6

Machine Readable Humanity

What's the Problem?

Margot Hanley, Solon Barocas, and Helen Nissenbaum

Over the past 15 years, Daniel Howe and Helen Nissenbaum, often working with other collaborators, have launched a series of projects that leverage obfuscation to protect people's online privacy. Their first project, TrackMeNot, is a plug-in that runs in the background of browsers, automatically issuing false search queries and thereby polluting search logs, making it more difficult or impossible for search engines to separate people's true queries from noise (TrackMeNot 2024). Howe and Nissenbaum later turned to online behavioural advertising, developing AdNauseam, another browser plug-in that automatically clicks on all ads. It is designed to obfuscate what people are actually interested in by suggesting – via indiscriminate, automatic clicks – that people are interested in *everything* (AdNauseam 2024).

Each of these projects has been accompanied by academic publications describing the teams' experiences developing the tools but also reflecting on the value of and normative justification for obfuscation (Howe and Nissenbaum 2017; Nissenbaum and Howe 2009). Among the many observations that they make in these papers, Howe and Nissenbaum conclude their article on AdNauseam by remarking that "trackers want us to remain *machine-readable* … so that they can exploit our most human endeavors (sharing, learning, searching, socializing) in pursuit of profit" (Howe and Nissenbaum 2017). Online trackers don't just record information but do so in a way that renders humans – their sharing, learning, searching, and socializing online – machine-readable and, as such, computationally accessible. For Howe and Nissenbaum, obfuscation is not only a way to protect people's privacy and to protest the elaborate infrastructure of surveillance that has been put in place to support online behavioural advertising, it is specifically a way to resist being made *machine-readable*.

At the time of the paper's publication, the concept of "machine readability" would have been most familiar to readers interested in technology policy from its

The work benefitted from comments on the early draft presented at the Being Human workshop and from outstanding editorial guidance from Val Steeves and Beate Roessler. We gratefully acknowledge the MacArthur Foundation for supporting this work through a grant to the Digital Life Imitative, Cornell Tech.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

important role in advocacy around open government (Yu and Robinson 2012), where there were growing demands that the government make data publicly available in a format that a computer could easily process. The hope was that the government would stop releasing PDFs of tables of data – from which data had to be manually and laboriously extracted – and instead release the Excel sheets containing the underlying data, which could be processed directly by a computer. "Machine readable" thus became a mantra of an open government movement, in service of the public interest. So why do Howe and Nissenbaum invoke machine-readability, in the context of online behavioural advertising, as a threat rather than a virtue, legitimating the disruptive defiance of obfuscation against an inescapable web of surveillance and classification?

In this paper, we address this question, in two parts, first, by theorizing what it means for humans to be machine-readable and, second, by exposing conditions under which machine-readability is morally problematic – the first, as it were, descriptive, the second, normative. Although the initial encounter with problematic machine-readability was in the context of online behavioural advertising and, as a justification for data obfuscation, our discussion aims for a coherent and useful account of machine-readability that can be decoupled from the practices of online, behavioural advertising. In giving the term greater definitional precision descriptively as well as normatively we seek to contribute to ongoing conversations about being human in the digital age.

6.1 MACHINE READABILITY: TOP DOWN

The more established meaning of "machine readable" in both technical and policy discourse applied to material or digital objects that are rendered comprehensible to a machine through *structured* data, expressed in a standardized format, organized in a systematic manner. Typically, this would amount to characterizing an object in terms of data, in accordance with predefined fields of a database, in order to render them accessible to a computational system. Barcodes stamped on material objects, as mundane as milk cartons, render them machine readable, in this sense. When it comes to electronic objects, there are even different degrees of accessibility; for example, conversion tools can transform photographic images and electronic pdf documents into formats more flexibly accessible to computation according to the requirements of various machines' data structures. Structure and standardization have been important for data processing because many computational operations cannot function over inconsistent types of inputs and cannot parse the different kinds of details in an unstructured record. In the near past, for example, a computer would not have been able to process a government caseworker's narrative account of interviews with persons seeking public benefits. Instead, it would have been coded according to discrete fields and predefined sets of possible answers, enabling government agencies to automate the process of

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

determining eligibility. Applied to *people*, machine readability, in this sense, would mean assigning data representations to them according to the predefined, structured data fields required by a given computational system. The innumerable forms we encounter in daily life, requiring us to list name, age, gender, address, and so forth, are instances of this practice.

If this were all there is to machine readability, it would not seem obviously wrong, nor would it legitimate, let alone valorize, disruptive protest, such as data obfuscation. It does, however, call to mind long-standing concerns over *legibility* in critical writing on computing. Here, the concept of legibility refers to the representation of people, their qualities, and their behaviours as structured data representations, the latter ultimately to be acted upon by computers. Representing people as data is not a passive act; rather data collection can be understood as an ontological project that defines what exists before seeking to measure it. Scholarship has tended to focus on how legibility is achieved by imposing categories on people which, in the very act of representing them in data, is morally objectionable in ways similar to practices of stereotyping and pigeon-holing. In the realm of computation, perhaps most famously, Bowker and Starr point out how information systems work to make people and populations legible to social and technical systems by subjecting them to rigid classificatory schemes, forcing them into often ill-fitting categories (Bowker and Starr 2000). This wave of critical writing emphasizes the crude ways in which information systems turn a messy world into tidy categories and, in so doing, both elides differences that deserve recognition and asserts differences (e.g. in racial categories) where they are unwarranted by facts on the ground (Agre 1994; Bowker and Starr 2000). Scholars like Bowker and Starr were keenly aware of the importance of structured information systems representing humans and human activity for the purposes of subsequent computational analysis.

Critiques such as Bowker and Starr's echo critical views of bureaucracies that go farther back in time, emphasizing the violence that their information practices inflict on human beings by cramming them into pigeonholes, blind to the complexities of social life (Scott 2020). The legacy of much of this work is a deeper appreciation of the indignities of legibility that is achieved through information systems, which have been drawn with top-down, bright lines. Incomplete, biased, inaccurate, and weighted by the vested interests of the people and institutions who wield them (see Agre 1994), these systems are also sources of inconvenience for humans who may have to contort themselves in order to be legible to them. Echoing these critiques, David Auerbach has observed that making oneself machine readable has historically demanded significant compromise, not just significant effort: "Because computers cannot come to us and meet us in our world, we must continue to adjust our world and bring ourselves to them. We will define and regiment our lives, including our social lives and our perceptions of ourselves, in ways that are conducive to what a computer can 'understand.' Their dumbness will become ours" (Auerbach 2012). During this period, the terms in which we could make ourselves

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

legible to computers were frequently so limited that achieving legibility meant sacrificing a more authentic self for the sake of mere recognition.

