Peeping HALs: Making Sense of Artificial Intelligence and Privacy

Ryan Calo

The field of artificial intelligence, broadly defined as the study and practice of designing intelligent agents,[1] is at least six decades into its existence as a formal discipline.[2] Sometimes called "computational" or "synthetic" intelligence, AI borrows from and informs a wide variety of subjects, including philosophy, psychology, linguistics, neuroscience, statistics, economics, and law.[3] Techniques of AI underpin all manner of industrial and consumer applications—from the complex neural nets used in data mining, down to the 'fuzzy logic' used by commercial washers and driers.

Insofar as "the issues of AI are directly related to [the] self-image of human beings", [4] and because the central projects and techniques of AI can often be articulated in lay language, few shy away from offering their insights and critiques.[5] This essay explores and updates a particular criticism—the long-standing claim that certain techniques and applications of AI violate human privacy—and discusses whether (U.S.) privacy law is adequately positioned to respond.

Historically, AI can be said to threaten privacy according to a specific pattern: AI substitutes for humans at various stages of observation or surveillance, allowing such activity to reach a previously impracticable scale. Whereas once telephonic surveillance required one listener per phone call, the development of voice recognition technology permits the substitution of a computer capable of monitoring thousands of calls simultaneously.[6] Whereas once hundreds of intelligence analysts might be required to pour over field records in search of connections, AI knowledge management techniques automatically spot patterns

[1] **S. RUSSELL and P. NOVIG**, *Artificial Intelligence: A Modern Approach*, Saddle River, Pearson Education, Inc., 2003, pp. 1-2.

[2] The term "artificial intelligence" was coined by John McCarthy at the Dartmouth Conference in 1956. **P. MCCURDOCK**, *Machines Who Think*, Natick, AK Peters, Ltd., 2004, p. 529.

[3] *Id.*, pp. 5-16.

[4] **H.R. EKBIA**, *Artificial Dreams: The Quest for Non-Biological Intelligence*, New York, Cambridge University Press, New York, 2008, pp. 30-31. The author of this essay is no exception.

[5] *See, e.g.*, **MCCURDOCK**, *Machines Who Think*, p. 406 (discussing Edward Fredkin's concern that AI will hyper-concentrate power in the hands of one country or institution); **B. JOY**, "Why the Future Doesn't Need Us," Wired Magazine, Apr. 2000 (worrying aloud that AI will turn hostile toward humanity).

[6] **J. WEIZENBAUM**, *Computer Power and Human Reason: From Calculation to Judgment*, San Francisco, W.H. Freeman and Company, 1976, p. 272.

and call them to the attention of agents.[7] These developments vastly amplifying the potential for data gathering and analysis, and hence underpin ubiquitous surveillance.[8]

Such advances in technology have played a key role in driving privacy law. The seminal 1890 article by Samuel Warren and Louis Brandeis, wherein the authors in may ways introduced a right to privacy protected by four causes of action, begins with a concern over "[r]ecent inventions and business methods" such as "instantaneous photography" that make possible previously unheard of behavior.[9] In the Fourth Amendment context, U.S. courts grapple with whether a given new technology permits humans to witness behavior in which, or occurring where, the individual has a reasonable expectation of privacy.[10] A state or federal regulator identify new information gathering practices—for instance, tracking online behavior for ad targeting—and hold it up against established fair information practice principles such as notice and control.[11]

A recent trend in AI complicates this dynamic. Increasingly, practitioners of AI and its subfield of robotics and human computer interaction are imbuing machines with 'social' characteristics.[12] These robots and computer interfaces contain key anthropomorphic signifiers such as eyes, expressive faces and gestures, and natural language capabilities designed to improve machine-human interaction. This set of techniques builds on extensive research suggesting that humans are exquisitely attuned to their own species, and that we react to computers, robots, and other social media as though it were actually human—including through the subconscious alteration of our attitudes and behavior. Its effect is to introduce apparent agents into a variety of new contexts.[13]

Some applications of social AI follow the original pattern of amplifying human capacities. For instance, commercial and governmental entities have begun to leverage social

[7] **T. ZARSKY**, "Mine your Own Business!: Making the Case for the Implications of the Data Mining of Personal Information in The Forum of Public Opinion," 5 Yale J. of L. & Tech. 4 (2004), p. 4.

[8] *Id.*

[9] **S. WARREN AND L. BRANDEIS**, "The Right to Privacy," 4 Harv. L. Rev. 193 (1890).

[10] *See, e.g, Kyllo v. United States*, 533 U.S. 27, 29 (2001) (discussing thermo-imaging devices).

[11] FTC Press Release, "FTC Staff Proposes Online Behavioral Advertising Privacy Principles," Dec. 20, 2007, available online at http://www.ftc.gov/opa/2007/12/principles.shtm.

[12] **T.M. HOLGRAVES ET AL**, *Perceiving artificial social agents,* Computer in Human Behavior 23 (2007) 2163 ("One of the major trends in human-computer interaction has been the development of more natural human-computer interfaces"); *id.* at 2171 ("There is no doubt that this trend will continue to increase.").

[13] **P.J. FOGG**, *Persuasive Technologies: Using Computers to Change What We Think and Do,* San Francisco, Morgan Kaufmann Publishers, 2003, p. 10 ("With the growth of embedded computers, computing applications are becoming commonplace in locations where human persuaders would not be welcome, such as bathrooms and bedrooms, or where humans cannot go (inside clothing, embedded in automotive systems, or implanted in a toothbrush)."). *See also* **H.R. EKBIA**, *Artificial Dreams*, p. 8 (discussing fact that "[c]omputers are everywhere."); **J. KANG and D. CUFF**, "Pervasive Computing: Embedding the Public Sphere," 62 Wash. & Lee L. Rev. 93 (2005), p. 94 ("[T]he Internet will soon invade real space as networked computing elements become embedded into physical objects and environments.").

machines and software to stand in for marketers, recruiters, and other organization representatives in gathering information and influencing consumers and citizens. Just a single AI program—for instance, the U.S. Army's virtual recruiter "SGT Star"—can engage with thousands of individuals simultaneously and record every interaction. Thus, social machines can stand in for a human as interviewer or interrogator, preserving the persuasive skills of humans but adding additional advantages such as massive scale, tirelessness, and an essentially limitless memory.[14]

Other consequences of social AI, however, fall outside the prevailing pattern. Rather than standing in for some specific human task such as listening, pattern-spotting, or questioning, a machine that presents as an independent agent can stand in for a human *as subject*. That is, a humanoid robot's mere presence will signal to individuals that they are not alone, even if the robot is neither collecting nor evaluating information on behalf of a human, which in turn can sharply alter attitudes and behavior. Relatedly, where a computer interface engages with a user socially—a direction the global leader in Internet search expressly contemplates within the next few years[15]—extensive social science research shows that the individual will feel and react as though she is engaging with an actual person.

