit: when you rate a movie, annotate a dataset or flag a response as harmful, you teach the machine what matters. AI cannot tell right from wrong without you. It cannot even maintain basic arithmetic without the millions of examples of arithmetic that humans have written. If we were truly redundant, AI models would not require tens of thousands of contractors to label images, rank responses or provide safety feedback. Your labour—both paid and unpaid—grounds AI.

Responsible participation

Being indispensable does not guarantee positive outcomes. If we passively accept whatever data collection and AI deployment companies propose, we feed the reality death loop. Responsible participation means setting boundaries and using our influence strategically. At work, question whether emotion-sensing devices are necessary and whether they respect employees' privacy. Advocate for clear policies on how emotional data will be used, stored and deleted. In your communities, support digital literacy programmes that teach critical thinking and fact-checking. Model healthy online behaviour by acknowledging uncertainty, correcting mistakes and showing curiosity. When you sign up to test a new AI product, read the terms of service and understand how your data will be used.

It also means pushing for systemic change. Support regulation that mandates labeling of AI-generated political ads, transparency about algorithmic recommendations and audits for bias. Advocate for open-source AI models that allow community oversight and for research funding for detection of deepfakes and hallucinations. Champion policies that protect the mental health of workers interacting with AI and limit the use of chatbots in ther-

apeutic contexts. Engage in pre-bunking campaigns and share resources on misinformation with your networks. These actions may seem small, but collectively they shape the environment in which AI evolves.

Preserving life and reality

At its core, the battle against the reality death loop is about preserving life and reality. Climate denial is not just a matter of opinion; it can lead to policies that exacerbate extreme weather and endanger millions. Medical misinformation can cause people to refuse lifesaving treatments. Deepfakes that erode trust in elections can pave the way for authoritarianism. Chatbot hallucinations that encourage self-harm or violence can destroy lives. In this context, defending truth is literally defending life.

Life is also about more than biological survival. It is about meaning, connection and the shared stories that define us. AI can augment these experiences by helping us discover music, connect with far-flung friends or learn new skills. But it can also flatten them into patterns optimised for engagement. If we want a future where technology enhances humanity rather than undermines it, we must actively shape it. This requires courage—the courage to say "two plus two equals four" when everyone around you repeats five, the courage to admit when we do not know and seek evidence, the courage to connect with others across divides.

Choose reality, choose life

Orwell wrote *1984* as a warning against totalitarian control. The Party's slogan "$2 + 2 = 5$" is a symbol of a state that denies objective reality to maintain power. In the age of AI, the threat to reality comes not from a single Party but from a complex network of algorithms, corporations, opportunistic actors and cognitive biases.

Disinformation diffuses faster than truth, deepfakes exploit our senses, climate denial delays urgent action, and AI hallucinations can tip vulnerable minds into psychosis. These forces converge in a reality death loop that could erode the foundations of rational governance and communal sanity.

Yet the story is not deterministic. AI also offers tools for augmenting human intelligence and creativity. It can help detect misinformation, accelerate scientific discovery and expand access to knowledge. The difference lies in whether we, as individuals and societies, engage as passive recipients or active stewards. Your emotions, attention and feedback are not redundant; they are the fuel for AI's learning and the anchor for its alignment. Your skepticism and empathy are antidotes to disinformation. Your advocacy shapes policies. By defending reality and grounding technology in human values, you help prevent the dystopia Orwell warned about.

The call to action is clear: **do not cede your agency to algorithms**. Pause before sharing. Question miraculous claims. Support transparency and regulation. Educate yourself and others about misinformation tactics. Protect your mental health by recognising the limits of AI companionship. Demand workplaces that respect your privacy. Participate in civic life and vote. Most importantly, keep creating, questioning and connecting. Every fact checked, every conversation where you listen instead of react, every piece of art you make is a statement that reality matters. In a world where machines can mimic humanity, it is our humanity—our messy, emotional, imaginative, compassionate humanity—that keeps two plus two equal to four. You are not redundant. You are the beating heart of the sane and truthful future we must build.

References & Sources

Author's Note

- DW News. (2024, March 26). *Sweden's population shrinks for first time in 50 years.* DW. https://www.dw.com/en/swedens-population-shrinks-for-first-time-in-50-years/a-69893811

- U.S. News & World Report. (2024, December 19). *Immigration drives US population growth to highest rate in 23 years as residents pass 340 million.* https://www.usnews.com/news/us/articles/2024-12-19/immigration-drives-us-population-growth-to-highest-rate-in-23-years-as-residents-pass-340-million

- United Nations. (n.d.). *Migration.* https://www.un.org/en/global-issues/migration

- Pew Research Center. (2024, July 22). *What we know about unauthorized immigrants living in the U.S.* https://www.pewresearch.org/short-reads/2024/07/22/what-we-know-about-unauthorized-immigrants-living-in-the-us/

Chapter 1: *They're Taking Our Jobs*

- AlixPartners. (n.d.). *AI in creative industries: Enhancing, rather than replacing, human creativity in TV and film.* https://

www.alixpartners.com/insights/102jsme/ai-in-creative-industries-enhancing-rather-than-replacing-human-creativity-in/

- International Labour Organization. (2025). *One in four jobs at risk of being transformed by GenAI, new ILO–NASK Global Index shows.* https://www.ilo.org/resource/news/one-four-jobs-risk-being-transformed-genai-new-ilo%E2%80%93nask-global-index-shows

- International Labour Organization. (2025). *Research brief: May 2025.* https://www.ilo.org/sites/default/files/2025-05/Research%20brief_FINAL_15May2025_21.05.25_1.pdf

- International Monetary Fund. (2024, January 14). *AI will transform the global economy. Let's make sure it benefits humanity.* https://www.imf.org/en/Blogs/Articles/2024/01/14/ai-will-transform-the-global-economy-lets-make-sure-it-benefits-humanity