6.2 MACHINE READABILITY: A NEW DYNAMIC

While the operative concerns of this early work on legibility remain relevant, they do not fully account for the perils of machine readability as manifested in online behavioural advertising and similar practices of the present moment. If we agree that automation via digital systems requires humans to be represented as data constructs and that top-down, inflexible, possibly biased and prejudicial categories undermine human dignity and well-being, we may welcome new forms of data absorption and analytics that utilize increasingly powerful techniques of machine learning and AI. Why? To answer, we consider ways that they radically shift what it means to make people machine readable – the descriptive task – and how this new practice, which we call dynamic machine readability, affects the character of its ethical standing. To articulate this concept of machine readability, we provide a sketch – a caricature rather than a scientifically accurate picture (for which we beg our reader's forbearance) – intended to capture and explain key elements of the new dynamic.

We begin with the bare bones: a human interacting with a machine. Less mysteriously, think of it as any of the myriad of computational systems, physically embodied or otherwise, that we regularly encounter, including websites, apps, digital services, or devices. The human in question may be interacting with the system in one of innumerable ways, for example, filing a college application, signing up for welfare, entering a contest, buying shoes, sending an email, browsing the Web, playing a game, posting images on social media, assembling a music playlist, creating a voice memo, and so on. In the approach we labeled "top down," the machine *reads* the human via data that are generated by the interaction, typically, but not always, provided by the human as input that is already structured by the system. Of course, the structured data through which machines read humans may also be entered by other humans, for example, a clerk or a physician entering data into an online tax or health insurance form, respectively. To make the human legible, the data are recorded in an embedded classification scheme, which typically may also trigger an appropriate form of response – whether directly by the machine (e.g. an email sent) or a human-in-the-loop who responds accordingly (e.g. an Amazon warehouse clerk assembles and dispatches a package).

Keyboard and mouse, the dominant data entry medium of early days, have been joined by others, such as sound, visual images, or direct behavioural monitoring, limited by predefined fields for input data fields. The new dynamic[1] accordingly

[1] We owe a debt of gratitude to Diyi Yang who patiently walked Nissenbaum through this setup. She should not, however, be blamed for mix-ups and errors.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

also involves humans and machines engaging in interaction, either directly or indirectly. Thus, it allows a vastly expanded set of input modalities, beyond the keyboard and mouse of a standard computer setup involving data entry through alpha-numerics and mouse-clicks. Machines may capture and record the spoken voice, facial and other biometrics, a myriad of non-semantic sensory data generated by mobile devices, and streamed behavioural data, through active engagement or passively in the background (e.g. watching TV and movies on a streaming service).

The new dynamic also incorporates machine learning (ML) models, or algorithms, which are key to making sense of this input. Machines capture and record "raw" data inputs while the embedded models transform them into information that the system requires in order to perform its tasks, which may involve inferring facts about the individual, making behavioural predictions, deriving intentions, or surmising preferences, propensities, and even vulnerabilities. (Although we acknowledge readers who prefer the plain language of probabilities, we adopt – admittedly – anthropomorphizing terms, such as infer and surmise, which are more common.) Instead of structuring input in terms of pre-ordained categories, this version of machine reading accepts streams of data, which are structured and interpreted by embedded ML models. For example, from the visual data produced by a finger pad pressing on the glass of a login screen, identity is inferred, and a mobile device allows the owner of the finger to access its system. From the streams of data generated by sensors embedded in mobile devices, ML models infer whether we are running, climbing stairs, or limping (Nissenbaum 2019), or, whether we are at home or at a medical clinic, and whether we are happy or depressed.

A transition to dynamic machine readability of humans means that it is unnecessary to *make* ourselves (and others) legible to computational systems in the formal languages that machines had been programmed to read. The caseworker we mentioned earlier may leapfrog the manual work of slotting as many details as possible into the discrete fields of a database and, instead, record the full narrative account (written or spoken) of meetings with their benefits-seeking clients. Language models, according to proponents, trained to extract relevant substantive details, would be able to parse these narratives, extract relevant information, and even, potentially, automate the decision-making itself.

A critical element of our high-level sketch has not yet been revealed, namely, key parties responsible for the existence of the systems we have been discussing – their builders (designers, software engineers, etc.) and their owners or operators (which may be their creators) or companies and other organizations for whom the systems have been developed. When attributing to ML models a capacity to make sense of a cacophony of structured and unstructured data, specifically, to *read* the humans with whom a system interacts, one must, simultaneously, bring to light the purposes behind the sense-making in turn dictated by the interests and needs of controlling

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

parties, including its developers and owner-operators. To serve the purposes of online advertising (highly simplistically), for example, a model must be able to *read* humans who land on a given webpage as likely (or unlikely) to be interested in a particular ad. Moreover, making sense of humans through browsing histories and demographics, for example, according to the dynamic version of machine readability, does not require the classification of human actors in terms of human comprehensible properties, such as, "is pregnant," or even in terms of marketing constructs, such as "white picket fence" (Nissenbaum 2019). Instead, the concepts derived by ML models may be tuned entirely to their operational success as determined by how likely humans are to demonstrate interest in a particular range of products (however that is determined).

Advertising, obviously, is not the only function that may be served by reading humans in these ways. Although lacking insider access, one may suppose that they could serve other functions just as well, such as, helping private health insurance providers determine whether applicants are desirable customers and, if yes, what premiums they should be charged – not by knowing or inferring a diagnosis of, say, "early stage Parkinson's disease" but by *reading* them as "desirable clients of a particular health plan." Importantly, a model that has been tuned to serve profitable advertisement placement is different from one that has been tuned to the task of assessing the attractiveness of an insurance applicant.[2] It is worth noting that machines reading humans in these functional terms may or may not be followed by a machine automatically executing a decision or an action on its grounds. Instances of the former include online, targeted advertising, and innumerable recommender systems, and, of the latter, human intervention in decisions to interview job applicants on the basis of hiring algorithms, or to award a mortgage on the basis of a credit score, etc. Earlier in this section, when we reported on natural language models that could extract relevant data from a narrative, we ought to have added that relevance itself (or efficacy, for that matter), a relational notion, is always tied to purposes. Generally, how machines read humans only makes sense in relation to the purposes embedded in them by operators and developers. The purposes themselves, of course, are obvious targets of ethical scrutiny.