Although this set of applications does not augment the human capacity to observe in the tradition sense, it nevertheless threatens core privacy values. As detailed below, a key role of privacy is to preserve solitude, in the sense of a temporary respite from interference with respect to curiosity, development, and thought.[16] Yet it is clear that people react to social machines as though they were human beings, including with respect to the sense of being observed. If, as many predict, social machines become ubiquitous—turning up in cars, bedrooms, bathrooms,[17] even within cell phones and mirrors[18]—possibilities for solitude may shrink intolerably. We may even witness a sea change of attitudes away from the prevailing view of computers as passive data conduits, in turn chilling curiosity at the borderline and creating discomfort around widespread machine custodianship of personal information.

It is exactly here, where AI begins to substitute for the human not as a gatherer or organizer of information but directly as subject, that privacy protections begin to break down.

[14] **P.J. FOGG**, *Persuasive Technologies, o.c.*

[15] *See* http://www.techcrunch.com/2008/12/10/marissa-mayer-at-le-web-the-almost-complete-interview/ (interview with Marissa Mayer).

[16] **A. WESTIN**, *Privacy and Freedom*, New York, Antheum 1970, p. 35.

[17] **P.J. FOGG**, *Persuasive Technologies*, p. 10.

[18] **P.J. O'RORKE**, "Future Shlock," The Atlantic, Dec. 2008 ("Various passages had caught my attention when I'd read it, and raised my blood pressure: 'Closets will help pick out the right dress for a party.' Imagine that: a talking mirror telling you, 'That makes your butt look big.'").

This is partly because American privacy law and theory focuses on the flow of information, on quantifiable harms, and on the level of notice and consent. It is also a function of limited imagination around the role of objects in our social lives. This essay concludes with thoughts on where to look in American law for legal analogues to vindicate the core privacy values threatened by social AI.

I. Traditional intersection of AI and privacy

On the traditional view, technology threatens privacy by increasing the power or reach of human observation. As prominent American privacy scholar Michael Froomkin sums up the space: "Privacy-destroying technologies can be divided into two categories: those that facilitate the acquisition of raw data and those that allow one to process and collate that data in interesting ways".[19] Speaking on the subject of privacy invasive technologies, Harvard Law School's Jonathan Zittrain identifies "three successive shifts in technology from the early 1970s: cheap processors, cheap networks, and cheap sensors".[20] He continues that "[t]he third shift has, with the help of the first two, opened the doors to new and formidable privacy invasions".[21] The thought is that humans will use cheaper and better technology to collect and organise information to greater effect, sometimes necessitating additional protections.

An early example of this analysis in the context of AI is synthetic intelligence pioneer Joseph Weizenbaum's concern over the use of AI in data mining.[22] In 1976, Weizenbaum wrote a scathing critique of artificial intelligence along multiple lines. Weizenbaum had developed a program called ELIZA that was designed to mimic psychoanalysis by engaging in a credible dialogue with a human operator, in keeping with the "Rogerian technique of encouraging a patient to keep talking".[23] ELIZA asked its users questions based on their previous answer and, where it did not have a response, merely supplied filler such as "I see" or "interesting". Weizenbaum claimed that he was profoundly disturbed by the tendency of

[19] **M. FROOMKIN**, "The Death of Privacy?," 52 Stan. L. Rev. 1461, 1468 (2000). *But see id.*, pp. 1469-70 (acknowledging that "[f]or some, just knowing that their activities are being recorded may have a chilling effect on conduct, speech, and reading").

[20] **J. ZITTRAIN**, *The Future of the Internet: And How to Stop It*, New Haven, Yale University Press, 2008, 205.

[21] *Id.*

[22] "Data mining is correctly defined as the 'nontrivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data." **T. ZARSKY**, "Mine your Own Business!: Making the Case for the Implications of the Data Mining of Personal Information in The Forum of Public Opinion," 5 Yale J. of L. & Tech. 4 (2004), p. 6.

[23] **J. WEIZENBAUM**, *Computer Power and Human Reason: From Calculation to Judgment*, San Francisco, W.H. Freeman and Company, 1976, p. 3.

humans to react ELIZA as though it were a person, which prompted him to write a book about what computers should never be pressed to do.

In one powerful passage, Weizenbaum argues that the most obvious application of some artificial intelligence techniques is massive surveillance. Weizenbaum observes that, as of 1976, there were "three or four major projects in the United States devoted to enabling computers to understand human speech".[24] According to the "principle sponsor of this work, the Advanced Research Projects Agency … of the United States Department of Defense", (now "DARPA") potential applications were uncontroversial and benign. For instance, the Navy wanted voice recognition technology in order to "control its ships, and the other services their weapons, by voice commands".[25] Weizenbaum rejects this explanation:

> Granted that a speech-recognition machine is bound to be enormously expensive, and that only government and possibly a few very large corporations will therefore be able to afford it, what will they be used for? … There is no question in my mind that there is no pressing human problem that will more easily be solved because such machines exist. But such listening machines, could they be made, will make monitoring of voice communications very much easier than it is now.[26]

Today, many varieties of sophisticated voice recognition technology exist.[27] Weizenbaum was wrong about the range of applications to which voice recognition would eventually be put—such technology has been used in everything from computers for the blind, to voice dialing, to hands-free wheelchairs. He was correct, however, that voice recognition would make massive government surveillance practicable.

Another concern closely related to Weizenbaum's insight that computers endowed with AI can stand in for human surveillants is the notion that AI can bring certain patterns of activity to the attention of humans. Thus, techniques of artificial intelligence have been used to decide where to point cameras or to 'flag' events such as the same face appearing in multiple transit stations. Weizenbaum hints in 1976 at this functionality as well:

> Perhaps the only reason that there is very little government surveillance in many countries of the world is that such surveillance takes so much manpower. Each conversation on a tapped phone must eventually be listened to by a human agent. But speech-recognizing machines could [recognise and] delete all

[24] *Id.*, p. 270.
[25] *Id.*, p. 271.
[26] *Id.*, p. 272.
[27] *See, e.g.*, Mass High Tech, "MIT adds robotics, voice control to wheelchair," Sept. 19, 2008, available online at http://www.masshightech.com/stories/2008/09/15/daily64-MIT-adds-robotics-voice-control-to-wheelchair.html (describing a voice-controlled wheelchair).

> “uninteresting” conversations and present transcriptions of only the remaining ones…[28]

More recently, Israeli legal scholar Tal Zarsky discusses the power of AI to sift through and organise data in seconds that would take a human an eternity. Zarsky argues that “[m]ere surveillance … is not grounds for concern, at least not on its own. The fact that there are an eye watching and an ear listening is meaningless unless the collected information is *recorded and emphasised*”.[29] Zarsky goes on to provide a detailed description of “knowledge discovery in databases” (or “KDD”), in which “complex algorithms, artificial intelligence, neural networks and even genetic-based modeling … can discover previously unknown facts and phenomenon about a database”.[30] These techniques are indeed central to AI applications, in which the ability to search for the right answer↓particularly in a complex and even dynamic environment—is the key to performance.[31] After exploring the dangers of consumer and citizen data profiling, Zarsky concludes that greater public awareness of the AI techniques involved in data mining—well understood within, but not beyond, the field of computing↓will lead to more ethical deployment of KDD.

