AI and the Ethics of Automating Consent

Meg Leta Jones, Ellen Kaufman, and Elizabeth Edenberg | Georgetown University

While supplementing consent with further mechanization, digitization, and intelligence—either through proffering notification on behalf of a consentee or choosing and communicating consent by the consenter—may improve take-it-or-leave-it notice and choice consent regimes, the goal for AI consent should be one of partnership development between parties, built on responsive design and continual consent.

Enthusiasm for artificial intelligence (AI) ethics is apparent from large funding gifts, conference themes, popular media coverage, and IEEE and ACM initiatives. In the call for this special issue of IEEE Security & Privacy on AI ethics, interdisciplinary submissions were encouraged from law and policy, computer science and engineering, social studies, and economics. Missing is perhaps ethics and moral philosophy. In an effort to explicitly integrate ethics and moral philosophy into this collection with the other relevant disciplines, we intend to take the opportunity new technologies often offer—to go back and reassess some foundational moral concepts. As such, we will tackle the complex issue of consent and AI.

Consent has been a fundamental aspect of ethics for centuries. As a moral concept, consent is significant because it plays a morally transformative role in interpersonal interactions. Valid consent can render permissible an otherwise impermissible action. It transforms the specific relations between the consenter and consentee about a clearly defined action—we can consent to sexual relations, borrowing a car, surgery, and the use of personal information. Without consent, the same acts become sexual assault, theft, battery, and an invasion of privacy. Thus, the conditions of consent involve two actors (an agent able to give consent—the consenter—and an agent seeking permission to do something with no moral right or entitlement to do so otherwise—the consentee) and a circumstance in which moral transformation is necessary. When consent is communicated under such conditions, a moral transformation occurs.

In the 20th century, control dominated the meaning and effectuation of privacy, and consent became central to this conception of privacy. Having played various roles in the development of data protection regimes worldwide in 1970s and 1980s, consent has been required (sometimes and in some places) for personal information in databases, international transfer, tracking, and profiling. As transparency is foundational to consent, digital information exchange is based on notice and choice wherein a user is notified of information collected and used and provided a choice of whether to continue with, object to, or explicitly accept (depending on the circumstances and jurisdiction) the information exchange. Although consent is central to data protection and privacy laws based on a notice and choice regime around the world, it is worth (re)questioning whether, why, and when consent is required in AI systems. This is particularly true both because it is common to perpetuate consent even when no authorization is necessary and because consent is such a contested aspect of global data protection and privacy policy.
After providing a background on the ethics of consent and its role in modern privacy law, we apply these principles to AI as both a system to which consent is solicited as well as a system that generates consent. AI systems collect, process, and generate data in ways that further exacerbate many existing problems with online consent, most notably issues of providing adequate notice, choice, and withdrawal to users. The unpredictable and even unimaginable use of data by AI systems is considered a feature, not a bug. Yet this feature creates problems for notifying users as well as assessing when consent might be required based on potential uses, harms, and consequences. We discuss whether these problems impact morally transformative consent in AI systems. Assuming consent is required for AI systems to collect and use personal information, one popular solution to the problems of obtaining consent in digital environments is to use AI to predict what information practices a user would consent to and have such preferences signaled to smart systems attempting to collect or use data about the user. Such signaling requires a great deal of personal information. We will discuss whether, and if so, how automatically generated consent could meet the moral demands of ethical consent.

Consent Background

In the broadest sense, an individual’s consent involves an effective communication of an intentional transfer of rights and obligations between parties. Consent thereby transforms the moral landscape between two parties, rendering permissible otherwise impermissible actions. Given its morally transformative power, consent generally requires a clearly defined scope of action that an individual gives permission for another person (or group) to do to that person. An individual must also have the relevant information so she knows what she consents to, and she must be free to choose (or refrain from) the transaction.

