# What We Should Do Before the Social Bots Take Over: Online Privacy Protection and the Political Economy of Our Near Future

Erhardt Graeff
MIT Media Lab
erhardt@media.mit.edu
Presented at MIT8, May 5, 2013

Direct interactions between humans and bots generally conjure up images from science fiction of *Terminator* robots or artificial intelligence gone rogue, like *2001*'s HAL or *The Matrix*. In reality, AI is still far from much of that sophistication, yet we are already faced with the ethical and legal ramifications of bots in our everyday lives. Drones are being used for collecting military intelligence and bombing runs. U.S. states have passed laws to address self-driving cars on public roads (Marcus 2012). And nearer the subject of this paper, the legality of search engine bots (*web crawlers*) has been openly questioned on grounds of intellectual property protection and trespassing (Plitch 2002).

Bots inspire fear because they represent the loss of control. These fears are in some ways justified, particularly on grounds of privacy invasion. Online privacy protection is already a fraught space, comprising varied and strong positions, and existing laws and regulations that are antiquated many times over by the rapid growth and innovation of the internet in recent decades. The emergence of social bots, as means of entertainment, research, and commercial activity, poses an additional complication to online privacy protection by way of information asymmetry and failures to provide informed consent.

In the U.S., the lack of an explicit right to privacy and the federal government's predilection for laissez faire corporate regulation expose users to a risk of privacy invasion and unfair treatment when they provide personal data to websites and online services, especially those in the form of social bots. This paper argues for legislation that defines a general right to privacy for all U.S. citizens, addressing issues of both access and control of personal information and serving as the foundation for auditable industry design standards that inherently value and honor users' rights to privacy.

## Defining Social Bots

> 1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
> 2. A robot must obey the orders given to it by human beings, except where such orders would conflict with the First Law.
> 3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.
>
> (Science fiction writer Isaac Asimov's "Three Laws of Robotics" quoted in Burger 2009)

Most contemporary (ro)bots are not in the business of directly killing or saving humans, and most lack the humanoid physique we are used to seeing in popular culture. In fact,

the category "robots" spans a broad array of artificial intelligences, made manifest physically and virtually, and with varying levels of autonomy. This paper is concerned with what are commonly called *bots*, a class of software agents that automatically perform digital tasks on behalf of users. Web crawlers are the most prevalent examples of bots online, often employed by search engines to index webpages, or by spammers to harvest email addresses and other personal data from public websites. Spammers also employ *spambots* to send email or post content on social media. However, both web crawlers and spambots are far from autonomous and have been designed with very specific tasks in mind.

More sophisticated and intelligent bots include software agents that learn a user's preferences over time and recommend products for purchase, such as in the context of Google Adsense or Amazon.com Recommendations, or they create customized news feeds like Google News or the Zite iPad app. In and of themselves, these personal assistant bots raise a host of important ethical and policy issues around human agency and the control of personal information, but the real focus of this paper is *social bots*, which introduce even more dramatic complications to these same issues.

By social bots, I mean software agents that engage in two-way, direct communication with human users through natural language. The classic example of a social bot is a *chatterbot*, which can engage in synchronous or near synchronous communication for entertainment or customer service purposes—a famous example being *SmarterChild* on AOL Instant Messenger, which answered simple questions about itself and looked up information online at a request. A *twitterbot* is a more contemporary example of social bot, and another one that inhabits a specific communication network, in this case, Twitter. Unlike *SmarterChild*, twitterbots have been shown to successfully engage with human users while hiding the fact that artificial intelligence is driving the conversation, and notably they have been deployed for both good and ill. Anti-social examples of twitterbots include those used during the 2012 Mexican election season by the Institutional Revolutionary Party to make it appear that a huge groundswell of users were against the opposing party's candidate (Orcutt 2012). Pro-social examples of twitterbots often aim to connect human user communities together, such as the social bot entrants to the competitions run by the Pacific Social Architecting Corporation (Nanis, Pearce, and Hwang 2011) and Greg Marra's *@Trackgirl* (McMillan 2012). *@Trackgirl*, which simply copies and pastes other users tweets as its own, illustrated the unique risk to privacy social bots pose when it received sympathetic notes from other Twitter users after posting a note about hurting an ankle (McMillan 2012). It's this potential for users to anthropomorphize and even empathize with a social bot that may open the gate for unprecedented invasions of privacy.[1]

[1] There are also social *robots* that take the more familiar humanoid or animal forms, such as Sony's *AIBO*. The added features of a face that "look" at you and/or express emotion, significantly increase the potential for anthropomorphizing these robots; their greater autonomy and physical qualities raise additional questions not only of risk shielding humans from them, but possibly extending legal protection to the robots too (Darling 2012). The privacy issues and regulatory proposals addressed in the follow sections for the case of social bots may be broadly applicable to physical social robots, but the additional psychological and legal complexities (think Google's self-driving cars) are beyond the scope of the paper.

## Unique Privacy Problems Posed by Bots

> A robot must establish its identity as a robot in all cases. (Science Fiction writer Lyuben Dilov's "Fourth Law of Robotics" quoted in Burger 2009)

The Laws of Robotics defined by science fiction writers were safety measures codified into the very decision structures of robots' artificial intelligence software; stories would then explore what happens when this software logic failed to capture the complexity and nuance of real world situations, such as what might happen if a police robot has to kill a malevolent human to save an innocent one (Sawyer 2007). The inherent connection between software and rules and regulations is now an important concept in jurisprudence and relevant to the issue of social bots and privacy. Lawrence Lessig (1999) argues for this view of software code as a regulator in and of itself by saying that the architecture it produces can serve as an instrument of social control on those that use it. Lessig's concern over what has been dubbed the *Code as Law* phenomenon is that these *de facto* laws have the potential to supersede our *de jure* rights, including privacy.

