Digital Wisdom

Searching for Agency in the Age of AI

Richard Lachman

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

Introduction

> It came without warning. Except for every possible warning.
>
> —*Jonathan Lethem, "The Arrest"*

Don't trust me. Don't trust anyone like me. Here's why:

In a New York Times interview, Sam Altman, the CEO behind ChatGPT, said, "Well, I am a believer that all real sustainable human progress comes from scientific and technological progress."

Maybe that sounds reasonable to you. Maybe you think "I love the things tech has given me—my smartphone, and streaming videos, and, uh, the wheel?"

But Altman's principle is a myopic and deeply troubling one. In one quick stroke, a man with his hand on the tiller of the future of Artificial Intelligence simplifies the innate, messy, chaotic nature of human society to a series of engineering problems. Not only does he discount the importance of all humanities and social sciences, of art, of politics, and of values or relationships, but he also pretends that the workings of our daily life are something we can solve for.

Artificial Intelligence is the latest digital tech driving massive change in seemingly every aspect of civil society, and we are rushing to fold tech solutions into areas we don't really understand. I'm an MIT-trained engineer and a professor, but my life and experiences give me just one perspective on the big questions of life. If I told you that my tech buddies and I were going to be in charge of education for your kids, choices for your healthcare, and your government's policies for housing, taxes, and the economy, you'd tell me to take a hike. You wouldn't be so polite about it. You didn't vote for me, and you have no idea if things I value, or the things I prioritize, are the same things you care about. Why would you let my choices dictate the world you live in?

Mark Zuckerberg famously instructed his teams to "move fast and break things." He meant that his developers should try a lot of ideas, without worrying about preserving existing income-streams or ways of doing things. But that's a behaviour of a certain time in life, both for companies and for humans. Toddlers move fast, and they break things, as part of their process of learning and experimentation, unrestricted by what is sometimes the paralyzing weight of adulthood.

But Facebook and the rest of big-tech should no longer be toddlers: the scale and reach of the platforms mean that when they break something, ethnic groups face forced migration,[1] public health information is lost,[2] and governments tremble.[3]

Google's early mantra was "Don't be evil," which is about as complex a code of ethics as a toddler can manage: knowing that there is such a thing as right and wrong, but with no sophisticated understanding of what that means in a complex world. Don't get me wrong: toddlers can do incredible things. My 2-year-old wandering the living room as I type this learns more in a day than I do in a year. Toddlers have energy, they ask fascinating questions, and they can dream big dreams. That's a powerful force, and it needs to be preserved and protected for all that it can do. But when those dreams turn not just into reality, but a reality which has scale, reach, and impact, they can't be governed by the same rules. They need to grow up. They need the insight, reasoning, care, and concerns of age. We need to focus not just on the digital revolution, but on Digital Wisdom. And these aren't abstract principles to debate in some Chatham House-Rules debating society. They affect us where we live, work, socialize, and sleep.

Let me make that point more concrete: When I sleep, I snore. It's just a fact of ageing, and genetics, and probably the glass of wine I had with dinner. Luckily, where I live, my publicly funded health-insurance means I'm covered for a CPAP machine: a nose-and-mouth covering mask with a corrugated tube that connects me to a small air-pump. I call it my SCOBA-tank ("Self-Contained Overwater Breathing Apparatus"). But, as with many medical interventions, it's a hard adjustment to make: I look ridiculous, I feel ridiculous, and the aforementioned toddler cracks up when he sees me in it. Unsurprisingly, as many as 83% of CPAP owners don't use their machines for the prescribed number of hours each night.[4] In medical terms, this is called "non-adherence" (or, for the more Orwellian among us, "non-compliance").

While it's any person's right to follow or ignore medical advice, you can see why this might be a problem for an insurer: Should they spend the money on an expensive medical device if the client won't use it? And so, my CPAP comes with a snitch and a threat: the snitch is a cellular modem with a tiny micro-SIM chip that sends data automatically to my health-care specialist, without even needing to be on my Wi-Fi. The threat: if I don't use the device at some minimum prescribed level, I have to pay back the cost.

This might seem like a small-scale solution to a localized problem. But it represents something real, and new, and on its way to being pervasive: networks, data, and digital technologies are rapidly changing our relationship to privacy, control, connection, and agency without our having had any kind of broad, basic conversation about what we want the rules of engagement to be. A machine on my bedside table is watching me while I sleep, and reporting what I do to someone who will charge me money if I cross the line. Where is the data stored? Why is it treated as more shareable than the rest of my medical history? Could the records be subpoenaed by an angry driver if my under-slept self got in a car accident? The effects of

AI, social media, and always-on networked devices are not just one of marketing, convenience, or, as tech companies want us to think, an inevitable and unstoppable path to the bright and shiny future. There are important conversations about personal agency and control that we should engage in while we still have some semblance of personal agency and control.