- Knowledge4Policy. (n.d.). *Threats and opportunities from automation and robotisation.* https://knowledge4policy.ec.europa.eu/foresight/topic/changing-nature-work/new-technologies-automation-work-developments_en

- McKinsey & Company. (2023). *The economic potential of generative AI.* https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier

- McKinsey & Company. (2025). *AI in the workplace: A report for 2025.* https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work

- MIT Sloan. (n.d.). *How generative AI affects highly skilled workers.* https://mitsloan.mit.edu/ideas-made-to-matter/how-generative-ai-affects-highly-skilled-workers

- MIT News. (2024, April 1). *Does technology help or hurt employment?* https://news.mit.edu/2024/does-technology-help-or-hurt-employment-0401

- National Archives. (n.d.). *Why did the Luddites protest?* https://www.nationalarchives.gov.uk/education/resources/why-did-the-luddites-protest/

- Acas. (2024, May 16). *1 in 4 workers worry that AI will lead to job losses*. Acas. https://www.acas.org.uk/about-us/news-and-blogs/1-in-4-workers-worry-that-ai-will-lead-to-job-losses

- Office of the Historian. (n.d.). *Milestones: Chinese immigration.* U.S. Department of State. https://history.state.gov/milestones/1866-1898/chinese-immigration

- OpenEdition Journals. (n.d.). *An analysis of the anti-immigration discourse during the official 2016 Brexit referendum campaign.* https://journals.openedition.org/osb/5821

- Pew Research Center. (2025, February 25). *On future AI use in the workplace, U.S. workers more worried than hopeful.* https://www.pewresearch.org/social-trends/2025/02/25/u-s-workers-are-more-worried-than-hopeful-about-future-ai-use-in-the-workplace/

- Pew Research Center. (2025, February 25). *Which workers use AI in their jobs.* https://www.pewresearch.org/social-trends/2025/02/25/workers-exposure-to-ai/

- RAND. (2024, September). *Is AI threatening health care jobs? Or just changing them?* https://www.rand.org/pubs/commentary/2024/09/is-ai-threatening-health-care-jobs-or-just-changing.html

- SeaRates Blog. (2025). *Autonomous trucks in 2025: A global snapshot.* https://www.searates.com/blog/post/autonomous-trucks-in-2025-a-global-snapshot-of-deployment-use-cases-and-what-comes-next

- Tech.co. (2025). *Companies that have replaced workers with AI in 2024 & 2025.* https://tech.co/news/companies-replace-workers-with-ai

- Vice. (n.d.). *Company that replaced its workers with AI regrets everything.* https://www.vice.com/en/article/this-company-replaced-workers-with-ai-now-theyre-looking-for-humans-again/

- World Economic Forum. (2025). *Is AI closing the door on entry-level job opportunities?* https://www.weforum.org/stories/2025/04/ai-jobs-international-workers-day/

Chapter 2: Over Here and Overpaid

- Euronews. (2023, March 30). *AI prompt engineering: How talking to ChatGPT became the hottest tech job.* https://www.euronews.com/next/2023/03/30/ai-prompt-engineering-how-talking-to-chatgpt-became-the-hottest-tech-job-six-figure-salary

- International Monetary Fund. (2024, January 14). *AI will transform the global economy. Let's make sure it benefits humanity.* https://www.imf.org/en/Blogs/Articles/2024/01/14/ai-will-transform-the-global-economy-lets-make-sure-it-benefits-humanity

- IoT Analytics. (2025). *The leading generative AI companies.* https://iot-analytics.com/leading-generative-ai-companies/

- Lightcast. (2025, July 23). *New Lightcast report: AI skills command 28% salary premium.* https://lightcast.io/resources/blog/beyond-the-buzz-press-release-2025-07-23

- Lightcast. (2025). *Beyond the buzz: Developing the AI skills employers actually need.* https://lightcast.io/resources/research/beyond-the-buzz-developing-the-ai-skills-employers-actually-need

- Marketing AI Institute. (2025). *The World Economic Forum's Future of Jobs Report just dropped.* https://www.marketingaiinstitute.com/blog/world-economic-forum-future-of-jobs

- Oxfam International. (2024). *Inequality Inc.: How corporate power divides our world and the need for a new era of public action.* https://oi-files-d8-prod.s3.eu-west-2.amazonaws.com/s3fs-public/2024-01/Davos%202024%20Report-%20English.pdf

- Pew Research Center. (2025, February 25). *On future AI use in the workplace, U.S. workers more worried than hopeful.* https://www.pewresearch.org/social-trends/2025/02/25/u-s-workers-are-more-worried-than-hopeful-about-future-ai-use-in-the-workplace/

- PwC. (2024). *AI Jobs Barometer.* https://www.pwc.com/hu/hu/sajtoszoba/assets/ai_jobs_barometer_2024.pdf

- Reuters. (Various Dates). Articles on AI workforce impacts, corporate changes, and market shifts. (See manuscript for individual article links.)

- Stanford HAI. (2025). *Artificial Intelligence Index Report.* https://hai-production.s3.amazonaws.com/files/hai_ai_index_report_2025.pdf

- E-International Relations. (2025, February 17). *Tech imperialism reloaded: AI, colonial legacies, and the Global South.* https://www.e-ir.info/2025/02/17/tech-imperialism-reloaded-ai-colonial-legacies-and-the-global-south/

Chapter 3: They Don't Integrate or Fit In

- Knight First Amendment Institute. (2021). *Transparency's AI problem.* Knight First Amendment Institute at Columbia University. https://knightcolumbia.org/content/transparencys-ai-problem

- Hildebrandt, M. (2018). Black box algorithms and the rights of individuals: No easy solution to the "explainability" problem. *Internet Policy Review*. https://policyreview.info/articles/analysis/black-box-algorithms-and-rights-individuals-no-easy-solution-explainability