The problem with Code as Law as a policy framework for things like bots is that it's too reductionist and deterministic—failing to account for the social embeddedness of technologies. Social bots are not coded to invade privacy or not invade privacy. That said, "artifacts" like bots can be imbued with politics—constructed both literally and socially according to the political goals and biases of their creators and users (Winner 1986). The history of internet technology is one in which the end use of technology is not always anticipated. For instance, Twitter was originally designed as a tool for group status updates via SMS to allow people to meet up offline more spontaneously. The politicization of the tool came less out of its design and more from subsets of users that saw its limited broadcast potential as a political organizing and information dissemination tool and used it as such (Lotan et al. 2011). Only in recent years has Twitter's co-founder Evan Williams described a primary principle of the social network as "be a force for good" (Siegler 2010).

Where these socio-cultural and technical principles become particularly blurry is when a platform itself starts taking a more active role rather than an expected passive role. The further an internet property deviates from the definition of a *common carrier*, which is a neutral conveyor of goods or content, the more the freedom and equality of users online are challenged.[2] Connections between humans and even between humans and corporations are often based on *social trust* arising from the presumption of shared norms (or laws) and values. The socio-cultural context of connections or transactions that determines social trust is convoluted by the information asymmetries created by anthropomorphic social bots.

This is, in part, by design. Anthropomorphism is the innovation in social bots that make them work as worthwhile interlocutors online thanks to sufficiently human-like visual, verbal, and/or textual cues. Brian R. Duffy has described anthropomorphic social bots as the "ideal interface:"

[2] This is the core argument for *net neutrality*, in which internet service providers would be classified as common carriers and liable if they failed to treat all content on their networks equally through guaranteed conveyance from source to destination (Wu n.d.). boyd (2010) has even suggested that social networks like Facebook should be regulated as "utilities."

> It has often been said that the ultimate goal of the human-computer interface should be to "disappear" the interface. Social robotics is an alternate approach to ubiquitous computing. The interaction should be rendered so transparent that people do not realise the presence of the interface. As the social robot could be viewed as the ultimate human-machine interface, it should similarly display a seamless coherent degree of social capabilities that are appropriate for interaction and to it being a machine. (Duffy 2003)

Whereas physical robots have only recently started to gain an ability to appear human, the barrier online is lower because bots need only interact via text and avatar, which are the same constraints facing humans when they participate on a website or social media network. It's the embodiment of the now classic cartoon adage, "On the internet, nobody knows you're a dog" (Steiner 1993). Users have become accustomed to the abstraction of talking to friends' static photos or simply their usernames alongside periodically updating text. A social bot can easily enter these spaces and reproduce all of the same characteristics we would expect from a human, and without informed consent might lead to an expectation of social trust and corollaries of personal privacy, wherein the human interlocutor shares more or different information with a bot than they otherwise would; in fact, social *robots* have already been shown to elicit intimate, confessional dialogue from users (Darling 2012).

Consider a hypothetical internet startup that sells widgets. They decide to employ social bots to interact online with likely buyers of widgets. The bots are part of an advertising strategy that human public relations employees already use on social media platforms—they attempt to create real relationships with users on a network in order to better understand their customer base and engender brand awareness and loyalty. Users may or may not be aware of the fact that they are interacting with a bot, but the conversation and relationship is continuous because the bot is always available and responsive. As the relationship between the social bot and the user matures, the conversation might span both public and private social media spaces (such as Twitter's direct messages), wherein a user might expect a greater degree of privacy or discretion from a human interlocutor. However, the bot may not acknowledge the nuances of such social norms and ethics; moreover, the company that runs the bot is collecting all of this data. While it's feasible that a human or team of humans could undertake such an advertising strategy on behalf of a company, it's unlikely to scale to the number of relationships necessary to make it cost effective. This poses no challenge to a social bot, which has perfect memory and requires no sleep or overtime pay. An unlimited number of relationships could be maintained through a social bot with the level of responsiveness necessary to produce intimate connections.

The better the machine learning algorithms powering a social bot's artificial intelligence the more data they can process and use to improve their social interactions. This means the potential creation of more intimate interactions based on historical data collected from you or from others in your friend network, including discussions of personal relationships—significant others and kids, work or life complaints and concerns, and hobbies (both conventional or embarrassing—the bot will simply meet you where you are at and affirm you). Extracted personal data can also go beyond text if you

share personal photographs and videos or link to those that you like; there are also data that may be invisible during social interactions with bots but which they are aware of: time, location (GPS data from mobile phones or IP addresses of networked computers), and even purchase records, depending on what corporation or even data sharing consortium the bot is affiliated with.

Many users are already sharing these data publicly online or with corporations and social networks—a situation that underlies broader, contemporary concerns over online privacy invasion. The data may be resold or otherwise proliferated online, or create situations whereby you are misrepresented online because of an error in the processing of information you shared sarcastically or in jest. Social bots may augment these potential harms because of the unprecedented volume of data they can collect, collate, and use without the informed consent of users lulled into false senses of social trust by their anthropomorphized interlocutors. A vicious cycle may unfold, wherein your, and your friends', data are used to produce more intense cases of simulated familiarity, empathy, and intimacy, leading to greater data revelations.

The current way in which online users are "informed" of how their data will be used and their privacy protected is insufficient for the risk posed by social bots. Users are forced to implicitly agree to publicly posted privacy policies that can only be found by clicking on tiny text links at the bottom of a website's pages. Users either agree by "clicking through" a Terms of Service (TOS) agreement when engaging with a site, or passively by continuing to use a web service with a TOS that goes into effect at first use, as specified in its legalese. Users usually don't know what they are agreeing to in these cases. For instance, a platform's porous privacy policy might extend implicit consent to cover data collection by third-party social bots acting on it—and in so doing indemnify the platform from negligence. We are already seeing such issues arise in response to the consolidation of privacy policies and corporate relationships in the cases of Google's subsidiaries and Facebook's recent acquisition of Instagram (Tickle 2012). Combining these existing forms of information asymmetry with the "invisible" quality of anthropomorphic interfaces significantly compromises the ability of users to be sufficiently informed about how and when their data is being stored and used.