Technology can shape our lives. Think of our bodies, pushed into office chairs, clawed-hands straining around a mouse, wrists and tendons aching as we hunt-and-peck type for hours on end. The growth of video-call platforms like Zoom during the pandemic lead to a change in the way American Sign Language (ASL) is used in the Deaf community, where the size of the video-frame, the cut-off view of the whole body, and the difficulty of picking out fine-grain finger movements led to vernacular alterations in the handforms and positions some Deaf communicators used online.[5] The details of how any particular piece of tech functions can change our bodies, our minds, and our relationships.

When I started my working-life in the late 1990s, I spent a lot of time trying to convince media companies that the internet should become part of their business. Now, in the 2020s, it's difficult to think of any aspect of our life that isn't touched by tech. It's part of how we raise our kids, educate ourselves, get healthy, treat illness, care for our parents, and run our countries. The CEOs of our technology companies are running a series of huge-scale experiments, in real-time, on how we communicate, form community, manage attention, develop a sense of well-being, and make a living. Don't you think we should have a greater say in how it all functions?

The famed urbanist Jane Jacobs, in talking about another great, complex, and messy part of life, wrote "cities have the capability of providing something for everybody, only because, and only when, they are created by everybody."[6] Just like cities, our digital life is chaotic, evolved, rich, and dangerous; it can be your source of employment, addiction, crime, love, arguments, life-lessons, and a fun Friday night. And just as in the cities where we live, we need to insist that we are more than economic consumers: we need to take an active role in understanding, growing, shaping, and claiming relevance. Yet, while it's tempting to critique "big tech," technology isn't like the Rotary Club. It isn't an organization, there are no membership cards, and there isn't a Mission Statement. The engineers, designers, entrepreneurs, and user-interface experts are individuals, working on disparate projects in disparate contexts. If we're in a situation that seems untenable, it's an emergent one, the product of a huge number of individual decisions by makers and users alike. And the path forward is unlikely to be monolithic in nature, but rather something that itself emerges from a sea-change, from something that affects all boats in the water, and becomes second nature.

I want us to take an active role in reshaping technology, and my job is to help you understand more about the underlying issues so we can pick a better path together. This book will explore how fundamental aspects of today's digital technology intersect with our rights and assumptions about private life, individual

autonomy, and social boundaries. Part I explores what all that data collected about us is for, and how companies, governments, and ourselves (as people who also work for those companies and governments) might want to alter the dynamic. We'll look at our changing sense of privacy, and how data is not just recorded but fused, correlated, shared, and acted upon in ways that seem fundamentally different than our laws or principles once assumed. We'll talk about technology's role in making us seem more divided than ever, affecting our politics, our pastimes, and even our basic ability to connect with one another. We'll explore the ways in which technology is used to change our minds, sometimes for our benefit and sometimes to our detriment. Part I closes by breaking down the increasing use of algorithms that impact our lives, and how we can build a more ethical approach to design into the practice.

Part II explores the mindset of Digital Wisdom and makes an argument for how we can reshape our approach to better live in the ways we want, with the features of technological advancement we choose. I'll suggest how our educational systems, industries, and governments could shift to better support Digital Wisdom. These ideas are not exhaustive: I'm proposing Digital Wisdom as a mindset, not a simple set of steps to follow. It's going to take creativity from all of us, in any sector that our personal or professional lives touch upon, to take concrete steps towards a more responsible digital future.

In her book *Alone Together: Why We Expect More from Technology and Less from Each Other*, Sherry Turkle writes: "Because we grew up with the net, we assume that the net is grown-up. We tend to see it as a technology in its maturity. But in fact, we are in the early days. There is time to make the corrections." This book makes the case for why we need to grow up, and what we can do to survive the growing-pains.[7]

Chapter 1

The Lessons of Realtechnik

> The problem with the future is that it keeps turning into the present.
>
> —*Calvin and Hobbes, Bill Watterson, December 30, 1990*[1]

When Diane Diller was trying to get pregnant, she did what so many of us do when we have a problem: she used her smartphone. Ovia, an app with millions of subscribers, helps users track their menstrual cycle, moods, basal body-temperatures, and even when they have intercourse. The company offers another app to track data during a pregnancy, measuring exercise, energy levels, diet, and symptoms. Diller heard about the app from her employer, who gave it to her, for free, as part of their health-and-wellness benefits. In fact, it was better than free: Diller's employer offered her $1/day in gift cards for each day she logged data.

As she told the Washington Post,[2] "Maybe I'm naive, but I thought of it as positive reinforcement: They're trying to help me take care of myself." Cue the sarcastic narrator: "They weren't trying to help her take care of herself." Or, if we're being generous, let's just say it's more complicated than that. Diane's employer, like any of Ovia's corporate customers, could get weekly anonymized updates about the app. They could see which articles were read the most by their employees—say, an article on depression, or miscarriage, or financial concerns, or information for queer parents. They could get weekly updates on the percentage of their employees with high-risk pregnancies, evidence of struggles with infertility, or a predicted timing for returning to work after parental leave. The company offered Diller something to help her, but they also gave themselves a window into her private life, turning her hopes, dreams, and disappointments into data points for the corporate planning process.