Implicit in what we have described, thus far, is the dynamic nature of what we have labeled dynamic machine readability. ML models of target attributes, initially derived from large datasets, may continuously be updated on the basis of their performance. Making people legible to functional systems, oriented around specific purposes, is not a static, one-off business. Instead, systems are constantly refined on the basis of feedback from successive rounds of action and outcome. This means that, to be successful, dynamic approaches to reading humans must engage in

[2] We note but are unable to give proper credit to the significant body of published work on proxies.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

continuous cycles of purpose-driven classification and, subsequently, modification based on outcomes. It explains why they are insatiably hungry for data – structured and unstructured – which may be collected unobtrusively as machines may monitor humans simply as they engage with the machines in question, to learn whether a particular advertisement yields a click from a particular individual, or a recommendation yields a match, and so on. Dynamic refinement, according to proponents, may be credited with their astonishing successes but also, we contend, a potential source of unethical practice.

To recap: dynamic machine readability is characterized: by an expansion of data input modalities and data types (structured and unstructured, semantic and non-semantic); by embedded ML models which are tuned to the purposes of machine operators and owners; and by the capacity of these models to be continuously refined in relation to these purposes. Dynamic machine readability releases us from the shackles of a limited lexicon – brittle and ill-fitting categories – and associated ethical issues. In a growing number of cases, the pressing concern is no longer whether we have to submit to a crass set of categories in order to be legible to computational systems; instead, many of the computational systems with which we interact take in massive pools of data of multifarious types and from innumerable sources, presumably, to read us *as we are*. Despite the scale and scope of the data and the power of ML, reading humans through models embedded in machines is constrained by the purposes laid out by machine operators and developers, which these models are designed to operationalize. From a certain perspective, the new dynamic is emancipatory. Yet, even if successful these model-driven machines raise a host of persistent ethical questions, which we reveal through a sequence of cases involving machines reading humans. Inspired by real and familiar systems out in the world whose functionality depends on making humans readable, we identified cases we considered paradigmatic in the types of ethical issues that machine readability raises. It turns out that, although the particular ways that a system embodies machine readability is relevant to its moral standing, moral standing depends on other elements of the larger system in which the human-reading subsystems are embedded.

6.3 THROUGH THE LENS OF PARADIGMATIC CASES: AN ETHICAL PERSPECTIVE ON MACHINE READABILITY

6.3.1 *Interactive Voice Response: Reading Humans through Voice*

Beginning with a familiar and quite basic case, one may recall traditional touch-tone telephone systems, which greet you with a recorded message and a series of button-press options. These systems, which have been the standard in customer service since the 1960s (Fleckenstein 1970; Holt and Palm 2021), require users to navigate a labyrinth of choices by choosing a series of numbers that best represent

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

their need. Generally, a frustrating experience, first, you are offered a limited set of options, none of which seems quite right. You listen with excruciating attention to make the best choice and to avoid having to hang up, call back, and start all over again. Although still unsure, eventually, you press a button for what seems most relevant – you chose "sales" but instantly regret this. Perhaps "technical support" would have been a better fit. Throughout the labyrinth of button pushes, you feel misunderstood.

Over time, touch-tone systems have been replaced by Interactive Voice Response (IVR) systems, enabling callers to interact with voice commands (IBM 1964). Using basic speech recognition, these systems guide you through a series of questions to which you may respond by saying "yes," "no," or even "representative." While the introduction of IVR was designed to ease the pain of interacting with a touch-tone system, they have their own interactional kinks. Along the way you may find that you have to change your pronunciation, your accent, your diction, or the speed of your speech: "kuhs-tow-mur sur-vis". You might add "please" to the end of your request, unsure of the appropriate etiquette, but then find that the unnecessary word confuses the system, which prompts it to begin reciting the list of options anew. While voice commands may, to a degree, have increased usability for the caller – for example, being able to navigate the system hands-free – the set of options was just as limited and the process just as mechanical, namely, choosing the button by saying it instead of pressing it.

We may speculate that companies and other organizations would justify the adoption of automated phone services by citing efficiency and cost-effectiveness. Like many shifts to automation that companies make, however, the question should not be whether they are beneficial for the company, or even beneficial overall, but whether the benefits are spread equally. Particularly for systems requiring callers to punch numbers or laboriously communicate with a rigid and brittle set of input commands, efficiency for companies meant effort (and frustration) for callers, not so much cost savings as cost shifting. If we were to attach ethically charged labels to such practices, we would call this lopsided distribution of costs and benefits unfair; we might even be justified in calling it *exploitative*. A company is exploiting a caller's time and effort to reduce its own.

Progress in natural language processing (NLP) technologies, as noted in Section 6.2, has transformed present-day IVR systems, which now invite you simply to state your request in *your own words*. "Please tell us why you're calling," the system prompts, allowing you to speak as if you were conversing with a human. Previously, where callers had to mold their requests to fit the predefined menu of options, now they may express themselves freely and flexibly. The capacity to extract intentions from an unstructured set of words – even if based on simple keywords – and the shift to a dynamic, ML-based approach has made callers more effectively machine readable. Increasingly sophisticated language models continue to improve the capacity to recognize words and, from them, to generate "meaning" and

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

"intention."[3] These developments have propagated throughout the consumer appliance industry, supporting a host of voice assistants from Siri to Alexa, and beyond.

Allowing for more "natural" – and thus less effortful – interaction with machines fulfills one of the long-standing goals of the field of Human–Computer Interaction (HCI), which aims to make computers more intuitive for humans by making humans more legible to computers (Nielsen and Loranger 2006; Shneiderman 2009). Following one of the early founders, Donald Norman, much work in HCI focuses on making interactions with computers materially and conceptually "seamless," effectively rendering interfaces invisible to the human user (Arnall 2013; Ishii and Ullmer 1997; Norman 2013; Spool 2005). IVR systems, which include NLP models, seem to have achieved seamlessness, sparing customers the exasperating and time-consuming experience of navigating a rigid, imperfect, and incomplete set of options. By enabling callers to express what they seek freely and flexibly, have IVR operators addressed the ethical dimensions of their systems – respecting callers time and even autonomy?