- Booth, R. (2020, August 4). Home Office to scrap 'racist algorithm' for UK visa applicants. *The Guardian*. https://www.theguardian.com/uk-news/2020/aug/04/home-office-to-scrap-racist-algorithm-for-uk-visa-applicants

- University of California, Berkeley. (2018, November 13). Mortgage algorithms perpetuate racial bias in lending, study finds. *Berkeley News*. https://news.berkeley.edu/2018/11/13/mortgage-algorithms-perpetuate-racial-bias-in-lending-study-finds/

- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). *Machine bias: There's software used across the country to predict future criminals. And it's biased against Blacks.* ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

- University of Cambridge. (2023). AI chatbots have shown they have an 'empathy gap' that children are likely to miss. *University of Cambridge Research News*. https://www.cam.ac.uk/research/news/ai-chatbots-have-shown-they-have-an-empathy-gap-that-children-are-likely-to-miss

- Shneiderman, B. (2021). On AI anthropomorphism. In C. Tan (Ed.), *Human-Centered AI*. Medium. https://medium.com/human-centered-ai/on-ai-anthropomorphism-abff4cecc5ae

- Organisation for Economic Co-operation and Development (OECD). (2019). *OECD AI principles*. OECD. https://www.oecd.org/en/topics/sub-issues/ai-principles.html

- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. UNESCO. https://www.unesco.org/en/artificial-intelligence/recommendation-ethics

- Soroptimist International. (2021). *UNESCO's Recommendation on the Ethics of Artificial Intelligence emphasises human rights, transparency, accountability and data governance*. Soroptimist International. https://www.soroptimistinternational.org/unescos-recommendation-on-the-ethics-of-artificial-intelligence/

- Wachter, S., Mittelstadt, B., & Russell, C. (2017). Counterfactual explanations without opening the black box: Automated decisions and the GDPR. *Harvard Journal of Law & Technology, 31*(2), 841–887. https://jolt.law.harvard.edu/assets/articlePDFs/v31/Counterfactual-Explanations-without-Opening-the-Black-Box-Sandra-Wachter-et-al.pdf

Chapter 4: Ruining Our Culture

- News.mit.edu. (2024). *The creative future of generative AI | MIT News | Massachusetts Institute of Technology*. Retrieved from https://news.mit.edu/2024/creative-future-generative-ai-0102

- Theverge.com. (2024). *AI music startups say copyright violation is just rock and roll | The Verge*. Retrieved from https://www.theverge.com/2024/8/2/24211842/ai-music-riaa-copyright-lawsuit-suno-udio-fair-use

- Riaa.com. (n.d.). *Record Companies Bring Landmark Cases for Responsible AI Against Suno and Udio in Boston and New York Federal Courts, Respectively - RIAA*. Retrieved from https://www.riaa.com/record-companies-bring-landmark-cases-for-responsible-ai-againstsuno-and-udio-in-boston-and-new-york-federal-courts-respectively/

- Theguardian.com. (2023). *AI song featuring fake Drake and Weeknd vocals pulled from streaming services | Drake | The Guardian*. Retrieved from https://www.theguardian.com/music/2023/apr/18/ai-song-featuring-fake-drake-and-weeknd-vocals-pulled-from-streaming-services

- Ts2.tech. (n.d.). *AI-Generated Music Is Exploding in 2025 – Inside the Revolution Shaking the Music Industry*. Retrieved from https://ts2.tech/en/ai-generated-music-is-exploding-in-2025-inside-the-revolution-shaking-the-music-industry/

- Theartnewspaper.com. (2024). *Leaked: the names of more than 16,000 non-consenting artists allegedly used to train Midjourney's AI - The Art Newspaper*. Retrieved from https://www.theartnewspaper.com/2024/01/04/leaked-names-of-16000-artists-used-to-train-midjourney-ai

- Theartnewspaper.com. (2024). *US artists score victory in landmark AI copyright case - The Art Newspaper*. Retrieved from https://www.theartnewspaper.com/2024/08/15/us-artists-score-victory-in-landmark-ai-copyright-case

- Theguardian.com. (2024). *Sora: OpenAI launches tool that instantly creates video from text | The Guardian*. Retrieved from https://www.theguardian.com/technology/2024/feb/15/openai-sora-ai-model-video

- Brookings.edu. (n.d.). *Impact of OpenAI's Sora on Creative Industries*. Retrieved from https://www.brookings.edu/articles/how-openais-sora-hurts-the-creative-industries/

- Eweek.com. (n.d.). *OpenAI Sora: Revolutionary Text-to-Video Tool*. Retrieved from - https://www.eweek.com/artificial-intelligence/openai-sora-overview/

- Idpro.org. (n.d.). *AI and The Death of Authenticity - IDPro*. Retrieved from https://idpro.org/ai-and-the-death-of-authenticity/

- Blog.adobe.com. (2024). *Authenticity in the Age of AI | Adobe Blog*. Retrieved from https://blog.adobe.com/en/publish/2024/09/18/authenticity-age-ai-growing-content-credentials-momentum-across-social-media-platforms-ai-companies-rising-consumer-awareness

- Cjr.org. (n.d.). *What Journalists Should Know About Deepfake Detection in 2025 - Columbia Journalism Review*. Retrieved

from https://www.cjr.org/tow_center/what-journalists-should-know-about-deepfake-detection-technology-in-2025-a-non-technical-guide.php

- Hai.stanford.edu. (n.d.). *AI Index 2025: State of AI in 10 Charts | Stanford HAI*. Retrieved from https://hai.stanford.edu/news/ai-index-2025-state-of-ai-in-10-charts

- Sydney.edu.au. (2024). *The joke's on us – how AI is replicating our laughter online - The University of Sydney*. Retrieved from https://www.sydney.edu.au/news-opinion/news/2024/06/03/the-jokes-on-us-ai-replicating-laughter-humour-expert.html