DOI: 10.1201/9781003592501-2

New parents and pregnant women are often discriminated against in the workplace, facing fewer opportunities for promotion, contract renewal, or plum assignments.[3] And while apps promise to only provide employers with aggregated information, keeping details private about any individual user, the practice of "de-anonymizing" data is a well-studied one.[4, 5] If you have a mid-sized company, with only a few potentially pregnant users, it's effortless to figure out who the "anonymized" info refers to. And, of course, putting information about your mental state, your habits, or your plans for having a family in the hands of a manager could change the course of your career.

Yet Diane Diller found real benefit in the app. Being able to chart the details of her journey and read suggested articles based on her own data was helpful to her. And she wasn't alone—in one study, over two-thirds of respondents confirmed using apps like Ovia before and during pregnancy,[6] and many users of similar apps reported finding the information and community-connections to be useful.[7]

So here's the question: why is this useful tool also a scary tool? Why is it that the business model of a market leader happens to be the one that endangers our rights, our privacy, and our quality of life? Is there no way for us to gain the benefits of our new technologies while limiting the dangers? Who is responsible for understanding the risks? Is it the user who sees the 6000-word terms-of-service agreement for an app and just clicks "accept"? Is it the junior software engineer assigned to implement a specific feature who happens to do a poor job of it? Consider the case of a competitor's pregnancy-app, which revealed sensitive user-reported information—such as whether they'd had an abortion, or when they last had sex—to anyone who simply knew the user's email address.[8] Or is it the responsibility of third-party data-brokers who, as an investigation by the Vice Motherboard blog revealed, sold information about user-visits to Planned Parenthood locations, including those that offer abortion services, for $160 to anyone with a credit card?[9] Are our priorities, privacy, or the impact of the tiny decisions we make in our digital lives completely out of sync with what's actually happening? How did we get here, and, more importantly, how do we leave?

Digital Promise

In the mid-1990s, at the dawn of the Web, I was a graduate student at the MIT Media Lab, a hotbed of research, invention, and digital dreaming. In the wee hours of any given weeknight, my colleagues and I were in the computer-lab, wired on too much caffeine, with access to far too much computing power. One lab-mate wired up a slot-car track so the cars would only work if you yelled at them. Another wrote code to constantly scan email for the word "pizza," so he could get the jump on after-meeting leftovers dropped off in the kitchen.

But on every one of those nights, we were hungry for more than Jolt Cola and free food. We were passionate about the promise of technology and the emerging

Internet and felt privy to a secret just moments before it exploded into the wider world. And along with small clusters of people in basement computer-labs around the globe, we were certain that the secret was going to change everything for the better. Citizen journalists would watchdog politicians, toppling authoritarian regimes with first-person on-the-ground reporting that circumvented state-controlled media. Lonely queer kids would log on in the middle of the night from a small town in the Midwest, connecting to people with a shared identity, having online conversations that might literally save their lives. Education would focus on the needs of each child individually, with intelligent-agent software charting progress and explaining concepts with infinite patience and boundless resourcefulness. New forms of art and storytelling would revolutionize moribund ways of thinking, ushering in a decentralized and disaggregated future in which corporate power was forever altered, and democracy could live up to some Athenian dream.

When I flash forward to the headlines of today, it's like looking at those ideals in a funhouse mirror. Facebook helps us connect with one another, but it also forms a massive store of photos and information about us, where it was vulnerable to being scraped by both facial-recognition company Clearview AI and political-consulting firm Cambridge Analytica, super charging products sold to law-enforcement officials or political campaigns worldwide. While schools provide students with free computers for work and learning, fears about school shootings have led school boards to purchase surveillance software like Geo Listening and Social Sentinel, which monitors all of the students' online activities and communications, resulting in young people being questioned or suspended for simply complaining about their teachers on social media.[10] And while civic-tech data can lead to better government decisions, those same cities are buying AI-driven software to assess who stays in jail and who goes free, without testing if the advice is any better than a random guess.[11]

Potential vs Actual

In 2020, the Pew Research Center interviewed academics and digerati about whether digital tech was inherently harmful or helpful. The experts made statements like "tech is a tool … how we choose to use these tools, the ethical choices we as human societies make along the way, will define us," or "the benefits or harms are determined by how we humans choose to use tools and technologies."[12] These answers are classic, and they are safe, but they vastly oversimplify what's going on. We can't just look at the potential for harm, as if the digital revolution is still a nascent trend and we're curious about what direction it might head. We have to accept the business reality in which those tools operate today. The "potential" has manifested some very real harms: surveillance, attention economies, addictive scrolling, isolated users, and polarized citizens. Tech platforms—say, social media—can be used for benefit (grassroots organizing; citizen science; creator-led fan relationships) or ill (spreading propaganda and misinformation), so it might