Thus, according to Weizenbaum, Zarsky, and others, AI plays a role in supporting human surveillance that might otherwise prove impossible. The issue is considered serious enough that a popular AI textbook has cited the potential to invade privacy as one of six principle ethical questions around AI.[32]

II. The role of social machines

A. Robotics and computer interfaces

A long-standing and accelerating goal of AI, especially within the subfields of robotics and human-computer interaction (“HCI”), has been to develop machines and programs that interact more naturally with humans. Due in part to increased worldwide demand for personal robotics (one research agency predicts that personal robotics will be a

[28] **J. WEIZENBAUM**, *Computer Power and Human Reason*, p. 272.

[29] **T. ZARSKY**, “Mine your Own Business!,” p. 4 (emphasis in original).

[30] Zarsky further observes that KDD can make predictions about the future. *Id.*, p. 8 (“After establishing the ‘clustering,’ both *descriptive* and *predictive* inquiries are possible.”) (emphasis in original). Beyond the scope of this essay is whether these techniques create new categories of relevant, invasive personal information that was never disclosed (or perhaps known) to the data subject.

[31] **H.R. EKBIA**, *Artificial Dreams*, p. 44.

[32] **S. RUSSELL and P. NOVIG**, *Artificial Intelligence*, p. 960.

$15 billion dollar industry by 2015),[33] roboticists have made close study of the human reaction to robots in the field. Some have reached the conclusion that humans are less likely to accept robots in certain capacities absent sufficient resemblance to humans and/or social complexity.[34] Thus, for instance, in developing the "Nursebot" Pearl for use in hospitals or elderly care facilities, researches at Carnegie Melon found that "if the Nursebot is too machine-like, her human clients ignore her, and won't exercise or take pills."[35] It was therefore necessary to make Nursebot appear more human and interact more naturally for it to be an effective tool in elderly care.

Other drivers behind socializing robots are the view that "to build systems that have human-level intelligence", it is necessary to "build robots that have not merely a physical body but in fact a humanoid form",[36] and the related hope that very complex behaviors can 'develop' over time through social interaction. Cynthia Breazeal, a pioneer in the emerging field of "social robotics" and the head of the influential MIT Media lab, has helped create a class of "Mobile/Dexterous/Social" robots capable of mimicking emotion and responding to social cues. In describing Kismet, among her first efforts in social robotics, Breazeal told the New York Times: "I hoped that if I built an expressive robot that responded to people, they might treat it in similar way to babies, and the robot would learn form that".[37] Her impressive work continues to advance in this direction.[38]

Software developers and computer engineers have similarly turned to more social interfaces. According to psychological science professor T.M. Holtgraves and colleagues, "[o]ne of the major trends in human-computer interaction … has been the development of more natural human-computer interfaces".[39] Moreover, "[w]ith the advent of the Internet, the appeal of even more natural, human-like interfaces has increased dramatically".[40]

[33] ABA Research, "Personal Robots Are Here," New York, Dec. 28, 2007.

[34] *See* http://robotic.media.mit.edu/projects/robots/mds/social/social.html ("Given the richness and complexity of human life, it is widely recognized that personal robots must be able to adapt to and learn within the human environment from ordinary citizens over the long term."); *see also* http://robotic.media.mit.edu/projects/robots/leonardo/socialcog/socialcog.html ("One way robots might develop socially adept responses that seem to reflect beliefs about the internal states of others is by attempting to simulate –in its own cognitive system – the behaviors of others.").

[35] **P. MCCURDUCK**, *Machines that Think*, p. 467. Conversely, the researchers worried that were Nursebot Pearl *too* humanlike, clients might form unnatural attachments to her. *Id.*

[36] **H.R. EKBIA**, *Artificial Dreams*, p. 259 (citing others).

[37] **P. MCCURDUCK**, *Machines that Think*, p. 454 (citing the New York Times).

[38] *See, e.g.,* **C. BREAZEAL, J. GRAY and M. BERLIN**, "An embodied cognition approach to mindreading skills for socially intelligent robots," International Journal of Robotics Research, 2008 (to appear); **A.L. THOMAZ and C. BREAZEAL**, "Teachable robots: Understanding human teaching behavior to build more effective robot learners." Artificial Intelligence, vol. 172(6-7), 2008, p. 716-37.

[39] **T.M. HOLTGRAVES et al.**, "Perceiving artificial social agents," *Computers in Human Behavior,* 2007, No 23, pp. 2163-2174, at p. 2163.

[40] *Id.*

As a consequence of such research, the number of applications that leverage social dimensions is growing. Companies and other institutions make use of virtual representatives, discussed in greater detail below, in order to handle customer service calls and even sales and recruitment.[41] We are also seeing the deployment of human-like robots into a variety of spaces, including the home for entertainment and service.[42] Many computer systems, particularly those running on cell phones or in an environment that requires a 'hands free' user experience, have moved toward spoken language and other, more natural interfaces.[43] After some initial setbacks,[44] websites are becoming more interactive and personalised.

B. Effects of social machines on humans

It turns out that making robots and computers more social yields a profound effect on humans. Consider social scientist Sherry Turkle's report after her encounter with the social robot Cog in the MIT Media Lab1990s:

> Trained to track the largest moving object in its field (because that will usually be a human being) Cog "noticed" me soon after I entered its room. Its head turned to follow me and I was embarrassed to note that this mad me happy. I found myself competing with another visitor for its attention. At one point, I felt sure that Cog's eyes had "caught my own."[45]

Studies across multiple disciplines have confirmed this human tendency to treat social objects as social, sometimes called the "ELIZA effect" in AI literature after Weizenbaum's program.[46] In their influential book *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places,* Byron Reeves and Clifford Nass detail their findings that humans treat computers as social actors.[47] Their method consists largely of reproducing experiments around known human behaviors toward other humans and substituting social computer for one set of people.[48] In this way, Reeves and Nass show that computers that evidence social characteristics have a similar, or, in some case, the exact same, effect on humans. Computers programmed to be polite, or to evidence certain personalities,

[41] **I. KERR**, "Bots, Babes," *o.c.*

[42] **P. MCCURDUCK**, *Machines that Think*, p. 467.

[43] *See* **D. GARLAN et al.**, "Project Aura: Toward Distraction-Free Pervasive Computing," IEEE Pervasive Computing, vol. 01, no. 2, pp. 22-31, Apr-Jun, 2002.

[44] *See* http://en.wikipedia.org/wiki/Microsoft_Bob (describing Microsoft's unpopular virtual helper).

[45] *Id.*, p. 277.

[46] *Id.*, p. 8.

[47] **B. REEVES and C. NASS**, *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places,* New York, Cambridge University Press, 1996.