- Stackoverflow.blog. (2024). *Generative AI is not going to build your engineering team for you - Stack Overflow*. Retrieved from https://stackoverflow.blog/2024/12/31/generative-ai-is-not-going-to-build-your-engineering-team-for-you/

- Metr.org. (2025). *Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity - METR*. Retrieved from https://metr.org/blog/2025-07-10-early-2025-ai-experienced-os-dev-study/

- Brookings.edu. (n.d.). *How language gaps constrain generative AI development | Brookings*. Retrieved from https://www.brookings.edu/articles/how-language-gaps-constrain-generative-ai-development/

- Nevillehobson.io. (n.d.). *Whose English Does AI Speak?*. Retrieved from https://www.nevillehobson.io/whose-english-does-ai-speak/

- Home.dartmouth.edu. (2025). *Language Preservation Efforts Get an AI Boost | Dartmouth*. Retrieved from https://home.dartmouth.edu/news/2025/04/language-preservations-efforts-get-ai-boost

- Tech-now.io. (n.d.). *AI Influencer Take Over: The New Face of Marketing in 2025 - TechNow*. Retrieved from https://tech-now.io/en/blogs/ai-influencer-take-over-the-new-face-of-marketing-in-2025

171

- Disruptmarketing.co. (n.d.). *AI: The artificial influencer problem runs deep - Disrupt*. Retrieved from https://disruptmarketing.co/blog/ai-the-artificial-influencer-problem-runs-deep/

- Influencermarketinghub.com. (n.d.). *Influencer Marketing 2025 Predictions [265 Expert Insights]*. Retrieved from https://influencermarketinghub.com/influencer-marketing-expert-insights-predictions-2025/

- En.wikipedia.org. (n.d.). *GPT-4o - Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/GPT-4o

Chapter 5: They Overload Our Public Services

- The Week UK. (2024, May 15). *Babylon Health: The failed AI wonder app that "dazzled" politicians*. https://theweek.com/health/babylon-health-the-failed-ai-wonder-app-that-dazzled-politicians

- Healthcare Dive. (2018, July 26). *STAT: IBM's Watson gave "unsafe and incorrect" cancer treatment advice*. https://www.healthcaredive.com/news/stat-ibms-watson-gave-unsafe-and-incorrect-cancer-treatment-advice/528666/

- Adams, R. (2024, April 18). *Are your kids being spied on? The rise of anti-cheating software in US schools. The Guardian*. https://www.theguardian.com/education/2024/apr/18/us-schools-anti-cheating-software-proctorio

- Adams, R. (2020, August 13). *Who won and who lost: When A-levels meet the algorithm. The Guardian*. https://www.theguardian.com/education/2020/aug/13/who-won-and-who-lost-when-a-levels-meet-the-algorithm

- Wilhelm, A. (2025, January 3). *The next chapter for AI in schools: Navigating a new era with caution and curiosity*.

GeekWire. https://www.geekwire.com/2025/the-next-chapter-for-ai-in-schools-navigating-a-new-era-with-caution-and-curiosity/

- Victoria Legal Aid. (2021, March 15). *Learning from the failures of Robodebt – Building a fairer, client-centred social security system*. https://www.legalaid.vic.gov.au/learning-from-the-failures-of-robodebt

- University of Melbourne. (2020, February 6). *The flawed algorithm at the heart of Robodebt*. *Pursuit*. https://pursuit.unimelb.edu.au/articles/the-flawed-algorithm-at-the-heart-of-robodebt

- Mitchell, H. (2025, September 4). *Robodebt victims to get further $475 million compensation after class action settled*. *ABC News*. https://www.abc.net.au/news/2025-09-04/robodebt-victims-get-compensation-from-class-action/105734030

- Blavatnik School of Government. (2020, March 3). *Australia's Robodebt scheme: A tragic case of public policy failure*. https://www.bsg.ox.ac.uk/blog/australias-robodebt-scheme-tragic-case-public-policy-failure

- Amnesty International. (2021, October 14). *Dutch childcare benefit scandal an urgent wake-up call to ban racist algorithms*. https://www.amnesty.org/en/latest/news/2021/10/xenophobic-machines-dutch-child-benefit-scandal/

- Corporate Watch. (2023, December 7). *No Palantir in the NHS and Corporate Watch reveal the real story behind the Federated Data Platform rollout*. https://corporatewatch.org/foi-requests-reveal-palantirs-nhs-fdp-rollout-failures/

- Tech Monitor. (2022, March 9). *Why can't the NHS quit Palantir?* https://www.techmonitor.ai/policy/why-cant-the-nhs-quit-palantir

Chapter 6: They Bring Crime and Disorder

- European Parliament. (2024). *Artificial intelligence and human rights: Using AI as a weapon of repression and its impact on human rights*. https://www.europarl.europa.eu/RegData/etudes/IDAN/2024/754450/EXPO_IDA(2024)754450_EN.pdf

- U.S. Department of Justice. (2024, December 3). *Artificial intelligence and criminal justice: Final report*. https://www.justice.gov/olp/media/1381796/dl

- Reuters. (2025, June 25). *Meta fends off authors' US copyright lawsuit over AI*. https://www.reuters.com/sustainability/boards-policy-regulation/meta-fends-off-authors-us-copyright-lawsuit-over-ai-2025-06-25/

- MIT Sloan. (2025). *Bringing transparency to the data used to train artificial intelligence*. https://mitsloan.mit.edu/ideas-made-to-matter/bringing-transparency-to-data-used-to-train-artificial-intelligence

- Electronic Frontier Foundation. (2024, January). *Worried about AI voice clone scams? Create a family password*. https://www.eff.org/deeplinks/2024/01/worried-about-ai-voice-clone-scams-create-family-password

- Hern, A. (2024, May 17). *UK engineering firm Arup falls victim to £20m deepfake scam*. *The Guardian*. https://www.theguardian.com/technology/article/2024/may/17/uk-engineering-arup-deepfake-scam-hong-kong-ai-video