make sense to think of them as neutral. But overwhelmingly, the business models, use cases, and ad-network policies that have developed reward the incendiary and the extreme. Internal documents from court cases and whistleblowers have repeatedly shown that these uses aren't incidental, and they aren't accidental. Leaked documents from Facebook whistleblower Frances Haugen showed that the company knew the use of the platform negatively affects body image, worsens disordered eating, and increases suicidal ideation in teen girls.[13] Unredacted documents internal to TikTok, revealed as part of a suit filed by over a dozen State attorneys-general, show the company knows about the dangers the app presents to children. TikTok's own research shows that "compulsive usage correlates with a slew of negative mental health effects like loss of analytical skills, memory formation, contextual thinking, conversational depth, empathy, and increased anxiety" and "interferes with essential personal responsibilities like sufficient sleep, work/school responsibilities, and connecting with loved ones."[14] That is not a potential. That is a series of negative effects experienced daily by millions of users around the globe. What percentage of tech business models depend on tricking users into doing something that's bad for them? There have always been businesses that do so, but we don't lionize them for their visions of making the world a better place.

In the foreign policy world, setting aside a theoretical or idealistic framing to operate more practically, in a day-to-day reality, is called *realpolitik*. It seems past time we engaged in the same framing—a *realtechnik*—to help us on the road ahead. If a company wants to be trusted with private data from its users, it needs to admit that misuse can happen and plan for reducing harm. If a branch of the government wants to use an AI algorithm to help in decision-making, it needs to be very clear about where the data came from, if it can be shared with any other branch, and how problems and mistakes can be reported and repaired. Anyone who proposes a tech solution should cite past problems and talk about what they'll do to learn from those problems.

It's possible to build a technology that uses personal data-tracking to help women trying to get pregnant. It's also quite possible that such information could leak, could be misused, and could cause harm to those same women. Realtechnik admits that this could happen and doesn't launch the tech without a serious, realistic plan for mitigating those harms. It doesn't pretend risks don't exist, nor does it pretend risk can be completely nullified.

It's possible that law enforcement or municipal authorities could benefit from computer vision data in a set of locations: say, red-light cameras with AI licence-plate scanning to automatically ticket offenders, or security cameras blanketing public spaces with real-time facial recognition. But investigations have found this type of data stored insecurely, on devices with default passwords (think "username: admin, password: 1234567"),[15] where it could be stolen and used to track the location and the movements of anyone on the street. Journalists and watchdog-groups like BigBrotherWatch have uncovered abuse of access, with some law enforcement officials using the databases to look up personal information and

harass women.[16, 17] One UK report tracked 2300 data-breaches in four years from the departments they monitored, with over 800 individuals accessing information for no valid policing reason.[18] Realtechnik suggests officials should be the first in line to state the risks and realistic, transparent, third-party-verified steps they're taking to limit them. How will the data be secured? How long will it be kept? Who can have access to it, how is that access logged/monitored, and who monitors or investigates wrongful uses? And the most important question: shouldn't these safeguards volunteered immediately by any agency that wanted to track data, before they're given the right?

It's easy to get lost in the abstract around these issues, so perhaps I should get back to another concrete example.

A Concrete Example

The Rite Aid pharmacy-chain had a problem. In the long history of commerce, it wasn't a new problem, but it seemed to be a growing one: people steal stuff instead of paying for it. They shoplift. Everything from teens stuffing a box of condoms under their coat to, as headlines and viral videos in the early 2020s showed, groups brazenly running for the door with bags of stolen goods. But Rite Aid was a modern corporation, not some ancient vendor hawking dates at the bazaar, and the team back at corporate decided it was time for a 21st century solution: an artificial intelligence facial-recognition algorithm tied into cameras trained on the front door. The AI would scan everyone who entered and compare their faces to a database of images of suspected shoplifters uploaded by Rite Aid employees from stores nationwide. If the AI found a match, it would alert staff by email or phone to shadow the suspect around the store, ban them from making purchases, or engage the target through direct confrontation.

This is the moment where I should say "there was just one problem…" but I can't. Because there were so, so many problems. A 2023 US Federal Trade Commission settlement with the company details the situation.[19] Facial-Recognition AI can be tricky under the best of situations, when it's trained on a database of well-photographed subjects, and operating under perfect lighting conditions. The pharmacy chain, however, used photos of suspects pulled from its own security-camera footage, or from phone-camera snapshots taken by staff, at a distance, at extreme angles, perhaps of a subject fleeing the scene. There was no transparency or oversight into who ended up in the database, and employees were pushed to add as many people as possible. These snapshots eventually numbered in the tens of thousands, stored in a database indefinitely. Shoppers were never notified that facial recognition was used in stores—in fact, staff were specifically told not to mention the system to the media.

But the real problem is deeper, and much more fundamental: facial-recognition systems used in North America are notoriously, and demonstrably, worse at

identifying faces that are Black, Brown, or Asian. It's not that the software couldn't be made more reliable, but rather, the tools are overwhelmingly trained using white faces because that is the labelled data most readily available for commercial purchase. Tools built in China, India, or Israel are similarly faulty, but for different races and skin-tones. The tech companies didn't have to set out to build racially biased tools; they just didn't explicitly set out to build racially impartial ones.