Seamlessness addresses some problems at the same time that it creates others. First, advances in machine readability don't necessarily go hand in hand with changes in the underlying business practices. If the back-end options remain the same, shunting callers into the same buckets as before ("sales," "technical support," etc.), defined by organizational interests, objectives, and operational constraints rather than by customers' granular needs, the ability to communicate in their own words actually misleads us into believing that the system is sensitive to our individual needs. If a supple *interface* is not accompanied by more adaptable options at the back end, the clunky button-pressing more honestly provides callers ways that a business is actually able to meet their needs.

Second, the transition to dynamic machine-interpretable, voice-based systems facilitates a richer exchange in more ways than people have reckoned. How one speaks, intonation, accent, vocabulary, and more communicate much more than the caller's needs and intentions, including approximate age, gender, socio-economic level, race, and other demographic characteristics (Singh 2019; Turow 2021b). Attributes of speech such as the sound of the voice, syntax, and tone have already been used by call centres to infer emotions, sentiments, and personality in real-time (Turow 2021b). With automation there is little to stop these powerful inferences spreading to all voice-mediated exchanges. The ethical issues raised by machines reading humans through the modality of voice clearly include privacy (understood as inappropriate data flow). They also include a disbalancing of power between organizations and callers, unfair treatment of certain clientele on the wrong end of fine-tuned, surreptitiously tailored, and prioritized calls, and an exposure to manipulative practices of consumers identified as susceptible and vulnerable to

[3] We use these terms in quotation marks to avoid a presumption that machines are grasping or interpreting in the ways humans do.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

certain pricing or marketing ploys. Scholars have already warned of the wide-scale deception and manipulation that the "voice-profiling revolution" might enable (Turow 2021a). Ironically, the very advances that ease customers' experiences with IVR systems now place customers at greater risk of exploitation, not by appropriating their time and effort but, instead, by surreptitiously reconfiguring their choices and opportunities.

The history of IVR systems highlights an irony that is not unique to it. Brittle systems of the past may have exploited time and effort but also protected against inappropriate extraction of information and laying out in the open the degree to which a business was invested in customer service. Dynamic, model driven IVR systems facilitate an outwardly smoother experience, while more effectively cloaking a rigid back end. Likewise, embedded NLP algorithms offer powers well beyond those of traditional IVR systems, including the capacity to draw wide-ranging inferences based on voice signal, semantics, and other sensory input. These, as we've indicated, raise familiar ethical problems – privacy invasion, disbalance of power, manipulation, unfair treatment, and exploitation. Each of these deserves far more extensive treatment than we can offer here. Although not a necessary outcome of machine readability, but of features of the voice systems in which they are embedded, machine readability both affords and suggests these extensions; it flips the default.

6.3.2 *Reading Human Bodies: From Facial Recognition to Cancer Detection*

Roger Clark defines *biometrics* as a "general term for measurements of humans designed to be used to identify them or verify that they are who they claim to be" (Clarke 2001). Measurements include biological or physiological features, such as a person's face, fingerprint, DNA, or iris; and behavioural ones, including gait, handwriting, typing speed, and so on. Because these measurements are distinctive to each individual, they are ideal as the basis for identification and for verification of identity (Introna and Nissenbaum 2000). The era of digital technology catapulted biometric identification to new heights as mathematical techniques helped to transform biometric images into computable data templates, and digital networks transported this data to where it was needed. In the case of fingerprints, for example, technical breakthroughs allowed the laborious task of experts making matches to be automated. *Datafied* and automated, fingerprints are one of the most familiar and pervasive biometrics, from quotidian applications, like unlocking our mobile phones, to bureaucratic management of populations, such as criminal registries.

6.3.2.1 Facial Recognition Systems

Automated facial recognition technology has been one of the most aspirational of the biometrics, and also one of the most controversial. Presented by organizations as

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

more convenient and secure than alternatives, facial recognition systems have been deployed for controlling access to residential and commercial buildings, managing employee scheduling in retail stores (Lau 2021), and facilitating contact free payments in elementary school lunch lines (Towey 2021). In 2020, Apple offered FaceID as a replacement for TouchID (its fingerprint-based authentication system) (Apple 2024) and in 2021, the IRS began offering facial recognition as a means of securely registering and filing for taxes (Singletary 2022).

In the United States, under guidance from the National Institute of Standards and Technologies, facial recognition has advanced since at least the early 2000s. *Verification* of identity, achieved by matching a facial template (recorded in a database or on a physical artifact such as a key fob or printed barcode) with an image captured in real time at a point of access (Fortune Business Insights 2022), has advanced more quickly than the identification of a face-in-the-crowd. It has also been less controversial because verification systems require the creation of templates through active enrollment by data subjects, presumably with their consent, whereas creating an identification system, in theory, requires the creation of a complete population database of facial templates, a seemingly insurmountable challenge. Unsurprisingly, in 2020 when news broke that Clearview AI claimed to have produced a reliable facial recognition system, a controversy was sparked. Clearview AI announced partnerships with law enforcement agencies and pitched investors its tool for secure-building access (amongst a suite of other applications) (Harwell 2022). The breakthrough it boasted was a database of templates for over 100,000,000 people, which it achieved by scraping publicly accessible social media accounts. Even though no explicit permission was given by accounts holders, Clearview AI took account access status as an implicit sanction.

Objections to automated facial recognition identification (FRI) run the gamut, with Phil Agre's classic, "Your Face is not a Barcode," an early critical perspective (Smith and Browne 2021; Stark 2019). To simplify the span of worthwhile writing on this topic, we propose two buckets. The first includes the societal problems created by FRI malfunctioning, prominently error and bias. The second includes societal problems associated with FRI when they're performing "correctly" or as intended. The second bucket holds insights for our discussion on machine readability.

The usual strawman rebuttal applies to FRI, too, viz. we always have had humans skulking around keeping people under watch. Automation simply improves the efficiency of these necessary practices. As in other cases, the counter-rebuttal insists that the scale and scope enabled by automation results in qualitative differences. Specifically, FRI systems fundamentally threaten a pillar of liberal democracy, namely, prohibitions against dragnets, against surveillance that chills freedoms in public spaces, and in favor of the presumption of innocence. The application of FRI technologies in public spaces impinges on such freedoms and the very existence of vast datasets of facial templates in the hands of operators exposes ordinary people to the potential of such threats. Particularly when there is not a clear alignment of

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

interests and purposes of individuals with the operators of FRI systems and a significant imbalance of power between them, individual humans are compromised by being machine readable.