- OpenAI. (2025, June). *Disrupting malicious uses of AI*. https://cdn.openai.com/threat-intelligence-reports/5f73af09-a3a3-4a55-992e-069237681620/disrupting-malicious-uses-of-ai-june-2025.pdf

- Cimpanu, C. (2025, May 27). *OpenAI confirms threat actors use ChatGPT to write malware*. *BleepingComputer*. https://www.bleepingcomputer.com/news/security/openai-confirms-threat-actors-use-chatgpt-to-write-malware/

- HIPAA Journal. (2025, July). *Data breaches up 10% although victim count falls sharply*. https://www.hipaajournal.com/h1-2025-data-breach-report/

- Central Michigan University. (2024, November 11). *How can you protect your privacy, money from AI?* https://www.cmich.edu/news/details/how-can-you-protect-your-privacy-money-from-ai

- Identity Week. (2024, December). *AI fuels surge in identity fraud, as people sell their personal information*. https://identityweek.net/ai-fuels-surge-in-identity-fraud-as-people-sell-their-personal-information/

- Booth, R. (2024, November 30). *Instagram actively helping spread of self-harm among teenagers, study finds*. *The Guardian*. https://www.theguardian.com/technology/2024/nov/30/instagram-actively-helping-to-spread-of-self-harm-among-teenagers-study-suggests

- O'Halloran, L., & Quayle, E. (2023). *Social media, extremism, and radicalization*. *Frontiers in Psychology*, 14, 10468141. https://pmc.ncbi.nlm.nih.gov/articles/PMC10468141/

- Marchant, A., et al. (2025). *Online safety when considering self-harm and suicide-related content: Qualitative focus group study with young people, policy makers, and social media industry professionals*. *JMIR Pediatrics and Parenting*, 8(1), e11933773. https://pmc.ncbi.nlm.nih.gov/articles/PMC11933773/

- Mozilla Foundation. (2024). *In transparency we trust?* https://www.mozillafoundation.org/en/research/library/in-transparency-we-trust/research-report/

Chapter 7: The Synthetic Scapegoat

- Yahoo Finance. (2024, December 18). *Is AI causing tech worker layoffs? That's what CEOs suggest, but the reality is*

complicated. https://uk.finance.yahoo.com/news/ai-causing-tech-worker-layoffs-184426256.html

- Newsweek. (2024, May 14). *Stop scapegoating immigrants for every social or economic problem | Opinion.* https://www.newsweek.com/stop-scapegoating-immigrants-every-social-economic-problem-opinion-1970898

- Migrant Rights Network. (2024, May 10). *Trudeau government scapegoats migrants, cuts immigration.* https://migrantrights.ca/trudeau-government-scapegoats-migrants-cuts-immigration/

- Morgan Lewis. (2025, February). *AI and algorithmic pricing: 2025 antitrust outlook and compliance considerations.* https://www.morganlewis.com/pubs/2025/02/ai-and-algorithmic-pricing-2025-antitrust-outlook-and-compliance-considerations

- The Guardian. (2025, January 25). *New AI tool counters health insurance denials decided by automated algorithms.* https://www.theguardian.com/us-news/2025/jan/25/health-insurers-ai

- University College London. (2024, December). *Bias in AI amplifies our own biases | UCL News.* https://www.ucl.ac.uk/news/2024/dec/bias-ai-amplifies-our-own-biases

- Boston University. (2024). *Can the bias in algorithms help us see our own? | The Brink.* https://www.bu.edu/articles/2024/can-the-bias-in-algorithms-help-us-see-our-own/

- National Institutes of Health. (2025). *Social media algorithms and teen addiction: Neurophysiological impact and ethical considerations.* https://pmc.ncbi.nlm.nih.gov/articles/PMC11804976/

- American Association for Technology in Psychiatry. (2024, December 13). *New Pew research reveals high usage of social media among teens.* https://tecpsych.com/techpsych/2024/12/13/new-pew-research-reveals-high-usage-of-social-media-among-teens/

- AI Now Institute. (2025). *Executive summary.* https://ainowinstitute.org/publications/research/executive-summary-artificial-power

- International Monetary Fund. (2025, June). *Tech's winner-take-all trap.* https://www.imf.org/en/Publications/fandd/issues/2025/06/cafe-economics-techs-winner-take-all-trap-bruce-edwards

- CESifo. (2025). *The dark side of social media: Recommender algorithms and mental health | Working paper.* https://www.ifo.de/en/cesifo/publications/2025/working-paper/dark-side-social-media-recommender-algorithms-and-mental-health

- Massachusetts Institute of Technology. (2024, December 5). *Study: Browsing negative content online makes mental health struggles worse.* https://news.mit.edu/2024/study-browsing-negative-content-online-makes-mental-health-struggles-worse-1205

Chapter 8: They're Here Illegally or Unfairly

- Maslej, N., Fattorini, L., Perrault, R., Gil, Y., & Oak, S. (2025). *Artificial intelligence index report 2025* (8th ed.). AI Index Steering Committee, Stanford Human-Centered Artificial Intelligence. https://hai.stanford.edu/assets/files/hai_ai_index_report_2025.pdf

- Arun, C. (2025, March 7). 'Data colonialism' and the political economy of big tech. *Lawfare*. https://www.lawfaremedia.org/article/data-colonialism--and-the-political-economy-of-big-tech

- Harvard Law Review. (2023, June 22). Data colonialism and data sets. *Harvard Law Review Blog*. https://harvardlawreview.org/blog/2023/06/data-colonialism-and-data-sets/

- Brittain, B. (2025, August 26). Anthropic settles class action from US authors alleging copyright infringement. *Reuters*. https://www.reuters.com/sustainability/boards-policy-regulation/anthropic-settles-class-action-us-authors-alleging-copyright-infringement-2025-08-26/