[48] *Id.*, p. 14.

have profound effects on test subjects.[49] Humans respond to flattery and criticism from computers,[50] and rate their experiences with computers more highly if the computer has a similar 'personality' (e.g., submissive) to their own.[51] The results applied to people of all ages and of diverse backgrounds, including those with a familiarity with technology.[52]

Further data around human-technology interaction suggests that the more human-like the medium, the greater the response. Canvassing the literature on human interaction with androids↓i.e., "artificial system[s] designed with the ultimate goal of being indistinguishable from humans in its external appearance and behavior"[53]↓informatics professors Karl MacDorman and Hiroshi Ishiguro conclude that "[h]umanlike appearance and behavior are required to elicit the sorts of responses that people typically direct toward one another",[54] and that "the more humanlike the robot, the more human-directed (largely subconscious) expectations are elicited".[55] In one cited study, test subjects exhibited greater unconscious eye contact behaviors (fixating on the right eye, typical of human-human interaction) when engaging with more humanoid robots.[56] In another, Japanese subjects only averted their gaze (a sign of respect) when engaging with the most human-like machines.[57] MacDorman and Ishiguro further offer several anecdotal examples of disparate treatment of robots. For instance, visitors to Ishiguro's lab could be convinced to treat more mechanical robots roughly, but show respect toward Uando, a robot with an enhanced "aura of human presence," due to automated response such as "shifting posture, blinking, and breathing".[58] One visitor reportedly asked his wife's permission before touching a 'female' robot.[59]

Importantly, research also shows that this tendency to anthropomorphise social media can also recreate in humans the sense of being observed. Thus, Terry Burnham and Brian Hare of Harvard University subjected 96 volunteers to a game in which they anonymously donate money or withhold it. Where players were faced with a mere photo of Kismet↓the

[49] *Id*, p. 24.
[50] *Ibid*. (Chapters 2, 4).
[51] *Ibid*. (Chapter 8).
[52] *Id.*, p. 252.
[53] **K. MACDORMAN and H. ISHIGURO** , "The uncanny advantage of using androids in cognitive and social science research," Interaction Studies 7:3, 2006, pp. 298-99.
[54] *Id.*, p. 316
[55] *Id.*, p. 309. There is an apparent point of similarity, often referred to as the "uncanny valley," at which humans can become repulsed by an android. Many theories exist to explain this phenomenon, including that almost human androids create certain expectations that they necessarily violate (in that they are not perfect replicas). *Id.*, p. 299.
[56] *Id.*, p. 316.
[57] *Id.*
[58] *Id.*, pp. 313-14.
[59] *Id.*, p. 317.

robot designed by Cynthia Breazeal to elicit a social reaction in humans↓they gave considerably more then those who were not.[60] In another experiment involving donation, subjects consistently donated more where the computer terminal they were using had eyespots on its screen.[61] In yet another study published in *Biology Letters,* UK psychologists found that the presence of a picture with eyes above a collection bin led people to pay for coffee on the honor system far more often then the presence of a picture of flowers.[62]

The standard explanation for this set of phenomena is that humans evolved at a time when representation was largely impossible, such that what appeared to be real was real in fact. As Reeves and Nass explain, "people are not evolved to twentieth-century technology. The human brain evolved in a world in which *only* humans exhibited rich social behaviors, and a world in which *all* perceived objects were real objects".[63] In evolutionary terms, we are not much further along than our oldest ancestors.

American cognitive science professor H.R. Ekbia puts it slightly differently: humans as highly social animals have developed an innate ability to identify with other humans. This confers a tremendous survival advantage in that it tends to foster cooperation. The ability is often indiscriminate, however, with the result that humans often unconsciously attribute human emotions to objects or animals. Ekbia adds: "The AI community has, often inadvertently, taken advantage of this human tendency, turning what could be called innocent anthropomorphism to a professional and often unjustified, technoscientific one".[64] That is, Ekbia believes that practitioner of AI have sometimes relied on the ELIZA effect to gloss over the difficulty in programming truly fulsome intelligent or social interactions.

II. The privacy threats posed by social AI

A. Old paradigm: AI as active gatherers of information

We have seen that social machines are on the rise and that humans treat social machines as though there were truly people. One application of this insight is to free computers from their historically hidden or passive role within surveillance and set machines

[60] **V. WOODS**, "Pay Up, You Are Being Watched," New Scientist, Mar. 18, 2005 (reporting a 30% increase in giving when faced with Kismet).
[61] **O. JOHNSON**, "Feel the Eyes Upon You," N.Y. Times, Aug. 3, 2008.
[62] **M. BATESON et al.**, "Cues of Being Watched Enhance Cooperation in a Real-World Setting," Biology Letters, 2(3), Sept. 22, 2006, pp. 412–14.
[63] **B. REEVES and C. NASS**, *The Media Equation*, p. 12 (emphasis in original).
[64] **H.R. EKBIA**, *Artificial Dreams*, p. 310.

to the active, interpersonal task of gathering of information. B.J. Fogg is a Stanford researcher who coined the term “captology”↓“an acronym based on the phrase computers as persuasive technologies”.[65] In his 2003 book *Persuasive Technology: Using Computers to Change What We Think And Do*, Fogg details some of the techniques of captology, many of which consist of embedding physical, psychological, and social cues in computer interfaces for a variety of purposes. It turns out that one of the primary applications of persuasive technology has been information gathering.

Building on the work of Reeves, Nass, and others, Fogg directly compares mechanical persuaders to persuasive people. He explains certain advantages thoughtfully modeled computers will typically have. Computers can be more persistent than humans, in that humans tire and respond to social cues such as anger and shame.[66] Machines have no necessary form or clear identity, and can therefore facilitate anonymous persuasion. Computers can also “store, access, and manipulate huge volumes of data”. They can leverage a variety of “modalities,” beyond speech and body language. Computers can “scale,” in the sense of reaching millions of people at once. Similarly, computers can go where ordinary human strangers cannot↓reaching into the home, a bathroom, or even a person’s clothing.

Fogg also details the dangers of persuasion by computer↓some of which overlap his advantages. He identifies six “unique ethical concerns related to persuasive technology”.[67] First, he notes that a technology’s novelty can mask its persuasive intent. Humans may not be ‘on alert’ to an agenda in a neat new gadget. Second, computers have a positive reputation as credible and unbiased; this reputation can be exploited to hide a persuasive intent. Third, unlike sales people, computers do not tire; they can reach thousands simultaneously and persistently. Computers also control all “interactive possibilities”, i.e., the computer decides what happens next and what the user can see or do. Fifth, computers “can affect emotions but can’t be affected by them”.[68] Programmers can expect a social reaction from humans but can control the reaction of the persuasive technology that elicits it. Finally, computers are not “ethical agents”, in the sense that they cannot take responsibility for an error.[69]

The gist of captology, then, is that computers and robots can be pressed into the task of persuading humans to engage in or refrain from behaviors through both direct and subtle social methods. This tasks has so far mostly involved gathering information. Canadian legal

[65] **P.J. FOGG**, *Persuasive Technologies*, p. xxv.
[66] *Ibid.*
[67] *Id.*, p. 213.
[68] *Id.*
[69] *Id.*, pp. 213-220.

scholar Ian Kerr has explored the use of virtual representatives and other online "bots" that leverage techniques of AI and human-computer interaction in order to establish trust with, gather information about, and ultimately influence consumers.[70] In an insightful 2004 law review article Kerr asks, "What if bots could be programmed to infiltrate people's homes and lives *en masse*, befriending children and teens, influencing lonely seniors, or harassing confused individuals until they finally agree to services that they otherwise would not have choose?"[71] The question proves a set up: Kerr observes that "[m]ost such tasks can be achieved with today's bot technologies".[72]