6.3.2.2 Biometric: Cancerous Mole

Computer vision has yielded systems that are valuable for the clinical diagnosis of skin conditions. Dermatologists, typically first to assess the likelihood that skin lesions are malignant, look at features, such as outline, dimensions, and color. Computerized visual learning systems, trained on vast numbers of cases, have improved significantly, according to research published in *Nature* in 2017 (Esteva et al. 2017). In this study, researchers trained a machine learning model with a dataset of 129,450 images, each labeled as cancerous or non-cancerous. Prompted to identify additional images as either benign lesions or malignant skin cancers, the model diagnosed skin cancer at a level of accuracy on par with human experts. Without delving into the specifics of this case, generally, it is unwise to swallow such claims uncritically. For the purposes of our argument, let's make a more modest assumption, simply that automated systems for distinguishing between cancerous and non-cancerous skin lesions function with a high enough degree of accuracy to be useful in a clinical setting.

Addressing the same question about automated mole recognition systems that we did about FRI; do they raise similar concerns about machine reading of the human body? We think not. Because machine reading necessarily is probabilistic, it is important to ask whether automation serves efficiency for medical caregivers at a cost to patients' wellbeing. Because systems such as these create images, which are stored on a server for immediate and potentially future uses, there may be privacy issues at stake. Much seems to hinge on the setting of clinical medicine and the decisive question of alignment of purpose. Ideally, the clinical provider acts as the human patients' fiduciary; the aims of dermatology and its tools aligned with those of the humans in their care.

Future directions for such diagnostic tools such as these are still murky. In 2017, there were 235 skin cancer focused dermatology apps available on app stores (Flaten et al. 2018). In 2021, Google announced that it would be piloting its own dermatological assistant as an app, which would sit within Google search. In these settings, questions return that were less prominent in a clinical medical setting. For one, studies have revealed that these applications are far less accurate than those in clinical settings and we presume that, as commercial offerings, they are not subject to the same standards-of-care (Flaten et al. 2018). For another, the app setting is notoriously untrustworthy in its data practices and the line between medical services, which have been tightly controlled, and commercial services, which have not, is unclear. Without tight constraints, there is clear potential for image data input to be utilized in unpredictable ways and for purposes that stray far from health.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

In sum, mole recognition systems offer a version of dynamic machine readability that may earn positive ethical appraisal because its cycles of learning and refinement target accuracy in the interest of individual patients. When these systems are embedded in commercial settings where cycles of learning and refinement may target other interests instead of or even in addition to health outcomes, their ethical standing is less clear.

6.3.2.3 Recommenders Reading Humans: The Case of Netflix

Algorithmically generated, personalized recommendations are ubiquitous online and off. Whereas old-fashioned forms of automation treated people homogeneously,[4] the selling point of advances in digital technologies – according to promoters – is that we no longer need to accept one-sized-fits-all in our interfaces, recommendations, and content. Instead, people can expect experiences catered to us, individually – our tastes, needs, and preferences. Ironically, these effects, though intended to make us feel uniquely appreciated and cared-for, nevertheless, are mass-produced via a cycle of individualized data capture and a dynamic refinement of how respective systems represent each individual. In general terms, it is difficult to tease apart a range of services that may, superficially, seem quite distinct, including, for example, targeted advertising, general web search, Facebook's newsfeed, Twitter feeds, TikTok's "For You Page," and personalized recommender systems such as, Amazon's "You might like," Netflix's "Today's Top Picks for You," and myriad others. There are, however, relevant differences, which we aim to reveal in our brief focus on Netflix.

Launched in 1996 as a DVD-by-mail service, Netflix began employing a personalization strategy early on, introducing a series of increasingly sophisticated rating systems, coupled with recommendation algorithms. In 2000, its first recommendation system called *Cinematch* prompted users to rate movies with a five star rating system (Biddle 2021). The algorithm then recommended movies based on what other users, with similar past ratings, had rated highly.[5] In its efforts to improve the accuracy of these recommendations, Netflix introduced a series of features on their site to capture direct user feedback – to add a star rating to a movie they had watched, to "heart" a movie they wanted to watch, or add films to a queue (Biddle 2021). All of these early features called on users to rate titles explicitly.

Over time, leveraging advances in machine learning and findings from Netflix Prize competitions (Rahman 2020), Netflix shifted to passive data collection practices, gathering behavioral data in the course of normal user–site interaction

[4] In the words of Henry Ford, "You can have any color car you want so long as it's black" (Alizon et al. 2008).

[5] If users a and b both rate movies x and y similarly, and user a also likes movie z, then Cinematch would recommend movie z to user b.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

(e.g., scrolling and clicking), instead of prompting users for explicit ratings. This involved recording massive amounts of customer activity data, including viewing behavioural data (e.g., when users press play, pause, or stop watching a program), viewing data on the programs they watch, at different times of day, viewing search query data, and applying cross-device tracking to collect data about which devices they are using at a given time. Infrequently, Netflix would ask customers for explicit ratings, such as, thumbs up or thumbs down. In addition to passively recording behavioural data, they also conducted A/B tests (approximately 250 A/B tests with 100,000 users each year), for example, to learn which display image performs best for a new movie so it can be applied to landing pages across the platform.

According to public reporting, this dynamic cycle of behavioural data gathering and testing shapes what Netflix recommends and how it is displayed. Factors, such as time-of-day, and a record of shows you have stopped watching midway, further affect recommendations and nudges (Plummer 2017). Algorithms comprising the recommender system shape not only what content is recommended to you, but, further, the design of your Netflix homepages, which (at the time of writing) is composed of rows of titles, each of which contains three layers of personalization; the choice of genre (such as comedy or drama), the subset of genre (such as "Imaginative Time Travel Movies from the 1980s"), and rankings within rows (Netflix 2012).