In the Rite Aid case, the effect was to target people of colour for more suspicion, harassment, and confrontation, in front of their friends and family, during what was supposed to be just a quick trip to pick up toothpaste. These types of mistakes, or "false positives," are a well-documented problem with facial-recognition systems, but Rite Aid staff received no training or warnings about the shortcomings inherent in the system. In fact, the pharmacy chain didn't ask the software makers about how accurate the system would be, and the contract with the vendor specifically "makes no representations or warranties as to the accuracy and reliability." The software was built to produce a "confidence rating" of how sure it was that the face of a shopper matched one in the database, but this number wasn't passed along to Rite Aid employees. Clerks were simply told that the fancy computer code had identified a malefactor and that they needed to leap into action.

This isn't an isolated incident. A pregnant woman, Porcha Woodruff, was getting her two daughters ready for school when six Detroit police officers showed up on her doorstep, charged her with robbery and carjacking, and dragged her away in handcuffs. Again, faulty facial recognition had flagged her photo based on a poor-quality security-camera screen-grab. And once again, the person falsely charged was Black.[20] In fact, an investigation of 23 police departments by the Washington Post found 15 of them, from 12 different states, were using facial-recognition AI to arrest suspects who had no other independent evidence to connect them with the crime. These actions were often in violation of the department's own policies, which require investigators to find additional evidence when generating a lead from AI.[21]

In all of these cases, we need to be careful: we can't just say "the tech didn't work well yet." That misunderstanding of the situation could, if we're not careful, lead to a false corollary: "…so, we just need to wait for the tech to work well." But in each case, the tech was part of a chain of decisions made by people. People who are over-eager to believe that a technology is akin to magic. Or people who need more time and resources to try to address a problem, but are instead given a digital solution that claims to do it better for cheaper. A one-time software purchase seems smarter on a corporate balance sheet than ongoing costs for staff time and attention. Tech isn't deployed in a vacuum, and realtechnik acknowledges that problems with the way people use technology are still problems with technology. Vendors are responsible for designing for the real world, and the dashboards, interfaces, and feature sets they launch need to account for the real-world problems we keep finding.

How to Critique New Tech

When a news story about technology makes me feel uncomfortable, I try to figure out what it is about the story that I should pay attention to. Not just which specific app is in the cross-hairs this time, but rather, what I can learn in general about how that form of tech is meshing with society. Is my problem that the tech isn't working very well at the moment? Or is it something more fundamental I'm worried about?

When Google Maps launched, it wasn't very good. The press shared stories of people being directed to drive forward into bodies of water, or down tiny side-streets. Digg.com reported drivers instructed to perform an endless series of U-turns to get to a location, quite literally acting out a software infinite-loop.[22] But these complaints weren't with the underlying idea of a digital map itself: it was simply that the data and the code behind the tech had to get better. And, with iterations, better maps, error checking, infrastructure redesign, and the gentle application of hundreds of millions of dollars, the usefulness of the product was much improved. Google Maps functions very well nowadays. But questions about the underlying technology remain as relevant now as they did when the tool launched. Every time you use Google Maps, you are gaining information, but you are also giving information. If you find your way to a tool called the Google Maps Timeline (go ahead, do a Google Search; I'll wait), you can see a collection of little red dots on a map showing everywhere you've been when using Google Maps. You can slice-and-dice by date or show the places you visit the most often. It will indicate what Google guesses to be your Home and Work, calculated by factors such as frequency and time-of-day you're present at those GPS locations.

Think about the sensitivity of that information for a moment. With your location history, we can make some guesses about whether you're religious and which church, mosque, or synagogue you go to. We can cross-correlate that guess with, say, the location data of a person-of-interest in a Homeland Security investigation and draw a circumstantial connection between you. We could make assumptions about your financial status (Prada store last month, pawnshop this month, with a flight to Vegas in between?). We can make guesses, correct or incorrect, about your sexuality (parking near a gay club every Friday?), your love-life (regular night-time visits to a non-home-address?), or your politics (were you near the site of a rally?). We could take a guess about your health problems based on how often you go to the clinic, or whether you park near the site of a weekly AA meeting. Every activity listed here is legal and reasonable, but each is also, above all, personal. You might not want them shared with companies, employers, advertisers, governments, political groups, family members, or insurance companies. Within the Google environment, you can turn Location Services off to prevent this from being stored, but most of us don't think to do so—and Google begs us to keep the switch on to make their services more useful and our accounts "more secure."

Every time you use Google Maps, you're also giving out information that helps make Google Maps more accurate. Think about connecting to a new Wi-Fi

network—say, at a cafe you're visiting for the first time. You ask the barista for the connection info; they roll their eyes and point wordlessly to the chalkboard with the network name and password. On your phone, you see a list of all kinds of other networks in the local neighbourhood—the guys in the apartment upstairs with a network called "CallOfDuty383," the house a few-doors-down with a network named "FluffyMcCuteyCat," the Wi-Fi at the real-estate office next door, and the cafe itself. You don't have the password for any of those other networks, so you can't connect to them. But your phone is able to see some public information about those networks: the name, the signal strength, and other unique info like a MAC hardware-address. From your seat at the cafe, that unique information—the fact that you can see four other networks, with these specific names, and these signal-strengths—forms a type of fingerprint for your location. When you load Google Maps, it looks up the fingerprint in a database, combining it with your GPS coordinates to give you a more accurate sense of where you are.