Kerr goes on to detail several "interactive agents" operating on the Web since 2000. One such agent is ELLEgirlBuddy, a text-based virtual representative for ELLEgirl.com that operates over instant messenger ("IM"). As Kerr explains: "ELLEgirlBuddy is programmed to answer questions about her virtual persona's family, school life and her future aspirations, occasionally throwing in a suggestion or two about reading ELLEgirl magazine." Although she has no actually body, she sometimes writes about her body image problems. Although she is in actuality only a few years old, ELLEgirlBuddy purports to be sixteen and seeks to replicate the lingo of a teenager, complete with emoticons.[73]

Among ELLEgirlBuddy's most alarming functions is straightforward data collection. Every single response the bot receives or elicits is recorded↓in all, millions of conversations over IM. This information is used in turn to further deepen the bond↓and therefore trust – between the bot and its interlocutor.[74] (In social robotic parlance, ELLEgirlBuddy is an "expressive robot that respond[s] to people" and, when people treat it like the teen it purports to be, the robot learns form that.) Kerr points that the data has other, commercial value in that it could be used to target advertisements.[75]

The use of virtual personalities is not limited to the private sector. The U.S. Army has deployed an interactive virtual representative for its recruitment website.[76] The program, SGT Star,[77] appears as an avatar. He speaks out loud in addition to displaying text. He can act both funny and agitated, as when in response to a command to do pushups he yells: "Hey, I'm the sergeant, here, YOU drop down and give me twenty! I CAN'T HEAR YOU!!!

[70] **I. KERR**, "Bots, Babes," *o.c.*
[71] *Id.*, p. 312.
[72] *Id.*
[73] Emoticons are faces drawn with text. ;o)
[74] **I. KERR**, "Bots, Babes," p. 316 ("In other words, these companies are constantly collecting incoming data from users and strong that information for the purposes of future interactions.")
[75] *Id.*
[76] *See* http://www.goarmy.com/ChatWithStar.do.
[77] The SGT stands for "strong, trained, and ready." *Id.*

COUNT 'EM!!!" He can also take a compliment; if you tell SGT Star that you like him, he responds: "Thanks, I try".

SGT Star purported function is to engage with users of the GoArmy website in order to answer questions and to provide other guidance such as the location of forms or local recruitment offices. Yet SGT Star also gathers information. As an initial matter, SGT Star prompts the user for his or her name before beginning the chat session. Moreover, the website invites users to sign in and provide more information (e.g., date of birth, address) for a more "personalised" SGT Star experience. SGT Star even invites users to "Tell A Friend" about him by submitting a name and email address, which will cause SGT Star to generate an email invitation to start a chat session with a third party.

According to the GoArmy privacy policy (in general a notoriously under-read document[78]), the Army records everything anyone says to SGT Star. The Army reserves the right to use all information gathered SGT Star for recruiting purposes, and to disclose such information as required by law.[79] The Army may therefore use chat transcripts in the aggregate to improve SGT Star's 'social skills', or to identify particularly promising candidates for eventual follow up by a human recruiter. It remains largely unclear, however, whether the Army might use a SGT Star chat transcript to reject a candidate↓for instance, by discovering the sexual orientation of a potential recruit on the basis of questions he asked about Army policy toward gays↓a question he might not ask of a human recruiter.[80]

In short, through a combination of powerful processing and sophisticated social mimicry, it appears possible for companies and other institutions to collect information from individuals beyond that which even a large human work force could accomplish. As in the context of data mining, a computer equipped with artificial intelligence is capable of engaging thousands of individuals simultaneously, twenty-four hours a day. But here the agent is able to leverage the power of computers to persuade via carefully orchestrated social tactics known to elicit responses in humans. In an age of national security and targeted advertising, citizen and consumer information is at an all time premium.[81] Techniques of AI and HCI create the opportunity for institutions to leverage the human tendency to anthropomorphise and other

[78] *See, e.g.*, **E. MORPHY**, "Consumers Trust Brands, Not Policies," CIO Today, Jan. 29, 2004 (citing research at Michigan State Univeristy).

[79] If you ask SGT Star about privacy, he responds: "I keep a record of all the chats I have with GoArmy users. My conversations are reviewed to ensure all potential recruits are getting the information that they need. However, your information will not be shared with the public."

[80] The U.S. Army uses a "don't ask, don't tell" approach wherein gays may serve as long as they do not self-reveal their orientation. *See* 10 U.S.C. Sec. 654.

[81] *See, e.g.*, **A. MCCLURG**, "A Thousand Words are Worth a Picture: A Privacy Tort Response to Consumer Data Profiling, 98 Nw. U. L. Rev. 63 (2003) (discussing institutional data demand and data mining trends).

advantages computers hold over humans (ubiquity, diligence, trust, memory, etc.) to facilitate and otherwise impracticable depth and breadth of data collection.

B. New paradigm: social AI as subject

We have seen various applications of AI that threaten privacy by substituting a machine or software function for a human task, thereby augmenting the human power to observe. Thus, AI can listen to phone conversations so that a human does not have to, comb through video or other data better and faster than a human, or be sent out to recruit or inquire on behalf of an organization using much of the social leverage of a real person, but with none of the limitations. Yet given the power of social AI to signal the presence of a human, the distribution of social machines may have still deeper, if unintended, impacts on privacy.

As an initial matter, the appearance of social AI in historically private spaces may lessen the opportunity for solitude and free reflection. Many privacy theorists have expounded upon the importance of private space, wherein one can "be themselves" and even transgress otherwise oppressive social norms. As Alan Westin famously writes in his 1970 treatise on privacy, *Privacy and Freedom*: "There have to moments 'off stage' when the individual can be 'himself'; tender, angry, irritable, lustful, or dream filled. … To be always 'on' would destroy the human organism".[82] Westin further cites the "need of individuals for respite from the emotional stimulation of daily life. … [T]he whirlpool of active life must lead to some quiet water, if only so that the appetite can be whetted for renewed social engagement."[83] According to Westin, "[p]rivacy provides the change of pace that makes life worth savoring".[84] For Westin, privacy protects "minor non-compliance with social norms" that "society really expects many persons to break", and the important opportunity to "deviate temporarily from social etiquette".[85]

Many other scholars have explored the same line of thought. In the words of political theorist Hannah Arendt, "[a] life spent entirely in public, in the presence of others, becomes … shallow. … A space apart from others has enabled people to develop artistic, political, and religious ideas that have had lasting influence and value when later introduced into the public sphere".[86] American law scholar Paul Schwartz argues that the belief that one is being

[82] **A. WESTIN**, *Privacy and Freedom*, New York, Antheum 1970, p. 35.
[83] *Id.*
[84] *Id.*
[85] *Id.*
[86] **D. SOLOVE**, "A Taxonomy of Privacy," pp. 554-55.

constantly observed interferes with self-determination.[87] Julie Cohn argues similarly that "pervasive monitoring of every first move or false start will, at the margin, incline choices toward the bland and the mainstream".[88] According to prolific privacy scholar Daniel Solove, "[n]ot only can direct awareness of surveillance make a person feel extremely uncomfortable, but it can also alter her behavior. Surveillance can lead to self-censorship and inhibition". Solove further notes that "[e]ven surveillance of legal activities can inhibit people from engaging in them".[89]