Without an inside view into Netflix and similar services, we lack direct, detailed insight into how the algorithms work and the complex incentives driving the relevant design choices. Yet, even without it, we're able to interpret elements of different stages of progressive shifts in terms of our analytic framework. To begin, the initial design is analogous to the primitive automated phone answering systems, discussed in Section 6.3.1, where customers were asked to deliberately choose from predetermined, fixed categories. Yet, effortfulness, a factor that raises questions about unfair exploitation, seems less relevant here. Whereas the automated answering services offered efficiency to firms while imposing inefficiencies on callers, in the Netflix case, the effort imposed on viewers, one might argue, results in a payoff to them. The shift to the dynamic form of machine readability relieves users of the effort of making deliberate choices, while, at the same time, yielding a system that is more opaque, less directly under viewers' control, and involves potentially inappropriate data flows and uses, which brings privacy into consideration.[6]

Champions point to the increase from 2% to 80% in the past 20 years in accurately predicting what users choose, as a justification for the use of behavioural approaches over those that rely fully on customers' direct ratings (Biddle 2021). In combination with the scrutiny that these numbers invite, we are, additionally, unconvinced that they are decisive in assessing the ethical standing of these practices. Specifically, no

[6] We have explained elsewhere why Privacy Policies are not satisfactory solutions to these issues (Barocas and Nissenbaum 2014).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

matter how it started out, Netflix, like most other online recommender systems with which we may be familiar, is not *solely* driven by their viewers' preferences and needs.[7] As the market for recommender systems has ballooned in all sectors (Yelp, TripAdvisor, local search services, banking, etc.) and competition for attention has mushroomed, there is pressure to serve not only seekers (customers, viewers, searchers) but also parties wishing to be found, recommended, etc. While managing the sheer magnitude of offerings (e.g. think of how many movies, TV shows, books, consumer items, etc. are desperate for attention) one can imagine the conflicts of interest confronting recommender systems, such as Netflix (Introna and Nissenbaum 2000).

Behavioural data is efficient, and the algorithmic magic created from it, which matches viewers with shows, may not be served by transparency. (Do we really need to know that there were ten other shows we might have enjoyed as much as our "top pick?") We summarize some of these points in the next, and final, section, of the chapter. In the meantime, as a purely anecdotal Postscript: Netflix members may have noticed that there has been a noticeable return to requests for viewers' deliberate ratings of content.

6.4 PULLING THREADS TOGETHER

Machine-readable humanity is an evocative idea, whose initial impact may be to stir alarm, possibly even repulsion or indignation. Beyond these initial reactions, however, does it support consistent moral appraisal in one direction or another? "It depends" may be an unsurprising answer but it begs further explanation on at least two fronts: one, an elaboration of machine-readability to make it analytically useful, and another, an exploration of the conditions under which machine-readability is morally problematic (and when it is not). Addressing the first, we found it useful to draw a rough line between two relevant developmental phases of digital technologies, to which we attributed distinct but overlapping sets of moral problems, respectively. One, often associated with critical discussions of the late twentieth century, stems from the need to represent humans (and other material objects) in terms of top-down, predefined categories, in order to place them in databases, in turn making them amenable to the computational systems of the day. As discussed in Section 6.1, significant ethical critiques honed on the dehumanizing effects of forcing humans into rigid categories, which, as with any form of stereotyping and pigeon-holing, may mean that similar people are treated differently, and different people are lumped together without regard for significant differences. In some circumstances, it could be argued that well designed classification schemes serve positive values, such as efficient functioning, security, and fair treatment but it's not difficult to see how the classification of humans into preordained categories could often lead to bias

[7] We have no insider view to the company's internal practices.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

(or unfair discrimination), privacy violations, authoritarian oversight, and prejudice. In short, an array of harms may be tied, specifically, to making humans readable to machines by formatting them, as it were, in terms of information cognizable by computational systems.

Machine readability took on a different character, which we signaled with the term *dynamic*, in the wake of the successive advances of data and predictive analytics ("big data"), machine learning, deep learning, and AI. Although it addresses problems of "lumping people together" associated with top-down readability, ironically, its distinctive power to mass produce individualized readings of humanity introduces a new set of ethical considerations. The list we offer here, by no means exhaustive, came to us through the cases we analyzed seen through the lens of characteristic elements of a dynamic setup.

To begin, the broadening of data input modalities, about which promoters of deep learning are quick to boast, highlights two directions of questioning. One challenges whether all the data is relevant for the legitimate purposes that the model is claimed to serve (e.g. increasing the speed with which an IVR addresses a caller's needs) or whether it is not (e.g., learning characteristics of callers that lead to unfair discrimination or violations of privacy) (Nissenbaum 2009; Noble 2018).

A second direction slices a different path through the issue of data modalities – in this instance, not about categories of data, such as race, gender, and so on, but about different streams feeding into the data pool. Of particular interest is the engagement of the human data subject (for lack of a better term), which is evident in personalized recommender systems. The Netflix case drew our attention because, over the years, it has altered course in how it engages subscribers in its recommender algorithm, a pendulum swinging from full engagement as choosers to no engagement to, at the present time, presumably somewhere in between. Similarly, in our IVR case, we noted that the powerful language processing algorithms that are able to grasp the meaning of spoken language and *read* and serve human-expressed intention are able to extract other features as well – unintended or against our will. Finally, in the context of behavioural advertising, paradigmatic of the dominant business model of the past three decades, the modality of recorded behaviour absent any input from expressed preference has prevailed in the *reading* of humanity by respective machines (Tae and Whang 2021; Zanger-Tishler et al. 2024).

In order to defend the legitimacy of including different modalities of input into the datasets from which models are extracted, an analysis would require a consideration of each of these streams – deliberate, expressed preference, behavioural, demographic, biometric, etc. – in relation to each of the cases, respectively. Although doing so, here, is outside the scope of this chapter, it is increasingly urgent to establish such practices as new ways to read humans are being invented, for example, in the growing field of so-called digital biomarkers (Adler et al. 2022; Coravos et al. 2019; Daniore et al. 2024), which are claimed to be able to make our mental and emotional states legible through highly complex profiles of sensory data from mobile devices

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

(Harari and Gosling 2023). Another wave of access involves advanced technologies of brain–machine connection, which claims yet another novel modality for reading humans – our behaviours, thoughts, and intentions – through patterns of neurological activity (Duan et al 2023; Farahany 2023; Tang et al. 2023).

In enumerating ethical considerations, such as privacy, bias, and political freedom, we have skirted around, but not fully and directly confronted the assault on human autonomy, which ultimately may be the deepest, most distinctive issue for machine readable humanity. Acknowledging that the concept of autonomy is enormously rich and contested, we humbly advance its use, here, as roughly akin to self-determination inspired by the Kantian exhortation introduced in most undergraduate ethics courses,[8] to "act in such a way that you treat humanity whether in your own person or in the person of any other, never merely as a means to an end, but always at the same time as an end" (Kant 1993, vii; MacKenzie and Stoljar 2000; Roessler 2021). When defending the automation of a given function, system, or institution, defenders cite efficiency, defined colloquially, as producing a desirable outcome with the least waste, expense, effort, or expenditure of resources. Our case of automated phone systems illustrated the point that efficiency for machine owners may produce less desirable outcomes for callers, more wasted time and expenditure of effort. In this relatively unsophisticated case, one may interpret the exploitation of callers as an assault on autonomy.