You push your location to Google, which, as we just learnt, writes it down for future use. But if you happen to transmit a network-name that isn't in the database, Google also takes the opportunity to update the record. You're giving it new information, linked to an address or location. And those who set up the network—say, the guys above the cafe—are also helping Google. They are providing a free service to the trillion-dollar company, and so are you. I use Google Maps all of the time, and I don't worry that this behaviour is nefarious—but it's also probably not something most of us have thought about. We imagine pulling information from the net, not pushing it to update semi-permanent database records about ourselves and the world. And we have to hope that this information is only used to help us, isn't incriminating, and is stored in a secure fashion. Maybe you are OK with this. But maybe we collectively want to come up with category-wide principles about how information is shared, how safely it's stored, and how we're informed about what's going on.

Let's say you're OK with sharing data to make Google Maps better. What about with an AI or digital voice-assistant's development team? What if you learnt that audio-recordings of your questions to Siri or Alexa were sometimes shared with teams of people hired to assess how well the software recognized the request, and who end up hearing snippets of "confidential medical information, drug deals, and recordings of couples having sex," as the Guardian uncovered in 2019?[23] What if you learnt that Niantic, in their popular game Pokémon Go, used your in-game photos to train an AI about the physical world, adding over a million photo-scans of the real world every week?[24] What if you ran a company and found your employees were pasting proprietary information like source code into ChatGPT to help make presentations, as happened to Samsung,[25] without knowing if that information could be recorded or otherwise used by other companies?[26] What if you learnt that 94% of US hospitals shared patients' medical conditions, prescriptions, and doctors' appointments with advertisers via third-party web cookies on their sites?[27]

Let's take a step back for a second. Early criticism about Google Maps was about it not being very good yet. Then it got better. None of that changes the important conversations we should have about how it functions in the first place. Here's my short-hand tool for this caveat:

> ***Lachman's Law****: Be wary of critiquing tech that doesn't work yet, when you mean to criticise whether it should work at all.*

Some of the issues we'll look at in this book hinge on a technology that simply needs to get better—maybe with more data, or different data, or more money and code. In some cases, it's actually the reverse: the tech is less troubling *because* it doesn't function well, and if it were operating more effectively, it would *then* become problematic. And still others are based on a more fundamental critique, where the tech, even if it worked exactly as promised, would change basic principles of society for the worse or endanger certain of its members. We'll use Lachman's Law to help us tell the difference.

How Did We Get Here?

Why are tech products like Google Maps or the Ovia pregnancy-tracker the way they are? Understanding the "why" of the details of our tech products is a relatively unexplored field; startup CEOs tend to spend more time figuring out what it should do next. But building without understanding the reason for what we're making is dangerous. Philosopher David Dennett explores how we describe the "reasons" for things being the way they are with two simple questions. In one example, he asks us to consider "What is the reason a planet is round?" and "What is the reason dice are not round?"

For a planet, we can answer quite clearly: "The effect of gravity on an object of a certain size forces it into a spherical shape." Dennett calls this a "how come?" question. Essentially, it is "How come planets are round? Gravity." But when we say "What is the reason dice aren't round?", we don't mean "how come?" anymore. There *is* a how-come—it has to do with the manufacturing process, with moulds and plastics that made it look the way it does. But what we usually mean by our second question is "for what purpose?" The purpose behind dice not being round is that they would keep rolling, instead of settling on a number. We use the question "what is the reason?" for both, but we don't mean the same thing. When we look at why so many of the characteristics of a social media platform or our online behaviours are the way they are, we can exhaustively chart feature releases, and incremental changes, and sub-cultural practices becoming mainstream. These are all "how come" answers. But if we ask "for what purpose?", we hit on a much more uncomfortable truth.

Why is misinformation easier to spread on Facebook than verifiable information? Why is security so hard to manage? Why don't individuals have rights over

their data? If we want to be generous, we could suggest that many details of our digital life aren't the result of considered design, but were emergent, happening after the fact, and sometimes to the surprise of the very people who programmed them. If we're less generous, and if we listen to everyone from Facebook whistleblower Frances Haugen to authors Tim Wu (*The Attention Merchants*) and Shoshana Zuboff (*The Age of Surveillance Capitalism*), the goal was always profit regardless of user safety. Design decisions reduced user-friction or kept people in-app longer, but the motivation was to increase our addictive behaviours, anxious app-refreshes, and outrage-driven consumption.

As we'll see in the chapters ahead, if we don't set out to design systems that value transparency and ethics, we get systems that are obfuscated and unethical. The market dictates it, especially in an era of monopoly power and "regulatory capture," where bodies tasked with managing an industry are outgunned, outfunded, and outmanoeuvred. We can't rely on Meta, TikTok, OpenAI, or other tech giants doing the right thing, empowering users, or prioritizing youth safety, simply because we think they are nice people and it is the right thing to do.