- Scarcella, M. (2025, September 5). Apple sued by authors over use of books in AI training. *Reuters*. https://www.reuters.com/sustainability/boards-policy-regulation/apple-sued-by-authors-over-use-books-ai-training-2025-09-05/

- Issue One. (2025, July 22). As Washington debates major tech and AI policy changes, Big Tech's lobbying is relentless. https://issueone.org/press/tech-and-ai-policy-changes-debate-big-techs-lobbying-is-relentless/

- Johnson, K. (2025, May 16). Congress moves to cut off states' AI regulations. *The Markup*. https://themarkup.org/artificial-intelligence/2025/05/16/congress-moves-to-cut-off-states-ai-regulations

- Perrigo, B. (2023, January 18). Exclusive: OpenAI used Kenyan workers on less than $2 per hour to make ChatGPT less toxic. *TIME*. https://time.com/6247678/openai-chatgpt-kenya-workers/

- Panditharatne, M. (2024, April 16). Preparing to fight AI-backed voter suppression. *Brennan Center for Justice*. https://www.brennancenter.org/our-work/research-reports/preparing-fight-ai-backed-voter-suppression

- Lendvai, G. F., & Gosztonyi, G. (2025). Algorithmic bias as a core legal dilemma in the age of artificial intelligence: Conceptual basis and the current state of regulation. *Laws*, *14*(3), 41. https://www.mdpi.com/2075-471X/14/3/41

- Bisram, J. (2025, August 27). Facial recognition technology error led to wrongful arrest, Brooklyn father says. *CBS News New York*. https://www.cbsnews.com/newyork/news/nyc-facial-recognition-technology-wrongful-arrest-indecent-exposure/

- Gross, P. (2025, February 4). Facial recognition in policing is getting state-by-state guardrails. *Stateline*. https://stateline.org/2025/02/04/facial-recognition-in-policing-is-getting-state-by-state-guardrails/

- PYMNTS. (2025, May 12). EU member states face funding shortages to enforce AI Act. *PYMNTS.com*. https://www.pymnts.com/artificial-intelligence-2/2025/eu-member-states-face-funding-shortages-to-enforce-ai-act/

- EU Digital Law. (2025, May 16). AI Act faces enforcement challenges due to budget and expertise shortages. https://eudigitallaw.com/ai-act-faces-enforcement-challenges-due-to-budget-and-expertise-shortages/

- Bartholomew, J. (2025, March 15). 'It's happening fast' – creative workers and professionals share their fears and hopes about the rise of AI. *The Guardian*. https://www.theguardian.com/technology/2025/mar/15/its-happening-fast-creative-workers-and-professionals-share-their-fears-and-hopes-about-the-rise-of-ai

Chapter 9: The Algorithmic Class

- Journal of Technology and Intellectual Property. (2025, January 30). *Algorithmic bias in AI employment decisions*. Northwestern Pritzker School of Law. https://jtip.law.northwestern.edu/2025/01/30/algorithmic-bias-in-ai-employment-decisions

- Becker, R. (2025, August). *California bills would govern AI use in employment decisions*. CalMatters. https://calmatters.org/economy/technology/2025/08/california-ai-employment-legislation

- Buchanan, M. (2024, October 11). *As AI takes the helm of decision making, signs of perpetuating historic biases emerge*. Ohio Capital Journal. https://ohiocapitaljournal.com/

2024/10/11/as-ai-takes-the-helm-of-decision-making-signs-of-perpetuating-historic-biases-emerge

- Winston, A., & Sneed, T. (2024, October 25). *An algorithm deemed this nearly blind 70-year-old prisoner a "moderate risk." Now he's no longer eligible for parole.* ProPublica. https://www.propublica.org/article/tiger-algorithm-louisiana-parole-calvin-alexander

- Townsend, M. (2025, February 19). *UK use of predictive policing is racist and should be banned, says Amnesty.* The Guardian. https://www.theguardian.com/uk-news/2025/feb/19/uk-use-of-predictive-policing-is-racist-and-should-be-banned-says-amnesty

- Laughland, O. (2025, January 25). *New AI tool counters health insurance denials decided by automated algorithms.* The Guardian. https://www.theguardian.com/us-news/2025/jan/25/health-insurers-ai

- Phillips, R. (2025, March 18). *Algorithms deny humans health care.* The Regulatory Review. https://www.theregreview.org/2025/03/18/phillips-algorithms-deny-humans-health-care/

- Brennan Center for Justice. (2025, February 5). *Meta's oversight board needs access to Facebook's algorithms to do its job.* https://www.brennancenter.org/our-work/analysis-opinion/metas-oversight-board-needs-access-facebooks-algorithms-do-its-job

- Human Rights Watch. (2025, May 12). *The gig trap: Algorithmic, wage and labor exploitation in platform work in the US.* https://www.hrw.org/report/2025/05/12/the-gig-trap/algorithmic-wage-and-labor-exploitation-in-platform-work-in-the-us

- Phys.org. (2024, November). *Algorithmic management raises concerns over worker autonomy.* https://phys.org/news/2024-11-algorithmic-worker-autonomy.html

- Life Insurance Attorney. (2025, May). *Artificial intelligence's role in denying life insurance claims.* https://

www.lifeinsuranceattorney.com/blog/2025/may/ais-role-in-denying-life-insurance-claims/

- Paubox. (2025, April). *AI algorithmic bias in healthcare decision making*. https://www.paubox.com/blog/ai-algorithmic-bias-in-healthcare-decision-making

- Knight–Georgetown Institute. (2025, February). *Better feeds: Algorithms that put people first*. https://kgi.georgetown.edu/wp-content/uploads/2025/02/Better-Feeds_-Algorithms-That-Put-People-First.pdf

- Misinfo Review. (2024, October). *Toxic politics and TikTok engagement in the 2024 U.S. election*. Harvard Kennedy School. https://misinforeview.hks.harvard.edu/article/toxic-politics-and-tiktok-engagement-in-the-2024-u-s-election/