If, as many contend, safeguarding occasional solitude is a central goal of privacy law, social AI may bypass existing protections by introducing the appearance and feeling of a human in an unlimited array of contexts. Consider just a few examples: Japanese software company MetaboInfo makes a virtual wife that appears on your cell phone and reminds you of chores;[90] an exhibit of a future house by Disney showcases a mirror that sees you and suggests outfits;[91] the Figera vacuum approaches its owner and responds to voice commands;[92] experiments are being conducted around whether placing robots or voice-interface computers in cars can stop road rage; etc.[93] As captology expert BJ Fogg explains, "computing applications are becoming commonplace in locations where human persuaders would not be welcome, such as bathrooms and bedrooms, or where humans cannot go (inside clothing, embedded in automotive systems, or implanted in a toothbrush)".[94]

Social AI may also impact personal privacy in the aggregate, by creating unease around the massive computer custodianship of human data. Hardly any aspect of human life today remains untouched by computers; this trend will only grow as computer become embedded into our streets, walls, and even our clothing. Meanwhile, the public sense of computer intelligence and evaluative capabilities↓fueled by our tendency to anthropomorphise, by the rise in prominence of tech media coverage, and by claims of competitive practitioners↓continues to develop. This synergy could, in theory, lead to widespread and intractable discomfort with computer information custodianship.

[87] *Id.*, p. 494.
[88] *Id.*
[89] **D. SOLOVE**, "'I've Got Nothing To Hide' and Other Misunderstandings of Privacy," 44 San Diego L. Rev. 745 (2007), p. 267.
[90] *See* http://www.metaboinfo.com/okusama/ (website in Japanese).
[91] **P.J. O'RORKE**, "Future Shlock," The Atlantic, Dec. 2008.
[92] *See* http://gizmodo.com/5105633/the-figura-vacuum-bot-allows-you-to-boss-it-around.
[93] **T.E. GALOVSKI and E.B. BLANCHARD**, "Road rage: A domain for psychological intervention?", *Aggression and Violent Behavior*, 2004, No 9, pp. 105—127.
[94] **P.J. FOGG**, *Pervasive Computing*, at p. 10.

Artificial intelligence has clearly seen its share of breakthroughs throughout its history, many of which have been widely reported by the media.[95] The field stands poised to make many more. In part by leveraging well-understood AI tactics and incredible but steady gains in computational power, projects such as the Defense Advanced Research Agency (DARPA)'s "Cognitive Agent that Learns and Organises" are making notable strides in advancing computer learning, and setting ambitious but attainable long-term goals.[96]

Indeed, computer and robotics insiders publicly predict that machines will be as or more intelligent than humans within a few decades. Jim Gray of Microsoft Research has speculated that computers will pass the famous Turing Test↓i.e., the test of machine intelligence devised by Alan Turing wherein a machine must fool a trained expert into believing it is human↓by the middle of this century.[97] A German software program competing in the International Turing Competition managed recently to fool three of twelve judges into believe it was a person.[98] Speaking as a keynote at a large technology conference, Justin Ratner, Intel's chief technology officer, observed in August 2008:

> The industry has taken much greater strides than anyone ever imagined 40 years ago. There is speculation that we may be approaching an inflection point where the rate of technology advancements is accelerating at an exponential rate, and machines could even overtake humans in their ability to reason, in the not so distant future.[99]

According to a recent report, the manager of the Adaptive Systems group at Microsoft estimated that "about a quarter of all Microsoft research is focused on AI efforts".[100] Google founders Sergey Brin and Larry Page have repeatedly articulated their goal of creating "obviously artificial intelligence", in the sense of a truly "smart" program that "understands"

[95] *See, e.g.*, http://www.sciencedaily.com/news/computers_math/artificial_intelligence/ (compiling artificial intelligence headlines); http://ai-depot.com/news/ (same); http://www.aaai.org/AITopics/pmwiki/pmwiki.php/AITopics/AINews (same).

[96] **R. BRACHMAN and Z. LENIOS**, "DARPA's New Cognitive Systems Vision," Computing Research News, Vol. 14/No. 5, pp. 1, 8. (Nov. 2002):

> A cognitive computer system should be able to learn from its experience, as well as by being advised. It should be able to explain what it was doing and why it was doing it, and to recover from mental blind alleys. It should be able to reflect on what goes wrong when an anomaly occurs, and anticipate such occurrences in the future. It should be able to reconfigure itself in response to environmental changes. And it should be able to be configured, maintained, and operated by non-experts.

[97] *Id.*, p. 501. *See also id.*, p. 460 (robotics pioneer Hans Moravec predicting strong AI by 2030).

[98] **WILL PAVIA**, "Machine takes on man at mass Turing Test," Times Online (Oct. 13, 2008), available online at http://technology.timesonline.co.uk/tol/news/tech_and_web/article4934858.ece.

[99] Intel News Release, Aug. 21, 2008, available online at http://www.intel.com/pressroom/archive/releases/20080821comp.htm?cid=rss-90004-c1-211570.

[100] **J. GASKIN**, "Whatever Happened to Artificial Intelligence?," Network World, Jul. 23, 2008 (emphasis added).

user queries and the universe of potential results to the point that searches as well as a human with immediate access to most of the Internet.[101]

Clearly the impact of "strong" artificial intelligence↓in the John Searle sense of actual self-awareness↓would be profound across all sectors.[102] Predictions of strong AI have fallen flat before, however, and many within the field argue that humans may never recreate actual intelligence.[103] This particularly achievement is at a minimum decades away. A potentially more interesting question in the short run (i.e., the next five to ten years) is whether computers will reach a level of sophistication at which humans become unsure of the AI's intelligence and, consequently, *uncomfortable* with their extensive 'knowledge'.

Today, humans appear to trust computers and computer servers with their personal information. The prevailing view of computers remains the desktop↓a complex but lifeless automaton that manipulates data without interest.[104] Thus, in seeking to allay fears over its practice of scanning web-based email messages in order to display contextual advertisements, the Internet giant Google is careful to represent that the scanning is conducted by a computer. "Google does NOT read your email… Gmail [or Google Mail] is a technology-based program, so advertising and related information are shown using a completely automated process".[105]

In the context of national security, American thought leaders debate whether machine shifting through public and private data can amount to a government invasion. Judge and scholar Richard Posner argues that "[m]achine collection and processing of data cannot, as such, invade privacy", such that computer data access or citizen surveillance does not in and of itself trigger a search or seizure for purpose of the Fourth Amendment.[106] Prosecutor turned legal scholar Orin Kerr also holds that no search occurs until "information from or about the data is exposed to human observation," not when it is simply "processed by a

[101] *See* http://ignoranceisfutile.wordpress.com/2008/09/13/google-founders-artificial-intelligence-quotes-archive/ (collecting AI quotes from Google principals). Brin reportedly said the following in November of 2002: "Hal could… had a lot of information, could piece it together, could rationalise it. Now, hopefully, it would never… it would never have a bug like Hal did where he killed the occupants of the space ship. But that's what we're striving for, and I think we've made it a part of the way there." *Id.*

[102] *See* **L. SLOCUM**, "Legal Personhood for Artificial Intelligence," 70 N.C. L. Rev. 1231 (1992) (discussing whether AI could serve as a trustee); *id*. (discussing John Searle). *See also* **C. STONE**, "Should Trees Have Standing? Toward Legal Rights for Natural Objects," 45 Cal. L. Rev. 450, 453-57 (1972) (discussing whether AI could have standing).