## Proposal for Online Privacy Protection covering Social Bots

> [A] properly designed Friendly AI does not, as in popular fiction, consist of endless safeguards and coercions stopping the AI from doing this, or forcing the AI to do that, or preventing the AI from thinking certain thoughts, or protecting the goal system from modification. That would be pushing against a lack of resistance—like charging a locked door at full speed, only to find the door ajar. If the AI ever stops wanting to be Friendly, you've already lost. (Singularitarian Eliezer S. Yudkowsky 2001)

New law in the EU ("Proposal for a DIRECTIVE…") and policy recommendations in the U.S. (Federal Trade Commission 2012) have been produced in the past year in attempts to reform grossly antiquated rules about digital data privacy. However, these regulations still fall short of anticipating the risks posed by social bots. Marc Rotenberg argues, 'the history of privacy protection is the history of the effort to regulate the design of

technology' (2001, I.A.15). The way this has been most recently and insufficiently addressed by U.S. regulators, according to Rotenberg, is by proposing *Privacy by Design*. The base concept of Privacy by Design is to create guidelines for industries that enable them to self-regulate by adding functionality to their products for proactive users to opt out of transactions in which their personal information could be collected (FTC 2012). Examples of this are the privacy settings that Facebook offers users the ability to control which other users can see their posts. However, this makes *opt-out* the default model of privacy protection.

An alternative system for privacy protection is proposed by the *Do Not Track* movement, which is gaining traction among internet companies and drawing the ire of the advertising industry (Singer 2012). The Do Not Track movement comprises: proposals by the FTC, internet industry action to create software mechanisms that opt-out users, and a program by the Digital Advertising Alliance to better notify users of their privacy options (FTC 2012). Mozilla, Microsoft, and Apple have created the biggest push in this area through adding functionality to their browsers that enables users to control the collection of their online browsing data. In the case of Microsoft, their latest version of Internet Explorer will have this feature enabled as the default, a decision that angered the advertising industry hoping to maintain a more notice-based system for Do Not Track (Singer 2012). What is needed to bring these acts of self-regulation into force would be the passing of legislation mandating both the inclusion of Do Not Track functionality in internet technologies and the compliance of these signaling mechanisms by companies like advertisers who collect this data. Bills along these lines have been proposed in the past two years but none have been passed, prominent examples being the *Commercial Privacy Bill of Rights Act of 2011* and the *Do-Not-Track Online Act of 2011*, both of which were referred to committees and never acted on.

The Do Not Track approach would greatly enhance online privacy protection in anticipation of social bots. However, this approach on its own is incomplete for the same reasons that Code as Law is incomplete as a framework—it discounts the importance of the socio-cultural context in which interactions and potential privacy invasions occur. A well-rounded approach must start with legislation enshrining a right to privacy on- and offline. Compliance with a legal right to privacy is not just about the technicalities of how data is collected and used but about social processes that value informed consent.

Legal scholars at the Haifa Center for Law & Technology propose a comprehensive and broadly applicable definition of the right to privacy that includes elements of access and control:

> The right to privacy is our right to keep a domain around us, which includes all those things that are part of us, such as our body, home, property, thoughts, feelings, secrets and identity. The right to privacy gives us the ability to choose which parts in this domain can be accessed by others, and to control the extent, manner and timing of the use of those parts we choose to disclose. (Onn et al. 2005, 12)

This definition or one based on it should be the centerpiece of legislation offering a right to privacy in the U.S. First, it's important to have a right to privacy in the U.S. to ensure

that the legal normativity of the right to privacy be established in the jurisdiction which is home to the first, and many of the most important, internet companies; it would also create greater regulatory harmonization with the EU and other foreign jurisdictions. Second, it's important to use a definition that addresses access and control to provide explicit, rights-based legal bases for future Do Not Track guidelines.

The robust set of Do Not Track guidelines should follow the principles articulated by the *vendor relationship management* (VRM) movement. VRM, in contrast to customer relationship management, is a model of interaction between consumers and companies that places the control over personal data in the hands of the consumer, including the right to correct data on record with a company and restrict use of that data at any time (McKay 2010). An ironic basis for VRM comes from the preamble to *The Cluetrain Manifesto*: "We are not seats or eyeballs or end users or consumers. We are human beings—and our reach exceeds our grasp. Deal with it" (Levine, Locke, Searls, and Weinberger 1999). Many of these principles are analogous to the Fair Information Practices (FIPs), which are guidelines for privacy protection measures put forward by regulatory bodies and other public institutions like the FTC and OECD, but which are not tied to specific values around privacy rights or how customers and businesses should interact.

The empowerment of the user over what data is stored and how it is used, with the ability to edit data about you, would be highly advantageous in addressing the problems arising from social bot conversations. Social conversations are fundamentally different from the types of communication that happen when inputting data into forms. Sarcasm, wordplay, and colloquial grammar are hard to discern to non-native speakers of a language, let alone artificial intelligence. With Do Not Track standards based on VRM, users could correct the mistakes on their own, otherwise permanent, online record. Next, the ability to later request that collected data not be used or be destroyed would allow social bot creators to collect data in cases where informed consent is unclear, but later give control over any stored personal data to the user. This would allow for innovation, like GOOG-411 for example, where individuals were able to call the 411 number and interact with a bot that conducted searches for information based on voiced requests. The data was used to train a speech recognition system, which Google later deployed as a new service (Manjoo 2011a). The history of user requests or any voice data that had been saved by Google after the training period (and possibly connected to users' personal phone numbers) should be able to be deleted at the request of the user. An upfront form of explicit consent would have made the service cumbersome; however, an *ex post facto* notification and data privacy control system would go a long way toward informed consent and honoring a right to privacy if properly designed.