In biomedical ethics, informed consent is sought to ensure that patients are respected as persons and can make informed choices about what is done with and to their bodies. The Belmont Report was a reaction against the paternalistic practices that frequently dominated doctor–patient interactions and the many abuses of people as research subjects (often without their knowledge that they were participating in research at all). In reaction to these practices, the goal of establishing informed consent requirements was to protect an individual’s ability to exercise control and decisional power over her body. Given the medical context, the focus on information was important because of the expected gap in knowledge about the relevant research or medical treatment. Healthcare professionals are expected to have a domain of expertise beyond the patients or potential research subjects, and requirements seeking to protect patients’ decisional liberty require that they understand the relevant choice situation. Often, this requires clear disclosure and explanation of information relevant to the medical decision at hand.

Informed consent requirements that are codified in law, however, frequently fail to live up to the underlying moral value that justified their creation. The practice of obtaining informed consent is justified on the grounds of respecting a patient’s autonomy or, at least minimally, her control over authorizing what is done to and with her body. Too frequently, however, medical practitioners give patients an “informed consent” disclosure sheet that must be signed before receiving any kind of treatment. Patients either face the choice of no treatment or sign away a wide scope of uses of their personal information, tissue, waive liabilities, and so forth. In these cases, a gap opens up between legally valid consent and morally transformative consent.

Similarly, digital consent in the form of “consent” to terms of service fails to do more than disclose the wide range of potential uses the company has for an individual’s personal data. While this may be legally valid consent, it may fail to live up to consent’s moral goal. Digital consent has been repeatedly criticized as being an unworkable, empty procedural act. Efforts to create tools to control personal information and improve digital consent have remained the focus of those who continue to resist paternalistic privacy regulations, promote the autonomous market participant, or believe no other form of protection is likely to surface.

Elsewhere in bioethics, the concept of consent was refined to include two relevant notions: continual consent and cooperative consent. Miller and Wertheimer, for example, highlight the importance of consent being a cooperative venture in which there are fair and predictable standards underlying the consent transaction for both the consentor and consentee. They emphasize the bilateral nature of the consent transaction that gives due weight to the interests of both parties. This is meant to capture both that the consenter should be able to make an informed decision that appropriately captures
her aims, and that the consentee needs clear standards for determining when it is permissible to proceed. The consentee is not always (nor is often) likely to be in a position to determine whether the token of consent was autonomously authorized. Far more likely, they argue, is the ability to determine the fair context that is the necessary precondition for cooperative interactions.

Related to the idea that consent is a transaction that involves the cooperation and effective communication of the transference of rights and obligations between two parties is the idea that this transaction is not always a one-off situation. Obtaining consent for a specific action at a moment in time is one issue. A single authorization of the transaction is sufficient. But what if the transaction extends over weeks, months, or years? In these cases, it seems that the consentor should retain the right to withdraw her consent at any time. When the transaction occurs over an extended period of time, consent may need to be obtained again for continued use, and the party who authorized the action initially has a right to change her mind.

These modifications of consent in bioethics have been offered as ways of adjusting consent to better fit the cooperative context in which transactions can extend beyond a single interaction. These tools may prove useful for reimagining consent for AI systems. Digital consent may be reimagined for AI systems as necessarily continual and aspirationally cooperative. We use these two concepts to scrutinize progress and potential for AI consent as it is currently developing.

**Digital Consent**

Morally binding digital consent is difficult to obtain, but that does not make it any less necessary. It may be that robust digital consent cannot realistically be achieved, and thus the action is prohibited or otherwise limited and heavily regulated. Many of today’s transnational debates about privacy hinge on how political cultures view this controversy. Online, notice is presented on a page users can find by clicking a “privacy policy” link usually placed at the bottom of each page or within app settings. Consent is sometimes signified by clicking an “OK” box in a cookie banner or settings pop-up (express action) or more commonly remaining on the site without leaving or changing settings (omission). The problems with relying on this form of consent to alter rights and liberties associated with personal information have been thoroughly investigated by privacy scholars for the past two decades. This form of notice and choice may serve as legally binding consent, but may still fall short of morally binding consent.