The expansive class of systems claiming to personalize or customize service (recommendations, information, etc.) illustrates a different assault on autonomy. Among the characteristic elements comprising dynamic machine readability is the dynamic revision of a model in relation to the goals or purposes for which a system was created. The general class of recommender systems[9] largely reflect a two-sided marketplace because it serves two interested parties (possibly three-sided, if one includes the recommender system itself as an interested party.) The operators of personalized services imply that their systems are tailored to the individual's interests, preferences, and choices but their performance, in fact, may be optimized for purposes of parties – commercial, political, etc. – seeking to be found or recommended. Purposes matter in other cases, too, specifically distinguishing between facial recognition systems, serving purposes of political repression of machine-readable humans, and mole identification systems, whose primary or sole criterion of success is an accurate medical diagnosis.

6.5 CONCLUSION: HUMAN BEINGS AS STANDING RESERVE

Martin Heidegger's "The Question Concerning Technology" introduces the idea of standing reserve, "Everywhere everything is ordered to stand by, to be

[8] Taught in the so-called Western tradition.
[9] Including, for example, Web search.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

immediately at hand, indeed to stand there just so that it may be on call for a further ordering." According to Heidegger, the essential character of modern technology is to treat nature (including humanity) as standing reserve, "If man is challenged, ordered, to do this, then does not man himself belong even more originally than nature within the standing-reserve?" (Heidegger 1977, 17). Without defending Heidegger's broad claim about the nature of technology, the conception of machine-readability that we have developed here triggers an association with standing-reserve, that is to say machine-readability as the transformation of humanity into standing reserve. Particularly evident in dynamic systems, humans are represented in machines as data in order to be readily accessible to the purposes of the controllers (owners, designers, engineers) that are embodied in the machine through the design of the model. The purposes in question may have been selected by machine owners with no consideration for ends or purposes of the humans being read. It is not impossible that goals and values of these humans (and of surrounding societies) *are* taken into consideration, for example, in the case of machines reading skin lesions; the extent this is so is a critical factor for a moral appraisal. Seen in the light of these arguments, AdNauseam is not merely a form of protest against behavioural profiling by the online advertising establishment. More pointedly, it constitutes resistance to the inexorable transformation of humanity into a standing reserve – humans on standby, to be immediately at hand for consumption by digital machines.

REFERENCES

Adler, Daniel A., Fei Wang, David C. Mohr, Deborah Estrin, Cecilia Livesey, and Tanzeem Choudhury. "A Call for Open Data to Develop Mental Health Digital Biomarkers." *BJPsych Open* 8, no. 2 (2022): e58. https://doi.org/10.1192/bjo.2022.28.

AdNauseam. 2024. Adnauseam.io.

Agre, Philip E. "Surveillance and Capture: Two Models of Privacy." *The Information Society* 10, no. 2 (1994): 101–127.

Alizon, Fabrice, Steven B. Shooter, and Timothy W. Simpson. "Henry Ford and the Model T: Lessons for Product Platforming and Mass Customization." *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* 43291 (2008): 59–66.

Apple. "About Face ID Advanced Technology," 2024. https://support.apple.com/en-us/HT208108.

Arnall, Timo. "Exploring 'Immaterials': Mediating Design's Invisible Materials." *International Journal of Design* 8, no. 2 (2013): 101–117.

Auerbach, David. "The Stupidity of Computers." *N+1, Machine Politics* no. 13 (2012). www.nplusonemag.com/issue-13/essays/stupidity-of-computers/.

Barocas, Solon, and Helen Nissenbaum. "Big Data's End Run around Anonymity and Consent." *Privacy, Big Data, and the Public Good: Frameworks for Engagement* 1 (2014): 44–75.

Biddle, Gibson. "A Brief History of Netflix Personalization." *Medium*, June 1, 2021. https://gibsonbiddle.medium.com/a-brief-history-of-netflix-personalization-1f2debf010a1.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

Bowker, Geoffrey, and Susan Leigh Star. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press, 2000.

Clarke, Roger. "Biometrics and Privacy." 2001. www.rogerclarke.com/DV/biometrics.html.

Coravos, Andrea, Sean Khozin, and Kenneth D. Mandl. "Author Correction: Developing and Adopting Safe and Effective Digital Biomarkers to Improve Patient Outcomes." *NPJ Digital Medicine* 2 (2019): 1–5. https://doi.org/10.1038/s41746-019-0090-4.

Daniore, Paola, Vasileios Nittas, Christina Haag, Jürgen Bernard, Roman Gonzenbach, and Viktor von Wyl. "From Wearable Sensor Data to Digital Biomarker Development: Ten Lessons Learned and a Framework Proposal." *npj Digital Medicine* 7, no. 1 (2024): 161.

Duan, Yiqun, Jinzhao Zhou, Zhen Wang, Yu-Kai Wang, and Chin-Teng Lin. "Dewave: Discrete EEG Waves Encoding for Brain Dynamics to Text Translation." *arXiv preprint arXiv:2309.14030* (2023).

Esteva, Andre, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and Sebastian Thrun. "Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks." *Nature* 542, no. 7639 (February 2, 2017): 115–118. https://doi.org/10.1038/nature21056.

Farahany, Nita A. *The Battle for Your Brain: Defending the Right to Think Freely in the Age of Neurotechnology*. St. Martin's Press, 2023.

Flaten, Hania K., Chelsea St Claire, Emma Schlager, Cory A. Dunnick, and Robert P. Dellavalle. "Growth of Mobile Applications in Dermatology: 2017 Update." *Dermatology Online Journal* 24, no. 2 (2018). https://doi.org/10.5070/D3242038180.

Fleckenstein, W. O. "Development of the Touch-Tone® Telephone." *Research Management*, vol. 13, no. 1 (1970): 13–25.

Fortune Business Insights. "Facial Recognition Market Rising at a CAGR of 14.8% to Reach USD 12.92 Billion by 2027." 2022. www.globenewswire.com/news-release/2022/02/08/2380458/0/en/Facial-Recognition-Market-Rising-at-a-CAGR-of-14-8-to-Reach-USD-12-92-Billion-by-2027.html.

Harari, Gabriella M., and Gosling, Samuel D. "Understanding Behaviours in Context Using Mobile Sensing." *Nature Reviews Psychology* 2, no. 12 (2023): 767–779. https://doi.org/10.1038/s44159-023-00235-3.