For Do Not Track to work effectively while respecting a user's right to privacy, there must be significant coordination between the makers of web platforms and services, the makers of browsers and other internet infrastructure, and internet users—what David Clark et al. call "the tussle" (2005). This coordination is already happening at the browser level through consumer demand and at the web platform level through the *Request for Proposal* (RFP) process used by the World Wide Web Consortium (W3C), the forebear of which was the process that made the internet work in the first place by being just specific enough to achieve its goals while adaptable enough to foster prolonged

innovation, described by Clark as "rough consensus running code" (1992). This kind of rough consensus standard is also what effectively governs bot activity on Wikipedia. Wikipedia's bot policy is based on the norms and values of the community's stakeholders and are continuously defined and refined by users, then enforced by those same users through community policing (Wikipedia 2012). A commonly understood right to privacy implemented procedurally and technically through a Do Not Track RFP that acknowledges the social trust dilemma inherent to online transactions with social bots can help address the information asymmetries caused by a lack of informed consent through poor labeling and anthropomorphism. Compliance with agreed standards would be based on a combination of industry self-policing at the technological level, strict liability tort and criminal litigation for invasions of privacy, and regular privacy audits by the FTC in the style of the mandate on Google following the Google Buzz privacy debacle (Manjoo 2011b).

Possible technological guidelines and procedures that would comprise a comprehensive Do Not Track standard should include the following considerations at a minimum:

1) All browsers, including mobile, should be required to have a Do Not Track feature that the user would be prompted to turn on or off when opening the application—the Do Not Track status (on or off) would be prominently displayed on the interface by default. This would require users to make the conscious decision to turn off such protections and in so doing learn that they have such power over access to personal data transmitted via their browser. Compliance with the software signals (*http headers*) that the browser sends to websites and services that users visit would be required. Severe civil penalties would be levied in cases of noncompliance, similar to what has been proposed in earlier bills (*Do-Not-Track Online Act of 2011*).

2) Web service providers should write privacy policies and terms of service that use clear language, tested in user studies, rather than legalese in order to be effective as sources of informed consent (Pollach 2005). Additionally, these providers should alert users as to reduced functionality when interacting with a website, service, or social bot that needs additional personal data from the users, similar to how UK websites have been complying with the EU Cookie Policy and attempt to secure explicit consent (McMonagle 2012).

3) To enable innovation, specific types of information should not necessarily be specified as out of bounds for a company to come in contact with online, including currently regulated data such as genetic information ("Genetic Information Nondiscrimination Act of 2008"). However, any misuse of that data should still be susceptible to criminal prosecution with strict liability for corporations that employ social bots, despite the way that artificial intelligence might challenge "our traditional notions of intentionality" (Asaro 2012, 170).[3]

[3] Ambiguities over intentionality will likely always exist, which is why privacy protection will remain a subject of tort litigation as both a result of and as a part of regulation. Specifically, defining rules around how to assign liability for data collection by social bots may be challenging due to the complexity inherent to designing bots and the platforms on which they perform their actions, either of which may or may not be produced by the same individual or consortium of firms. While liability should be strict for misuses of data, problems around invasions of privacy at the collection level will probably need to be decided on a case-by-case basis to determine if liability should be applied severally or strictly.

## How we got into this Mess (Barriers to Reform)

> The development of AI is a business, and businesses are notoriously uninterested in fundamental safeguards—*especially* philosophic ones. (A few quick examples: the tobacco industry, the automotive industry, the nuclear industry. Not one of these has said from the outset that fundamental safeguards are necessary, every one of them has resisted externally imposed safeguards, and none have accepted an absolute edict against ever causing harm to humans.) (Science Fiction writer Robert J. Sawyer 1991)

The goal of good regulation should be to invoke the *precautionary principle* when warranted and otherwise wait on *proof before action*. Although social bots may not seem to invoke the precautionary principle in the same way an environmental health risk might, the unique risks they pose to personal privacy may dramatically increase the harms that follow as a result of existing under-regulation of online privacy protection.

Recent legal scholarship (Solove 2008; Lipton 2010) has tried to reconceptualize privacy in light of issues posed by the internet, moving away from a focus on "identifying and compensating harms that can be economically quantified" (Lipton 2010, 486), such as damages to career prospects as result of disclosure. Limited privacy rights in the U.S. may be the result of regulatory capture and lobbying (Rotenberg 2001, II.C.49), most prominently undertaken by the advertising industry, which has contributed to the lack of relevancy and success that laws like the *Electronic Communications Privacy Act of 1986* and the *Computer Fraud and Abuse Act of 1986* have had in protecting privacy online—particularly when it comes to overly broad requests for data from law enforcement (Helft and Miller 2011) and companies tracking users against their will (Wolk, McGaraghan, and Hantover 2012).

The key concept in online privacy protection law is FIPs. And the key difference between the U.S. and the rest of the world is how these have been applied. European privacy protection is based on an explicit and general *right to privacy* (*Convention for the Protection of Human Rights and Fundamental Freedoms* 1950, Article 8). FIPs date back to U.S. privacy reports from 1973 (Gellman 2012), and have been a through line of most privacy regulation since then, although with important modifications over time. One important modification was in the FTC's 2000 report on privacy, which specified that corporations collecting personal information must abide by a "notice and choice" principle that allows for implicit consent, which was a change from previous principles endorsing more of a "notice and consent" principle that requires explicit consent. Rotenberg (2001) argues that this was a case of a regulatory capture by the Direct Marketing Association in 1996 to create an industry-wide standard of privacy protection in which implicit consent is the default that consumers must opt-out of. The FTC has since revised the FIPs to collapse both principles into "choice/consent." However this "principle" still offers corporations the option of simply pursuing the "choice" approach, i.e. default opt-in or opt-out, which inevitably leads to most privacy policies being opt-out.