When we think about the manipulative and one-sided nature of privacy-tracking, or the persuasive nature of app design, the answer to "for what purpose?" is "to maximize advertiser-driven profit, perhaps with a cursory regard to public-opinion rather than public safety, but only when we get caught." The tech giants invest massive dollars and developer hours into keeping us onsite and online, releasing scores of features, duplicating the functionality of their competitors, and hyper-targeting their personalization algorithms; but when asked about fake news,[28] or keeping teens safe,[29] or moderation of harmful content,[30] the answer is always "we're working on it." A realtechnik point of view accepts the evidence before us: tech companies do not make products that prioritize our values. If we can imagine tools and apps that do, then that same point of view can help us use laws, public pressure, advertiser influence, personal habits, and new business models to get there. Realtechnik also suggests we can't pretend all harm can be removed (nor should we want to—learning how to navigate complexity is an important part of our digital maturity), but we need to acknowledge the harms in order to help reduce them.

But Wait … There's More (to Worry about)

Paying attention to the tech as it is, instead of what it seems to be, isn't limited to the world of social media. However, just at a time we are reckoning with the implications of social media, we're leaping ahead with the next development, far in advance of any regulator's attempts to learn from the past. Artificial intelligence is driving massive investment by technologists, crystal-ball gazing by corporations, and handwringing from intellectual property lawyers and labour analysts. While aspects of AI change from week to week, Lachman's Law pushes us to find deeper critiques that will remain true at a fundamental level for at least the medium-term future.

While the range of approaches covered by the term artificial intelligence is vast, the leaps and bounds of recent memory concern a subfield called Machine Learning and within that Generative AI (GenAI). Broadly speaking, GenAI applies statistical techniques to create text, images, video, or sound. Large Language Models (LLMs) are a form of GenAI that bring those statistical techniques to bear specifically on text. Unlike other AI approaches, these popular and wildly successful techniques don't apply rules or explicit reasoning to the world; rather, they point massive amounts of computing power at massive amounts of data to come up with their own statistical weightings for possible responses to any prompt.

Imagine wanting to think of a word that starts with the letter "A," assigning probabilities to the chance that the next letter would be "b," or "c," etc., and then rolling dice. You could assign equal probabilities to each letter, but you'd get more readable results if you paid attention to the popularity of letters in English (if that's the language you're operating in). If your first letter is a "Q," the odds are higher that the following letter will be a "u"; not 100% certain, as any Scrabble aficionado will confirm ("Qi," or "Qat" are my favourites), but certainly higher. Now imagine doing the same thing for whole words: if a sentence starts with "True north," you might assign higher probabilities to "…strong and free." LLMs are far, far more complex than this simplistic example; they turn words into "tokens" and analyze vast amounts of text to "learn" how one token might relate to another, adjusting parameters to better predict what comes next. When you chat via the friendly interface, it doesn't understand your question, and it doesn't understand its answer. It is playing off relationships between tokens and probabilities and delivering a credible, statistically most likely response. AI chatbots are actually statbots. Their responses are, of course, amazing and useful and have a huge potential for all sorts of applications we have yet to dream of. Yet they also can make up ideas that only sound plausible, "hallucinating" information presented in the same fluent style as actual facts. And instead of just counting on this getting better over time, of course, Lachman's Law wants us to pay attention to our relationship with AI chatbots regardless of the current state of tech.

Releases of AI machine-learning systems trumpet how many parameters they process, or how much data was ingested, or how they compare to benchmarks. But LLMs don't work on defined problems in a lab; rather, they have to function in a messy world populated, inconveniently, by humans. Humans who formulate questions and then trust, mistrust, or interpret the AI's responses. Humans who bring their own biases, experiences, conceptions, and misconceptions into the experience. Realtechnik says we should pay attention to how we incorporate this technology into our everyday lives.

If an AI tool makes recommendations to a human, say, in healthcare benefits administration,[31] or to decide who should get assistance like food stamps[32] or unemployment insurance,[33] as is happening today, how does the actual process of human oversight work? How many click-through reviews of AI decisions happen before that person's attention drifts? When will that experienced human be

replaced by a cheaper person who doesn't have the benefit of years of experience, implicit knowledge, and on-the-ground insight? How long before the human staffing is reduced to a skeleton crew through budget cutbacks, since the machine seems able to produce a credible result? How often will a human actually question the results of what is described, by the vendor, as a cutting-edge machine-learning tool trained on millions of records? When we talk about the risks of deploying AI, we have to pay attention to the people who are going to be using the machines, not just the tools themselves.

When we talk to someone—a real someone, a live human being—it cues a lifetime of our experience in dealing with people. And when something responds in kind, it's very hard not to react as if the hallmarks of a conversation—taking something in, thinking about it, responding in context—are being followed. Conversational interfaces like Alexa and Siri, or LLMs like ChatGPT, Claude, or Gemini, can trigger deep-seated reflexes in regular users and experts alike, no matter how aware of the limits they are. While researchers and writers have detailed some of the biases that machine-learning systems have based on the data they've trained with,[34] there's an additional risk in these modern systems: the bias that humans bring to the interaction.