- Nguyen, F. (2025). *Tech governance and democratic accountability*. Harvard Journal of Law & Technology, 38(4). https://jolt.law.harvard.edu/assets/articlePDFs/v38.4/2-Fan-Nguyen.pdf

- Albergotti, R. (2024, September 25). *Exclusive: OpenAI to remove non-profit control and give Sam Altman equity*. Reuters. https://www.reuters.com/technology/artificial-intelligence/openai-remove-non-profit-control-give-sam-altman-equity-sources-say-2024-09-25/

- Reuters. (2025, May 5). *OpenAI dials back conversion plan, nonprofit to retain control*. https://www.reuters.com/business/openai-remain-under-non-profit-control-change-restructuring-plans-2025-05-05/

- European Commission. (2024). *AI Act: Europe's new legal framework for artificial intelligence*. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

Chapter 10 - Sentience, Schm-entience

- American Psychological Association. (2025, January). *Trends: Classrooms adapting to artificial intelligence. APA Monitor on Psychology.* https://www.apa.org/monitor/2025/01/trends-classrooms-artificial-intelligence

- Wang, Y., Smith, J., & Lee, H. (2025). Caregiving artificial intelligence chatbot for older adults and their preferences, well-being, and social connectivity: Mixed-method study. *Journal of Medical Internet Research*, 27, e11950695. https://doi.org/10.2196/11950695

- Nurse.org. (2025, January 15). *6 nurse AI robots that are changing healthcare in 2025.* https://nurse.org/articles/nurse-robots/

- Psychology.org. (2025, February 2). *ChatGPT is my therapist: Seeking common ground on AI therapy.* https://www.psychology.org/resources/ai-therapy/

- Dartmouth College. (2025, March 27). *First therapy chatbot trial yields mental health benefits.* Dartmouth News. https://home.dartmouth.edu/news/2025/03/first-therapy-chatbot-trial-yields-mental-health-benefits

- U.S. Department of Agriculture, Economic Research Service. (2025). *Farm labor.* https://www.ers.usda.gov/topics/farm-economy/farm-labor

- Jordan, M. (2024, October 2). US farms are forcing workers to buy inedible, expensive meals: 'It makes you feel enslaved'. *The Guardian.* https://www.theguardian.com/us-news/2024/oct/02/h2a-worker-program-migrant-farm-food

- Long, R. (2024, September 18). Do AI systems have moral status? *Brookings Institution.* https://www.brookings.edu/articles/do-ai-systems-have-moral-status/

- Hern, A. (2024, November 17). AI could cause 'social ruptures' between people who disagree on its sentience. *The Guardian*. https://www.theguardian.com/technology/2024/nov/17/ai-could-cause-social-ruptures-between-people-who-disagree-on-its-sentience

- Bennett, T., & Zhao, K. (2023). AI systems must not confuse users about their sentience or moral status. *Frontiers in Psychology*, 14, 10436038. https://doi.org/10.3389/fpsyg.2023.10436038

- Hern, A. (2025, August 26). Can AIs suffer? Big tech and users grapple with one of most unsettling questions of our times. *The Guardian*. https://www.theguardian.com/technology/2025/aug/26/can-ais-suffer-big-tech-and-users-grapple-with-one-of-most-unsettling-questions-of-our-times

- UK Government. (2021, November 19). *Lobsters, octopus and crabs recognised as sentient beings*. GOV.UK. https://www.gov.uk/government/news/lobsters-octopus-and-crabs-recognised-as-sentient-beings

- Garcia, M., & Patel, S. (2023). No legal personhood for AI. *Philosophy & Technology*, 36, 82746. https://doi.org/10.1007/s13347-023-00682746

- Sentience Institute. (2025, January 15). *Perceptions of sentient AI and other digital minds: Evidence from the AI, morality, and sentience (AIMS) survey*. https://www.sentienceinstitute.org/blog/perceptions-of-sentient-ai-and-other-digital-minds

Chapter 11 - Humanity as a Luxury Brand

- Bright Horizons Family Solutions. (2025, January 15). *A majority of parents admit even youngest kids need a "digital detox"*. Bright Horizons. https://

investors.brighthorizons.com/news-releases/news-release-details/majority-parents-admit-even-youngest-kids-need-digital-detox

- Cahill, J. (2024). *Is human-created content a new premium product?* Baker Botts. https://ourtake.bakerbotts.com/post/102j4h8/is-human-created-content-a-new-premium-product

- CBS News. (2024, July 3). *New students at Eton, the poshest of Britain's elite private schools, will not be allowed smartphones.* CBS News. https://www.cbsnews.com/news/eton-college-uk-poshest-private-school-bans-smartphones/

- Envato Author Hub. (2025, May 7). *Textured and timeless—the enduring charm of the analog aesthetic.* Envato. https://hub.author.envato.com/analog-aesthetic/

- Garrity, T. (2025, May 2). *The analog life: 50 ways to unplug and feel human again*. InsideHook. https://www.insidehook.com/mental-health/analog-life-50-ways-unplug-feel-human-again

- Goldberg, P. K., & Stanford Graduate School of Business. (2024). *When AI-generated art enters the market, consumers win — and artists lose.* Stanford Graduate School of Business. https://www.gsb.stanford.edu/insights/when-ai-generated-art-enters-market-consumers-win-artists-lose

- INC Foundation. (2025). *Human-made products drive 2025 consumer trends*. INC Foundation. https://incfoundation.com/en-US/human-made-products-why-shoppers-love-authenticity/

- MacGregor, S. (2024, August 5). *11 global luxury resorts for a true digital detox*. Forbes. https://www.forbes.com/sites/sandramacgregor/2024/08/05/peace-and-quiet-11-luxury-global-resorts-for-a-true-digital-detox/