This problem of regulatory capture and lowest common denominator compliance in the U.S. is in part because of the lack of a right to privacy. This is in contrast to EU

---

And as Asaro (2012) points out, when bots achieve near-sentient levels of artificial intelligence and autonomy, it may not even be clear that their creators should be held fully liable if they go rogue.

regulation, which is based on a more comprehensive version of FIPs—the *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*, which recognize EU language on the right to privacy and endorse approaches that strive for informed consent through "knowledge" or explicit consent for both data collection and disclosure (1980).[4]

Unfortunately, the goal of passing right to privacy legislation is the most fundamental barrier to implementing policy reform. As mentioned earlier, two Do Not Track bills from 2011 were killed by inaction in committee. The key reasons for this are not clear but cases like this are generally the result of inertia created by the other barriers to reform: economic ideology, coordination, regulatory capture, and timing. Passing a "right to privacy" bill is further complicated by the inertia of decades of jurisprudence treating privacy otherwise. Influential voices like Judge Richard Posner have argued on economic grounds that there is no need for a right to privacy (1981), believing case law to be sufficient, in particular the traditional definitions of privacy invasion for modern U.S. tort law:

1. Intrusion upon the plaintiff's seclusion or solitude, or into his private affairs.
2. Public disclosure of embarrassing private facts about the plaintiff.
3. Publicity which places the plaintiff in a false light in the public eye.
4. Appropriation, for the defendant's advantage, of the plaintiff's name or likeness. (Prosser 1960, 389)

Such cases of a data privacy tort are based on the legal test of the *reasonable expectation of privacy*—defined as either the subjective opinion of a plaintive that holds a situation as private or that which is commonly accepted by society as private, as determined by judge or jury—in contrast with the *right to privacy*, typical in Europe as the explicit "right to respect for private and family life" (Council of Europe 1950).

However, in more recent years, Congress has see it fit to extend a right to privacy over personal data in specific sectors: health information ("Health Insurance Portability and Accountability Act of 1996" or HIPAA), children age 13 and under ("Children's Online Privacy Protection Act of 1998"), and consumer financial information ("Gramm-Leach-Bliley Act of 1999"). Building off these special cases and invoking the precautionary principle in light of risks posed by social bots could overcome some inertia. HIPAA requires minimal disclosure of information to achieve objectives, a right to request that inaccurate information be corrected, reasonable steps be taken to maintain confidentiality of communications, that individuals must be notified of how their information is used, and all disclosures be tracked and privacy policies and procedures documented. This is an ideal precedent for a right to privacy and Do Not Track law, which could then ease the way for new comprehensive legislation.

Ideological opposition, particularly on economic grounds, still poses a significant barrier. Many members of Congress share Posner's opinion that the need for a right to privacy is not justified economically. Certainly, command and control regulation on privacy could have adverse outcomes on advertising-driven businesses (Tucker 2011).

[4] While the United States is a member of the OECD, the Guidelines are non-binding on member countries. U.S. laws continue to be based on the FTC's FIPs, which do not recognize an explicit right to privacy or require explicit consent.

Critics have already lined up with their evidence and made open calls for Congress to require cost-benefit analysis (CBA) for all internet regulation (Szoka and Sperry 2012).

One counterargument to these claims may be to attack the conventional view of protection through torts litigation and the incentives of monetary compensation for harms. If we look through a rights-based lens then the validation of regulation by quantifying harm is irrelevant. You cannot quantify the *right* to privacy like a private good in a marketplace because it is a "social state" (Sen 2000).

Another counterargument takes CBA on its own turf of justifying policy in terms of utility. On the one hand, online privacy policy is based on a utilitarian perspective—often justified via CBA—whereby the use of data by a company is measured according to its ability to create the most good for the greatest number of people. Positive externalities like the network effects of online social spaces enhance utility as data is pooled on common platforms like Facebook. It may be possible to illustrate that social bots when properly regulated with Do Not Track standards produce more utility than those that are unregulated, suffering from negative externalities as greater numbers of consumers are affected by privacy invasion. Certainly, prototypes in the form of Twitter bots have been shown to be capable of connecting disparate networks around common interests—a potential public good when privacy is protected (Nanis, Pearce, and Hwang 2011).

A third counterargument might be based on VRM as a driver of new economic potential around personal data. VRM maintains a copyright-like view of information as a tradable, rivalrous good rather than as a non-rival good. If personal information is exclusively owned by the person and can be traded, the right to use personal information doesn't need to be permanently transferred to other actors like social bots. Continued access to personal data could be given on a limited license rather than with full rights to data collectors; an example of how this might be manifested is a single loyalty card, which you use to give your limited loyalty and personal data to whichever store you want to buy from at that point in time—rather than the current case of being a "member" of each store (McKay 2011). This creates a market for personal data, which may stimulate innovation through competition for the license to use personal data and the benefit of continued patronage.

A final counterargument to economy-centric ideology might come from applying the *Porter Hypothesis*, which predicts that a clear set of guidelines for companies to follow, especially if they are harmonized across jurisdictions, may perhaps lead to greater innovation in services like social bots (Porter and van der Linde 1995). However, the Porter Hypothesis depends on coordination between governments, and among industries and industry actors. Laissez faire is sometimes assumed to be the soundest policy not only from an ideological belief in perfect market efficiency, but because the government is "bad" at picking winners and losers, and any proposed regulation may fail because of interoperability issues from top-down standards or simply by regulatory capture.

History suggests that regulatory capture will be a barrier, and that advertising firms will be on the vanguard of lobbying the government to kill the privacy reform legislation, invoking creative knowledge assessment to argue for proof before action against the "unproven risks of social bots" or trumpeting inflated economic arguments. The Direct Marketing Association already influenced privacy policy once in the U.S., and

their competing definition of Do Not Track policy based on a self-regulated opt-out system with new labels—called the Digital Advertising Alliance—is a potential barrier to stronger regulation passing Congress and enforced by the FTC if it's deemed sufficient by politicians and officials (Singer 2012). The ACLU's Jay Stanley argues that stronger privacy fences around personal data are needed sooner rather than later precisely because of incipient revenue streams based on that data, once "companies begin to make hay from those areas that should be forbidden, [...] it becomes much harder to get them to stop, because they can deploy the money they make there to stop rules and protections from being erected" (2012).