Harwell, Drew. "Facial Recognition Firm Clearview AI Tells Investors It's Seeking Massive Expansion beyond Law Enforcement." *The Washington Post*, February 16, 2022. www.washingtonpost.com/technology/2022/02/16/clearview-expansion-facial-recognition/.

Heidegger, Martin. *The Question Concerning Technology*, translated by William Lovitt. New York: Harper & Row, 1977. Originally published in German as *Die Frage nach der Technik*, 1954.

Holt, Jennifer, and Michael Palm. "More Than a Number: The Telephone and the History of Digital Identification." *European Journal of Cultural Studies* 24, no. 4 (2021): 916–934.

Howe, Daniel, and Helen Nissenbaum. "Engineering Privacy and Protest: A Case Study of AdNauseam." *International Workshop on Privacy Engineering*, San Jose, CA, May 25, 2017.

IBM. "IBM Product Announcement 7770." (1964). https://ed-thelen.org/comp-hist/IBM-ProdAnn/7770.pdf.

Introna, Lucas D., and Helen Nissenbaum. "Shaping the Web: Why the Politics of Search Engines Matters." *The Information Society* 16, no. 3 (2000): 169–185.

Ishii, Hiroshi, and Brygg Ullmer. "Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms." *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, Atlanta, GA, March 22–27, 1997.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

Kant, Immanuel. *Grounding for the Metaphysics of Morals*, translated by James W. Ellington. Cambridge, MA: Hackett Publishing, 1993.
Lau, Pin Lean. "Facial Recognition in Schools: Here Are the Risks to Children." *The Conversation*, October 27, 2021. https://theconversation.com/facial-recognition-in-schools-here-are-the-risks-to-children-170341.
Mackenzie, Catriona, and Natalie Stoljar, eds. *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*. New York: Oxford University Press, 2000.
"Netflix Recommendations: Beyond the 5 Stars (Part 1)." Medium, April 6, 2012. https://netflixtechblog.com/netflix-recommendations-beyond-the-5-stars-part-1-55838468f429.
Nielsen, Jakob, and Hoa Loranger. *Prioritizing Web Usability*. London: Pearson Education, 2006.
Nissenbaum, Helen. "Contextual Integrity Up and Down the Data Food Chain." *Theoretical Inquiries in Law* 20, no. 1 (2019): 221–256.
Privacy in Context: Technology, Policy, and the Integrity of Social Life. Redwood City, CA: Stanford University Press, 2009.
Nissenbaum, Helen, and Daniel Howe "Trackmenot: Resisting Surveillance in Web Search." In *Lessons from the Identity Trail: Anonymity, Privacy, and Identity in a Networked Society*, edited by I. Kerr, C. Lucock, and V. Steeves, 417–440. Oxford: Oxford University Press, 2009.
Noble, Safiya Umoja. "Algorithms of Oppression: How Search Engines Reinforce Racism." In *Algorithms of Oppression*. New York: New York University Press, 2018.
Norman, Donald. *The Design of Everyday Things: Revised and Expanded Edition*. New York: Basic Books, 2013.
Plummer, Libby. "This Is How Netflix's Top Secret Recommendation System Works." *Wired*, August 22, 2017. www.wired.co.uk/article/how-do-netflixs-algorithms-work-machine-learning-helps-to-predict-what-viewers-will-like.
Rahman, Was. "The Netflix Prize: How Even AI Leaders Can Trip Up." *Medium*, January 11, 2020. https://towardsdatascience.com/the-netflix-prize-how-even-ai-leaders-can-trip-up-5c1f38e95c9f.
Roessler, Beate. "*Autonomy: An Essay on the Life Well-Lived*." Hoboken, NJ: John Wiley & Sons, 2021.
Scott, James C. *Seeing like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven, CT: Yale University Press, 2020.
Shneiderman, Ben. "Creativity Support Tools: A Grand Challenge for HCI Researchers." In *Engineering the User Interface*, edited by M. Redondo et al., 1–9. London: Springer, 2009.
Singh, Rita. *Profiling Humans from Their Voice*. London: Springer, 2019.
Singletary, Michelle. "Despite Privacy Concerns, ID.me Nearly Doubled the Number of People Able to Create an IRS Account." *The Washington Post*, February 25, 2022. www.washingtonpost.com/business/2022/02/25/irs-idme-account-success-rate/.
Smith, Brad, and Carol Ann Browne. *Tools and Weapons: The Promise and the Peril of the Digital Age*. New York: Penguin, 2021.
Spool, Jared M. "What Makes a Design Seem 'Intuitive'?" UX Articles by Center Centre (blog), January 10, 2005. https://articles.centercentre.com/design_intuitive/.
Stark, Luke. "Facial Recognition Is the Plutonium of AI." *XRDS: Crossroads, The ACM Magazine for Students* 25, no. 3 (2019): 50–55.
Tae, Ki Hyun, and Steven Euijong Whang. "Slice Tuner: A Selective Data Acquisition Framework for Accurate and Fair Machine Learning Models." *Proceedings of the*

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008

2021 International Conference on Management of Data, Virtual Event China, June 20–25, 2021, 1771–1783.

Tang, Jerry, Amanda LeBel, Shailee Jain, and Alexander G. Huth. "Semantic Reconstruction of Continuous Language from Non-invasive Brain Recordings." *Nature Neuroscience* 26, no. 5 (2023): 858–866.

Towey, Hannah. "The Retail Stores You Probably Shop at that Use Facial-Recognition Technology." *Business Insider*, July 19, 2021. www.businessinsider.com/retail-stores-that-use-facial-recognition-technology-macys-2021-7.

TrackMeNot. 2024. TrackMeNot.com.

Turow, Joseph. "Shhhh, They're Listening – Inside the Coming Voice-Profiling Revolution." *The Conversation*, April 28, 2021a. http://theconversation.com/shhhh-theyre-listening-inside-the-coming-voice-profiling-revolution-158921.

The Voice Catchers: How Marketers Listen in to Exploit Your Feelings, Your Privacy, and Your Wallet. Yale University Press, 2021b.

Yu, Harlan, and David Robinson. "The New Ambiguity of 'Open Government'." *UCLA Law Review Discourse* 59 (2012): 180–208.

Zanger-Tishler, Michael, Julian Nyarko, and Sharad Goel. "Risk Scores, Label Bias, and Everything but the Kitchen Sink." *Science Advances* 10, no. 13 (2024): eadi8411. https://doi.org/10.1126/sciadv.adi8411.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 12 Nov 2026 at 19:30:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009383851.008