[103] *See* **S. RUSSELL and P. NORVIG**, *Artificial Intelligence*, pp. 947-60 (canvassing the literature).

[104] It is precisely this human view of computers as unbiased, trustworthy data processors that creates the opportunity for persuasion present in captology. **B.J. FOGG**, *Persuasive Technologies*.

[105] *See* http://mail.google.com/support/bin/answer.py?answer=6599&topic=12787.

[106] **R. POSNER**, "Our Domestic Intelligence Crisis," The Washington Post, Dec. 21, 2005.

computer".[107] Law professor Larry Lessig also uses the example of a search by a government computer program that mindlessly borrows through citizen data (a so-called "worm") to test the parameters of search and seizure law in cyberspace.[108]

This image of a passive conduit may change, however, if and when computers reach a threshold of apparent intelligence wherein processing begins overly to resemble human judgment. Given a handful of factors↓namely, the human tendency to anthropomorphise discussed in detail above, the aggressive claims of AI practitioners and critics, the occasional hyperbole of the media, and the lack of any definitive test of intelligence↓humans could come to equate computer mentality with human mentality in the relatively near term. This in turn could lead to an uncomfortable reexamination of computers as passive custodians of consumer and citizen data.[109]

III. The (in)adequacy of U.S. privacy law

American privacy law already contains the seeds of a solution to many of the emerging privacy harms identified in this essay. A notorious patchwork, American privacy law nevertheless requires notice about the collection, use, and disclosure of personal and other information[110] and is relatively steadfast in its protection against invasions into private space without adequate process.[111]

The privacy community has already begun to propose concrete solutions to perceived abuses of sophisticated and widespread data mining by government and industry. Andrew McClurg argues, for instance, for a resuscitation of the U.S. common law tort of appropriation (discussed by Warren and Brandeis in *The Right to Privacy*[112]) as a response to the creation and use of consumer profiles.[113] Appropriation refers to the use of another's

[107] **O. KERR**, "Searches and Seizures in a Digital World," 119 Harv. L. Rev. 531, 551 (2005).

[108] **L. LESSIG**, *Code 2.0*, New York, Basic Books, 2006, pp. 20-23.

[109] Changes to user interfaces may have more immediate effects, however, in the realm of voice-driven search. In a recent interview, Google's vice president for search Marissa Mayer discussed the hope that, within a few years, users might be able to interact with Google orally by "asking questions by voice." *See* http://www.techcrunch.com/2008/12/10/marissa-mayer-at-le-web-the-almost-complete-interview/. Query whether individuals will search for the same things when it feels to the user that they are speaking with a person.

[110] *See, e.g.*, California Online Privacy Protection Act of 2003, Bus. & Prof. Code Sec. 22575-22579 (California statute requiring companies that collect personal information to link to a privacy policy). The FTC also holds companies to their claims about data and sets minimum thresholds of notice for material changes to policy. *See, e.g.*, *In the Matter of Gateway Learning Corp.*, FTC File No. 042-3047 (2004).

[111] **D. SOLOVE**, "A Taxonomy of Privacy," 154 U. Pa. L. Rev. 477 (2006), p. 552 ("For hundreds of years, the law has strongly guarded the privacy of the home.").

[112] **S. WARREN AND L. BRANDEIS**, "The Right to Privacy," 4 Harv. L. Rev. 193 (1890).

[113] **A. MCCLURG**, "A Thousand Words are Worth a Picture."

identity↓generally, their name or “likeness”↓to one’s own benefit without consent. Such a use can amount to an invasion of privacy under American common law.[114] McClurg argues convincingly that the digital profile that results from sophisticated data mining constitutes an “inner identity” that can trigger the tort.

American law professor Daniel Solove also urges a more comprehensive understanding of privacy law that encompasses the “Kafkaesque” nature of modern surveillance[115] and has shown how to arrive at an appropriate balance between security interests and privacy rights by meticulously cataloguing the harms of data mining. Digital rights groups such as the San Francisco based Electronic Frontier Foundation have brought suit against telephone providers and the government itself in an effort to understand and domesticate government data mining.[116]

Similarly, the use of social media to persuade consumers to give up information or to purchase particular products has a ready analog in tactics already being investigated by national and local consumer protection agencies. In the United States, Section 5 of the FTC Act prohibits “unfair or deceptive trade practices,” broadly defined.[117] The Federal Trade Commission is charged with enacting and enforcing policy aimed at prohibiting unfair, deceptive, or anti-competitive practices within the industries in its jurisdiction. The agency has turned its attention in recent years to online data collection practices such as the traffic of users’ surfing habits,[118] as well as the use of ‘buzz’ or subliminal marketing wherein products are promoted without notice that the speaker is affiliated with an advertising company. State attorneys general have also investigated online information gathering practices and, in cases, reached agreements with companies perceived to gather or use data too aggressively.[119] Ian Kerr explains that the use of AI bots particularly for marketing and consumer information gathering may violate similar Canadian consumer protection regulations.[120]

[114] *See* Restatement (Second) of Torts Sec. 652C (1977).
[115] **D. SOLOVE**, “‘’I’ve Got Nothing To Hide,’’” p. 756.
[116] *See* EFF Press Release, “EFF Sues NSA, President Bush, and Vice President Cheney to Stop Illegal Surveillance,” Sept. 18, 2008, available online at http://www.eff.org/press/archives/2008/09/17-0.
[117] Federal Trade Commission Act, 15 U.S.C. Secs. 41-58, as amended.
[118] FTC Press Release, “FTC Staff Proposes Online Behavioral Advertising Privacy Principles,” Dec. 20, 2007, available online at http://www.ftc.gov/opa/2007/12/principles.shtm.
[119] Online adverting company DoubleClick entered into a consent decree with a coalition of state attorneys general in 2001, agreeing not to combine certain categories of information following a merger with offline consumer profiler Abacus. *See, e.g.*, Washington State Office of the Attorney General Press Release, “States Settle with DoubleClick,” April 2001, available online at http://www.atg.wa.gov/pressrelease.aspx?&id=5848.
[120] **I. KERR**, “Bots, Babes,” at p. 321 (“The fair information practices set out in Appendix 2 of the Canadian Code contain a number of requirements that are clearly not respected by ActiveBuddy and many other bot-based business models.”).