Advertising is not the only industry that has an interest in weak privacy regulation. Many corporations rely on user data to improve their products and services. By requiring data collection to be opt-in there is a fear in the industry that they will not be able to collect enough data to fuel their internal innovation processes. And some internet services only work when personal data is available, including location-based services. The burden of regulation can and does occasionally impede innovation, and the U.S. should be particularly careful when regulating industries such as artificial intelligence and internet commerce in which it enjoys a market advantage with national strategic implications in terms of both economics and security. Attempts at standardization and harmonization can sometimes be inappropriate depending on the maturity of a technology. However, bottom-up standards development can create more robust standards, with higher levels of compliance, thanks to buy-in to the process by stakeholders who want to enjoy the positive externalities of interoperability. There is reason to believe that online privacy is on the cusp of this due to the previously mentioned evidence of early movement by internet industry actors to enact their own Do Not Track standards. Hopefully, this opens the door for the federal government to support such efforts by passing "right to privacy" legislation and working with industry to find a mutually beneficial set of Do Not Track standards.

A lot of policy reform depends on "timing." If there is insufficient concern by the public around social bots and privacy, it may be hard to trigger the precautionary principle against Congress's anti-regularity inertia. There may be a lack of evidentiary experience with social bots and their effects on online privacy at this time. However, the initiative taken by industry itself to implement Do Not Track may represent good timing on balance, offering Congress an opportunity for easier passage and compliance of new legislation. Conversely, members of Congress may perceive the same initiative as the sufficiency of self-regulation, negating the need to guarantee a right to privacy and mandate compliance.

Strong support for policy reform could also result in poor timing for policy reform. For instance, at the writing of this paper, U.S. national attention has been captured by online privacy protection issues raised by the investigation of former CIA Director David Petraeus's indiscretions over email (Romm and Byers 2012). If a moral panic is created around such an issue, it may lead to over-regulation by Congress, or rushed and poorly considered legislation, which in the case of social bots might mandate too many or the wrong kind of standards on industry and jeopardize coordination and compliance, interoperability, or innovation.

## Conclusion

Science fiction has provided its readers with much to think about concerning the possibilities and pitfalls of new technologies. Who knows? Social bots may be "life-savers" for a few power users. But as we benefit from new technology, someone needs to consider the costs. Take for example another emerging set of technologies called the "internet of things," which may pose further unprecedented risk to privacy. The source of the problem lies in the science fiction-like ubiquity of internet-connected sensors to be built into every object we use. Objects like coffee cups will initially collect data through their sensors for specific purposes—perhaps to understand better how you like your coffee. However, if data is comprehensively collected and stored, it may provide levels of privacy invasion on par or greater than the "vicious cycle" case of social bots. The key difference is that the internet of things will be in the form of seemingly innocuous appliances that inhabit not the virtual but the real world, specifically our homes—what have traditionally been considered our "private" spaces and spheres.

This issue of surveillance demands its own separate research papers and policy details, but I want to mention it briefly here because the relatedness among social bots, the internet of things, and their parallels to *1984*'s "Big Brother" underscores the importance of the precautionary principle and reform. The U.S. may not have a right to privacy in a general sense, but it does have one with respect to the government thanks to the Fourth Amendment and the "Privacy Law of 1974," and its amendments. However, the Fourth Amendment took a battering at the hands of *The Patriot Act* in its loosening of restrictions on the government's ability to access personal data (Smith et al. 2002), which was combined with a continued vagueness in current privacy protections around electronic communications as described earlier. As the volume and detail of personal data increases dramatically—accumulated automatically through web services, social bots, and the internet of things—the risk of that data being used against users rises along with the temptation of the state to pursue such valuable data. A few internet companies, notably Google (*Google Transparency Report* 2012) and Twitter ("Twitter Transparency Report" 2012), have publicly discussed the enormous number of requests they receive from officials at all levels of government, asking for huge chunks of data that go well beyond what should be narrow investigations.

My hope is that enshrining privacy as a right and enforcing Do Not Track will not only address the privacy problems tied to transactions like targeted advertising but also help reinforce the Fourth Amendment. And I strongly encourage a continued research agenda of technology and policy analysis to consider more threats to privacy posed by innovations like the internet of things that are sure to follow social bots in the never too distant future.

## References


Asaro, Peter M. 2012. "A Body to Kick, but Still No Soul to Damn: Legal Perspectives on Robots." In *Robot Ethics: The Ethical and Social Implications of Robots*, edited by Patrick Lin, Keith Abney, and George A. Bekey, 169–186. Cambridge: The MIT Press.

Burger, David. 2009. "Ethics of Robot Behaviour." *IEEE Global History Network*. Accessed November 16, 2012: http://www.ieeeghn.org/wiki/index.php?title=Ethics_of_Robot_Behaviour&oldid=21293.

"Children's Online Privacy Protection Act of 1998." Title 15 *U.S. Code*, Pts 6501–6505. Accessed November 16, 2012: http://www.ftc.gov/ogc/coppa1.htm.

*Commercial Privacy Bill of Rights Act of 2011*. S.799. 112th Cong. 1st sess. Accessed November 16, 2012: http://epic.org/privacy/consumer/Commercial_Privacy_Bill_of_Rights_Text.pdf.

Council of Europe. 1950 (November 4). *Convention for the Protection of Human Rights and Fundamental Freedoms*. Accessed November 16, 2012: http://conventions.coe.int/treaty/en/Treaties/Html/005.htm.

boyd, danah. 2010 (May 15). "Facebook is a utility; utilities get regulated." *apophenia*. Accessed November 16, 2012: http://www.zephoria.org/thoughts/archives/2010/05/15/facebook-is-a-utility-utilities-get-regulated.html.