In the climax of the 1939 film adaptation of The Wizard of Oz, the little dog Toto tugs away the drapery to reveal that the Great and Powerful Oz is, in fact, a man behind the curtain. AI presents a symmetry to this critique: We have to pay attention to the fact that there is someone behind the curtain from either point of view. Humans make decisions about how AI systems are built, and we also interpret the results. Human biases are at play in how the tools are designed (What data is included in training-sets? What guardrails and safety-systems are built? When are systems ready for use and for what purposes?) and how they are used (Are recommendations cross-checked or rubber stamped? Are tools used appropriately, with transparency and accountability? Are safeguards understood and evaluated?). We need to be just as aware of what we, as users, bring to the table as we are of the complex computer algorithms.

AI literacy involves knowing a bit about how systems work, about vulnerabilities in their logics and in our own, in order to take responsible actions. Using systems when you don't know what a reasonable answer looks like can be OK in some situations (say, brainstorming ideas) or to speed up activities you know how to do (such as an experienced developer using GitHub Co-Pilot for routine programming tasks). If you understand a problem-area, you can split the task into elements that LLMs or other AI tools can help with and make huge leaps in productivity and problem-solving. But when using tools in cases where accountability and transparency are important (say, recommending which candidate to hire) or where your information needs to remain privileged (say, corporate decisions or healthcare monitoring), we need to be much more careful. And, dangerously, our human biases can play havoc with understanding the best, safest, or most reasonable responses.

The massive amounts of data crunched by machine-learning systems bring an incredible weight to its recommendations, which functions as a type of intimidation. We find it hard to question an AI recommendation even if common sense tells us to. At the same time, some studies have found evidence of "AI Aversion,"[35] where we decide to mistrust AI recommendations based on dynamics of human interaction rather than proven or statistical reliability (e.g. "It was wrong once, so I won't listen to it again"). With people placed in the position to purchase AI systems, apply software recommendations, and be the final line between a machine and a human subject, we each need to learn a bit about ourselves as well as the tools we use. And that process, learning and applying more knowledge about ourselves in our relationship to technology, can help us understand what problems of tech we want to solve, and how to approach the process.

Towards Digital Wisdom

Another Pew Foundation survey, this one in 2020, found that while 72% of Americans used social media sites, 64% of them felt the sites had a mostly negative effect on the way things were going in their country.[36] The same group found, when asking Canadians and Americans if they felt the artificial intelligence which underlies so many modern technologies was a good thing or a bad thing, that respondents were almost evenly split.[37] We're engaging daily, even hourly, with things we think are bad for us. And the influx of applications that are going to use data to shape our careers, our ability to get an apartment, our political values, our relationships, and our future is only going to increase.

I'm going to explore, in Part II of this book, the idea I call Digital Wisdom. It centres around applying humility, attention to the wider context, an awareness of limitations, and an openness to listening, to our relationship with technology. For now, just imagine what today's tech could look like through the hopes of that idealistic grad student I used to be. If we're worried that social media can cause harm around body image, disinformation, and isolation, what would a social media platform look like that worked on your behalf? What would an app look like if it tracked your interests not on behalf of advertisers, but to focus your attention, help you in your goals, broaden your outlook, or boost your mood when you're down? How might a social-media feed be different if you wanted to pay attention to social, civic, or pro-social outcomes? Or if you wanted it to maximize a breadth of points-of-view, or verifiability, or potential for impacting your community? What would it look like to have an AI agent that worked entirely in your interests, keeping your data safe, and sharing only what was necessary to get what you wanted from vendors, websites, or institutions? What if you had a right to see what information was being stored about you, in a way that was easy to parse, easy to control, and easy to correct?

Sociologist Ruha Benjamin's book *Imagination: A Manifesto* reminds us that we have to develop and practise the skill of thinking of a world different from the

one around us if we want any chance of achieving it.[38] As you read the critiques of technologies in the pages ahead, treat each as an opportunity to acknowledge that technology is not inevitable and that any design choice could have been a different one. Technological, business, and societal constraints are real, but so is our ability to steer a different path on the road ahead. With technologies like AI moving so rapidly, trying to develop practices about one particular capability or feature set seems futile—we'd continually be out-of-date, a few steps behind in our rules and regulations. What we need is a set of principles we can use to help define and delimit the role of technology in our lives.

I want to be clear about something: I'm still, deep down inside, that idealistic graduate student. I still believe that new technologies can give us huge benefits. I'm excited by the capabilities of machine learning, of data-driven research into healthcare, of personalized education, and of digital arts and culture, and I'm engaged every day helping students and companies turn their ideas, perspectives, and experiences into new marvels of the digital age. But if we want the good from our digital future, we need to actively work on mitigating the bad.

Lachman, R. (2026). DIGITAL WISDOM: Searching for Agency in the Age of Ai. Taylor & Francis.

http://www.digitalwisdom.ca