In other cases, however, the law may have no obvious starting point in addressing these emerging privacy harms. As discussed above, the effect of social media is often a direct but subconscious one. It is not that humans will use a technology to invade one another's privacy; rather, the object will be treated as itself human. The danger is that voice-driven, natural language interfaces will become the norm; that computers will increasingly be endowed with personalities; and that robots with anthropomorphic features will come to be voluntarily accepted as a daily part of life (as is increasingly the case in Japan). Simultaneously, but at an examined level, privacy will be eroded by the subconscious perception that we are always being watched and evaluated.

An extreme example with intentional and obvious chilling effects on speech, such as a holographic police officer that follows around each citizen, could in theory trigger the First Amendment of the U.S. Constitution.[121] But there may be no immediate legal solution to a diffuse introduction of social media into private space by natural means. Similarly, the discomfort we may begin to feel at AI custodianship of data may not be reducible to a legally cognizable injury. Although real anxiety could result, perhaps little more can be said about AI capable of extremely accurate judgments or vested with the appearance of common sense is that it is 'creepy'. American privacy law may be ill-suited to protect against such subtle and (for now) speculative harms.[122]

Solutions may ultimately come from outside of privacy law. It turns out that in other contexts, American law forces consideration of subjective interests such as fear or discomfort. Thus, for instance, in the (pun-ridden) case of *Stambovsky v. Ackely*, a New York appeals court recognised a buyer's right to rescind purchase of a home after he learned that it was haunted by a poltergeist.[123] It was no reply that poltergeists do not exist. The buyer could not be forced to live with a ghost merely because the existence of ghosts has not been established.[124] Sellers and brokers must also disclose other stigmas such the occurrence in a home of a multiple murder.[125] In the context of pollution, litigants have pursued a variety of

[121] *Cf. Laird v. Tatum*, 408 U.S. 1, 11 (1972) ("In recent years this Court has found in a number of cases that constitutional violations may arise from the deterrent, or 'chilling,' effect of governmental regulations that fall short of a direct prohibition against the exercise of First Amendment rights.").

[122] **D. SOLOVE**, "A Taxonomy of Privacy," p. 562-63 ("Too many courts and policymakers struggle even identifying the presence of privacy problems. . . . Unfortunately, due to conceptual confusion, courts and legislatures often fail to recognise privacy problems . . .").

[123] 169 A.D. 2d 254 (N.Y. Ct. App. 1991).

[124] In *Stambovsky*, the tongue-in-cheek court actually held the house to be haunted "as a matter of law." *Id.*

[125] *See Reed v. King*, 145 Cal. App. 3d 261 (1983) (holding that plaintiff stated a cause of action for defendant-broker's failure to disclose that house was site of multiple murder).

harms bred of unrealised fears.[126] Accordingly, one can imagine a requirement that an entity disclose that it is using sophisticated AI and offer to store user information separately.

Another useful analog might be the requirement of warning labels for non-obvious product defects. The Food and Drug Administration and individual states often require harmful goods to contain warnings as to their contents, and the existence of a warning label can sometimes provide a meaningful defense against a civil action for product liability. Where a car comes equipped with an active AI passenger, for instance, its user manual could warn that humans react to social machines as though they were truly humans, and that constantly being in the presence of others can lead to discomfort.

Viable solutions are equally likely to come from outside the law, especially in the short term. They might include the inclusion of privacy in ethics discussions around social media, the participation of developers of AI in efforts to build privacy protections into emerging technology,[127] and sustained efforts at public education by industry and government.[128] In his aforementioned book on captology, Fogg creates a framework by which to assess the ethical implications of a given instance of persuasive technology. He concludes that:

> Ultimately, education is the key to more ethical persuasive technologies. Designers and distributors who understand the ethical issues … will be in a better position to cerate and sell ethical persuasive technology products. Technology users will be better positioned to recognise when computer products are applying unethical or questionably ethical tactics to persuade them.[129]

Calling attention to and discussing these phenomena is a necessary first step to heading off or addressing a novel set of privacy threats.

[126] *See, e.g., City of Santa Fe v. Komis*, 845 P.2d 753, 757 (N.M.1992) (awarding land owner damages due to fear of nuclear waste); *Lunda v. Matthews*, 613 P.2d 63, 67-68 (Or.Ct.App.1980) (allowing emotional distress damages for fear of air emissions from cement plant); *Heddin v. Delhi Gas Pipeline Co.*, 522 S.W.2d 886, 888 (Tex.1975) (awarding damages to landowner due to fear that pipeline on adjoining land would explode); *Texas Elec. Serv. Co. v. Nelon*, 546 S.W.2d 864, 871 (Tex.Civ.App.1977) (allowing landowner to recover for fear of nuclear waste transported nearby).

[127] UK's Information Commissioner's Office has, for instance, commission the Enterprise Privacy Group to produce a new report on the impact on personal privacy of various activities across multiple industries. Applications of social AI should be included in such a report.

[128] *See also* **T. ZARSKY**, "Mine Your Own Business!," Sec. III (discussing the role of public education in addressing AI data mining techniques).

[129] **P.J. FOGG**, *Persuasive Technologies*, p. 235.

IV. Conclusion

Our conception of what constitutes an invasion of personal privacy continues to evolve↓over time, dramatically. Consider the origin of the term "Peeping Tom." Tom was an adolescent with the bad luck to be within the city limits of Coventry when Lady Godiva made her (in)famous naked ride to protest taxes. Unlike other young men, Tom openly gawked at Lady Godiva's naked form as she passed. Today, were a young man *not* to gawk at naked woman on a horse, we might be amazed. We would certainly give no credence to a complaint by or on behalf of the naked woman. (We would say that she willingly exposed herself in public where she has no expectation of privacy.) At the time of the legend, circa 1050, Tom was blinded for his impudence.[130]

Even as our privacy norms evolve, however, a set of basic biological facts remains constant: humans react to social media as though it were human.[131] This disconnect between the state of evolution and the state of our technology continues to be exploited↓sometimes inadvertently↓by developers of certain types of AI in order to develop machine intelligence, foster machine acceptability, and improve user experiences. As a consequence, humans may face a meaningful reduction in their already waning privacy. Upon a thorough canvass of the literature, German privacy theorist Beate Rössler concludes that "a person's privacy can be defined, therefore, in these three ways: as illicit interference in one's actions, as illicit surveillance, as illicit intrusions in rooms or dwellings".[132] Particular techniques of artificial intelligence can be said to violate each of these definitions.

Clearly, artificial intelligence has led to important medical, commercial, and other benefits, and promises many more. And where AI merely supports a human practice↓as in the case of data mining or interviewing consumers↓the law seems well-equipped to provide a meaningful solution. All that may be needed is to expand the law through ordinary methods to encompass and limit the underlying offensive activity. In other cases the solution is not as simple. More subtle and comprehensive changes may be required to mitigate the impact of sophisticated social agents in our midst. Ultimately, however, it may be that "we won't know enough to regulate [AI] until we see what it actually looks like".[133]

[130] **D. SOLOVE**, "A Taxonomy of Privacy," p. 492.

[131] **B. REEVES and C. NASS**, *The Media Equation*.

[132] **B. ROSSLER**, *The Value of Privacy*, Cambridge, Polity Press, 2005, p. 9.

[133] **J. MCCARTHY**, "Problems and Projections in CS for the Next 49 Years," Journal of the ACM, 2003.