Clark, David D. 1992. "A Cloudy Crystal Ball: Visions of the Future." At *24th meeting of the Internet Engineering Task Force*, Cambridge, July 13–17. Accessed November 27, 2012: http://www.ietf.org/old/2009/proceedings/prior29/IETF24.pdf.

Clark, David D., Wroclawski, John, Sollins, Karen R., and Robert Braden. 2005. "Tussle in Cyberspace: Defining Tomorrow's Internet." *IEEE/ACM Transactions on Networking* 13, no. 3: 462–475.

Darling, Kate. 2012. "Extending Legal Rights to Social Robots." Presentation at We Robot Conference, University of Miami, April 21–22. Accessed November 16, 2012: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2044797.

*Do-Not-Track Online Act of 2011*. S.913. 112th Cong. 1st sess. Accessed November 16, 2012: http://www.gpo.gov/fdsys/pkg/BILLS-112s913is/pdf/BILLS-112s913is.pdf.

Duffy, Brian R. 2003. "Anthropomorphism and the Social Robot." *Robotics and Autonomous Systems* 42, no. 3–4: 177–190.

Federal Trade Commission. 2012. *Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policymakers*, http://www.ftc.gov/os/2012/03/120326privacyreport.pdf.

Gellman, Robert. 2012 (November 12). "Fair Information Practices: A Basic History." Version 1.91. *bobgellman.com*. Accessed November 16, 2012: http://bobgellman.com/rg-docs/rg-FIPShistory.pdf.

"Genetic Information Nondiscrimination Act of 2008." P.L. 110-233. *United States Statutes at Large*. 122 Stat. 881. Accessed November 16, 2012: http://www.govtrack.us/congress/bills/110/hr493/text.

*Google Transparency Report*. 2012. Accessed November 16, 2012: http://www.google.com/transparencyreport/.

"Gramm-Leach-Bliley Act." P.L. 106-102. *United States Statutes at Large*. 38 Stat. 1338. Accessed November 16, 2012: http://www.gpo.gov/fdsys/pkg/PLAW-106publ102/html/PLAW-106publ102.htm.

"Health Insurance Portability and Accountability Act of 1996." P.L. 104-191. *United States Statutes at Large*. 110 Stat.1936. Accessed November 16, 2012: http://www.gpo.gov/fdsys/pkg/PLAW-104publ191/html/PLAW-104publ191.htm.

Helft, Miguel, and Claire Cain Miller. 2012 (January 9). "1986 Privacy Law Is Outrun by the Web." *The New York Times*. Accessed November 16, 2012: http://www.nytimes.com/2011/01/10/technology/10privacy.html.

Lessig, Lawrence. 1999. *Code and Other Laws of Cyberspace*. New York: Basic Books.

Levine, Rick, Locke, Christopher, Searls, Doc, and David Weinberger. 1999. *The Cluetrain Manifesto*. Accessed November 27, 2012: http://www.cluetrain.com/.

Lipton, Jacqueline. 2010. "Mapping Online Privacy." *Northwestern University Law Review* 104, no. 2: 477–515.

Lotan, Gilad, Graeff, Erhardt, Ananny, Mike, Gaffney, Devin, Pearce, Ian & danah boyd. 2011. 'The Revolutions Were Tweeted: Information Flows during the 2011 Tunisian and Egyptian Revolutions.' *International Journal of Communication*, 5.

Manjoo, Farhad. 2011a (April 6). "Now You're Talking!" *Slate*. Accessed November 27, 2012: http://www.slate.com/articles/technology/technology/2011/04/now_youre_talking.html.

Manjoo, Farhad. 2011b (April 7). "No More Privacy Paranoia." *Slate*. Accessed November 27, 2012: http://www.slate.com/articles/technology/technology/2011/04/no_more_privacy_paranoia.html.

Marcus, Gary. 2012 (November 27). "Moral Machines." *News Desk* (*The New Yorker*). Accessed 27, 2012: http://www.newyorker.com/online/blogs/newsdesk/2012/11/google-driverless-car-morality.html.

McKay, Lauren. 2010 (May). "It's Not Your Relationship to Manage." *CRM Magazine*. Accessed November 27, 2012: http://www.destinationcrm.com/Articles/Editorial/Magazine-Features/Its-Not-Your-Relationship-to-Manage-66870.aspx.

McMillan, Robert. 2012. "Man builds Twitter Bot that Humans Actually Like." *Wired*. Accessed November 16, 2012: http://pacsocial.com/files/pacsocial_field_test_report_2011-11-15.pdf.

McMonagle, Christie. 2012 (May 28). "Complying or Not? Approaches to the EU cookie directive." *The Attacat Brain*. Accessed November 16, 2012: http://www.attacat.co.uk/brain/16-brands-approach-cookie-directive.

Nanis, Max, Pearce, Ian, and Tim Hwang. 2011. “PacSocial: Field Report Test Report.” *PacSocial.com*. Accessed November 16, 2012: http://pacsocial.com/files/pacsocial_field_test_report_2011-11-15.pdf.

*OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*. 1980. Accessed November 16, 2012: http://www.oecd.org/internet/interneteconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm.

Onn, Yael, et al. 2005. *Privacy in the Digital Environment*. Ed. Niva Elkin-Koren and Michael Birnhack. The Haifa Center of Law and Technology Publication Series, no. 7. Accessed November 27, 2012: http://books.google.com/books?id=yeVRrrJw-zAC.

Orcutt, Mike. 2012 (June 21). “Twitter Mischief Plagues Mexico’s Election.” *MIT Technology Review*. Accessed November 27, 2012: http://www.technologyreview.com/news/428286/twitter-mischief-plagues-mexicos-election/.

Plitch, Phyllis. 2002 (September 16). “Are bots legal?” *The Wall Street Journal*. Accessed November 27, 2012: http://online.wsj.com/article/0,,SB1031785433448138595-search,00.html.