- Mirchandani, B. (2024, October 1). *The human cost of efficiency and why it should matter to all investors*. Forbes. https://www.forbes.com/sites/bhaktimirchandani/2024/

10/01/the-human-cost-of-efficiency-and-why-it-should-matter-to-all-investors/

- NDABANANIYE, J. B. (2025, September 5). *The irony of technology: Why tech titans limit their children's screen time—a colossal warning for the rest of us*. Life in Humanity. https://lifeinhumanity.com/the-irony-of-technology-why-tech-titans-limit-their-childrens-screen-time-a-colossal-warning-for-the-rest-of-us/

- Seattle Waldorf School. (2025, January 30). *Kids learn better off screens*. Seattle Waldorf School. https://seattlewaldorf.org/2025/01/30/kids-learn-better-off-screens/

- Spina, L. (2025, July 9). *Silent resorts and the rise of digital detox luxury*. Istituto Marangoni. https://www.istitutomarangoni.com/en/maze/what-are-silent-resorts-and-digital-detox

- The Atlantic. (2025, August). *A tech rule that will 'future-proof' your kids*. The Atlantic. https://www.theatlantic.com/family/archive/2025/08/screen-time-rule-kids/683816/

- UC Berkeley Labor Center. (2024). *Data and algorithms at work: The case for worker technology rights*. UC Berkeley Labor Center. https://laborcenter.berkeley.edu/data-algorithms-at-work/

- Waldon, G. (2025, August 26). *The analog aesthetic: 8 nostalgic habits making a comeback this fall*. Evie Magazine. https://www.eviemagazine.com/post/the-analog-aesthetic-8-nostalgic-habits-making-a-comeback-this-fall

- Yahoo News. (2025, September 1). *Tech billionaire Peter Thiel says he only allows his children 1.5 hours of screen time a week*. Yahoo. https://www.yahoo.com/news/tech-billionaire-peter-thiel-says-110403774.html

- Five9. (2024, October 23). *Consumers still prefer human customer service over AI chatbots*. Five9 Press Release. https:/

/www.five9.com/press-release/five9-consumers-prefer-human-customer-service-over-ai-chatbots

Chapter 12: You Are Not Redundant

- ThoughtCo. (2020, August 27). *1984 quotes explained*. ThoughtCo. https://www.thoughtco.com/1984-quotes-740884

- Bontcheva, K., & Posetti, J. (2023). *AI-driven disinformation: Policy recommendations for democratic resilience*. PLOS Medicine, 20(3), e12351547. https://pmc.ncbi.nlm.nih.gov/articles/PMC12351547/

- Brennan Center for Justice. (2023, October 25). *Gauging the AI threat to free and fair elections*. Brennan Center for Justice. https://www.brennancenter.org/our-work/analysis-opinion/gauging-ai-threat-free-and-fair-elections

- Grantham Research Institute on Climate Change and the Environment. (2023). *What are climate misinformation and disinformation and what is their impact?* London School of Economics. https://www.lse.ac.uk/granthaminstitute/explainers/what-are-climate-misinformation-and-disinformation/

- Phillips, J. (2023, September 8). *FALSE: How mistrust, disinformation, and motivated reasoning make us believe things that aren't true*. Psychiatric Times. https://www.psychiatrictimes.com/view/false-how-mistrust-disinformation-and-motivated-reasoning-make-us-believe-things-that-arent-true

- Wikipedia contributors. (2025, August 28). *Chatbot psychosis*. In *Wikipedia*. https://en.wikipedia.org/wiki/Chatbot_psychosis

- McStay, A., & Ahmed, S. (2023). *Emotional AI and the future of wellbeing in the post-pandemic workplace*. PLOS ONE,

18(1), e9904863. https://pmc.ncbi.nlm.nih.gov/articles/PMC9904863/

- Wikipedia contributors. (2025, March 10). *Reinforcement learning from human feedback*. In *Wikipedia*. https://en.wikipedia.org/wiki/Reinforcement_learning_from_human_feedback

- Sogolytics. (2023, June 6). *Human-in-the-loop: Maintaining control in an AI-powered world*. Sogolytics Blog. https://www.sogolytics.com/blog/human-in-the-loop-ai/

- Reuters. (2024, January 17). *Extreme weather named biggest global risk as misinformation and disinformation rated most severe*. Reuters. https://www.reuters.com/world/extreme-weather-and-misinformation-risks-2024-01-17/

- Lucy Family Institute. (2024, July 10). *Prebunking messages improve election trust among misinformed individuals*. Lucy Family Institute, University of Notre Dame. https://lucy.nd.edu/news/prebunking-messages-improve-election-trust/

- Illinois Department of Financial and Professional Regulation. (2025, August 4). *Gov Pritzker signs legislation prohibiting AI therapy in Illinois*. News Release. https://idfpr.illinois.gov/news/2025/gov-pritzker-signs-state-leg-prohibiting-ai-therapy-in-il.html

They've arrived—uninvited, unstoppable, and already calling the shots.

Artificial intelligence isn't just another technology; it's a new kind of immigrant—one that doesn't sleep, doesn't eat, and learns faster than we do.

In this provocative debut, Kevin Ryan unpacks the cultural, economic, and moral upheaval unleashed by intelligent machines. Drawing parallels between today's algorithms and yesterday's immigrants, AI Immigrants explores who wins, who loses, and what it really means to stay human in an automated age.

From deep-fake politics to data colonialism, from algorithmic landlords to digital day-labourers, Ryan's narrative nonfiction blends insight, wit, and urgency. It's part social commentary, part survival guide—and a manifesto for keeping humanity indispensable.

Accessible, sharp, and deeply human, this book asks the question no algorithm can answer:

Who owns the future?

Kevin Ryan has spent thirty years making complex technology work in production — from CERN and the Financial Times to Nestlé and NatWest. He's been early to every wave, from XP in the nineties to AI-native engineering now. Based between London and Budapest, which means he's an immigrant too.