Pollach, Irene. 2005. “A Typology of Communicative Strategies in Online Privacy Policies: Ethics, Power and Informed Consent.” *Journal of Business Ethics* 62, no. 3: 221–235.

Porter, Michael, and Claas van der Linde. 1995. “Toward a New Conception of the Environment-Competitiveness Relationship.” *Journal of Economic Perspective* 9, no. 4: 97–118.

Posner, Richard A. 1981. *The Economics of Justice*. Cambridge: Harvard University Press.

“Privacy Act of 1974.” P.L. 93-579. *United States Statutes at Large*. 88 Stat. 1896. Accessed November 16, 2012: http://www.llsdc.org/attachments/wysiwyg/544/PL093-579.pdf.

“Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the protection of individuals with regard to the processing of personal data by competent authorities for the purposes of prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and the free movement of such data,” COM (2012) 10. *EUR-Lex*. Accessed November 16, 2012: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52012PC0010:en:NOT.

Prosser, William L. 1960. “Privacy.” *California Law Review* 48, no. 3: 383–423.

Romm, Tony, and Alex Byers. 2012 (November 16). “David Petraeus affair scandal highlights email privacy issues.” *Politico*. Accessed on November 27, 2012: http://www.politico.com/news/stories/1112/83984.html.

Rotenberg, Marc. 2001. “Fair Information Practices and the Architecture of Privacy: (What Larry Doesn’t Get).” *Stanford Technology Law Review*. Rev. 1. Accessed November 16, 2012: http://stlr.stanford.edu/STLR/Articles/01_STLR_1.

Sawyer, Robert J. 1991 (1994). “On Asimov’s Three Laws of Robotics.” *Science Fiction Writer*. Accessed November 16, 2012: http://www.sfwriter.com/rmasilaw.htm.

Sawyer, Robert J. 2007. “Robot Ethics.” *Science* 318, no. 5853: 1037.
Sen, Amartya. 2000. “The Discipline of Cost-Benefit Analysis.” *Journal of Legal Studies* 29, no. S2: 931–952.
Siegler, MG. 2010 (March 15). ‘Ev Williams: Twitter's First Principle, “Be A Force For Good.”’ *TechCrunch*. Accessed November 27, 2012: http://techcrunch.com/2010/03/15/ev-williams-sxsw/.
Singer, Natasha. 2012 (October 13). “Do Not Track? Advertisers Say ‘Don’t Tread on Us’.” *The New York Times*. Accessed November 13, 2012: www.nytimes.com/2012/10/14/technology/do-not-track-movement-is-drawing-advertisers-fire.html.
Smith, Marcia S., Seifert, Jeffery W., McLoughlin, Glenn J., and John Dimitri Moteff. 2002. *The Internet and the USA PATRIOT Act: Potential Implications for Electronic Privacy, Security, Commerce, and Government*. CRS Report for Congress. Accessed November 16, 2012: http://epic.org/privacy/terrorism/usapatriot/RL31289.pdf.
Solove, Daniel. 2008. *Understanding Privacy*. Cambridge: Harvard University Press.
Stanley, Jay. 2012 (October 17). “Protections Against Commercial Internet Spying: Why Delay is Deadly.” *Free Future*. Accessed November 29, 2012: http://www.aclu.org/blog/technology-and-liberty/protections-against-commercial-internet-spying-why-delay-deadly.
Steiner, Peter. 1993 (July 5). “On the Internet, nobody knows you’re a dog.” *The New Yorker* 69, no. 20: 61.
Szoka, Berin, and Ben Sperry. 2012 (November 16). “Congress Delays Requiring Cost-Benefit Analysis of Internet Regulation.” *TechFreedom*. Accessed November 16, 2012: http://techfreedom.org/publications/congress-delays-requiring-cost-benefit-analysis-internet-regulation.
Tickle, Glen. 2012 (November 23). “Facebook Announces New Privacy Changes That Can Lead to External Advertising.” *Geekosystem*. Accessed November 27, 2012: http://www.geekosystem.com/facebook-external-advertising/.
Tucker, Catherine. 2011. *Empirical Research on the Economic Effects of Privacy Regulation*, http://cetucker.scripts.mit.edu/docs/law_summary_2011.pdf.
“Twitter Transparency Report.” 2012. *Twitter Help Center*. Accessed November 16, 2012: https://support.twitter.com/groups/33-report-abuse-or-policy-violations/topics/148-policy-information/articles/20170002-twitter-transparency-report#.
Veruggio, Gianmarco. 2006. “The EURON Roboethics Roadmap.” Presentation at IEEE-RAS International Conference on Humanoid Robots, Genova, Italy, December 4–6. Accessed November 16, 2012: http://www.nd.edu/~rbarger/ethics-roadmap.pdf.
Wikipedia. 2012 (November 15). “Wikipedia:Bot policy.” *Wikipedia*. Accessed November 27, 2012: http://en.wikipedia.org/w/index.php?title=Wikipedia:Bot_policy&oldid=523208915.

Winner, Langdon. 1986. “Do Artifacts have Politics?” In *The Whale and the Reactor: A Search for Limits in an Age of High Technology*, 19–39. Chicago: University of Chicago Press.

Wolk, Michael, McGaraghan, John, and Lixian Hantover. 2012. “Dismissal of Amazon Privacy Suit Signals Uphill Battle for Plaintiffs’ CFAA Claims.” *Eye on Privacy*, July 2012. Accessed November 16, 2012: http://www.wsgr.com/publications/PDFSearch/eye-on-privacy/July2012/index.html#3.

Wu, Tim. n.d. “Net Neutrality FAQ.” *timwu.org*. Accessed November 27, 2012: http://timwu.org/network_neutrality.html.

Yudkowsky, Eliezer S. 2001 (May 3). “What is Friendly AI?” *KurzweilAI*. Accessed November 27, 2012: http://www.kurzweilai.net/what-is-friendly-ai.