Solove\(^3\) succinctly describes the challenges of managing one’s digital information using a consent model. First, individuals cannot possibly read all the terms of services and privacy policies they are exposed to every day. Second, notification language is very hard to write clearly and accurately so users can really understand what information processing and exchanging may happen. Finally, even if individuals could read and understand all of the policies, taking action and assessing risks and harms based on the notices in the policies is often futile or impossible. The prevailing approach to privacy attempts to protect individual choice by remaining neutral on the types of policies accepted. Solove refers to this as “privacy self-management.” However, consent to the collection, use, and disclosure of personal data is often not meaningful in the current system.

Nissenbaum and Barocas argue that even if informed consent were achievable, it would not be effective against contemporary information harms because modern data practices revolve around future and unanticipated uses.\(^2\) Some, including Nissenbaum and Barocas, have turned to a more paternalistic approach in which the law restricts certain permissible uses by data controllers. Limiting what one can consent to is argued for by a number of privacy researchers including Schermer, Custers, and van der Hof, who explain that consent overload, information overload, and the absence of meaningful choice leads to “consent desensitization,” wherein users are consenting to risks they do not want to take on and controllers are relying on an unstable form of compliance.\(^4\) They call for a set of commonly held “fair use” data practices that do not require consent. Solove points out that this leads to a puzzling result of trying to secure individual autonomy by limiting an individual’s control over what can be done with her data.

Much of the scholarly energy exerted by privacy researchers went to consent during the early 2000s, but the conversation has evolved from one focused on individual control to one of power and harms. Nonetheless, consent has remained contentiously central to data protection ethics, policy, and international debate from the early days of digital computing through the commercialization of the web and will further complicate the integration of AI systems and the Internet of Things (IoT).

AI’s exacerbation of the consent dilemma offers an opportunity for us to reconsider consent anew. In an AI consent model, the data subject continues to be the consenter and a similar range of controllers could be the consentee. The course of action between the parties dictates whether consent is required. Similar to digital consent, the action takes place within a particular context, which may relieve consent requirements by providing justification to the would-be consentee. Also similar to digital consent, contextual integrity may not morally overcome the overwhelming risks and unknowns associated with AI data collection and processing, in which case consent supplementation (either explicit consent or other regulatory means) may be justified.
However, AI does present three special problems for a notice and choice consent model. First, AI systems are defined by their unpredictability and opacity. These are considered features, not bugs, but they only serve to complicate the existing problems with digital consent by drawing unforeseeable connections between information, generating novel uses, and mounting challenges to explainability. Second, AI systems in social settings can induce personal information from individuals in unexpected and even manipulative ways. Finally and relatedly, AI plays an important role in the integration of the IoT and future smart environments, wherein connected objects, people, and spaces not only challenge the screen-based form of notification but also present novel challenges because users may not have a direct relationship with the systems collecting and processing their information (a problem referred to as the “Internet of other people’s things”).

Continual (as opposed to conclusive) consent, which includes the ability to withdraw consent, is perhaps even more justified for emergent AI behavior than run-of-the-mill digital data processing. Tools that support continual consent are often dashboards provided by the platform itself, such as Facebook settings and the Google privacy dashboard. However, tools to manage one’s data once it has been given to a third party are usually limited to removing information from public people searches with free opt-out procedures. Thus, the most effective privacy tools have been those that signify a rejection of consent through blocking as much data collection as possible. A new set of management concepts and tools have been presented in recent years that may further empower individuals to consent and so increase consent’s popularity as a governance tool. Our analysis below suggests that these tools may improve control and notification but do little to establish the cooperative partnership consent can build between data controllers and subjects, or address the problem of creating an information infrastructure to support continual consent.

**Improving Consent**

Information scientists working to improve digital privacy, known now as privacy engineers, have developed a number of systems to improve the notice and choice consent regime. One of the earliest systems designed to enhance privacy self-management is also one of the most prescient. Decades before privacy scholars considered artificial intelligence as a solution for managing user privacy preferences, there was the Platform for Privacy Preferences (P3P), an initiative spearheaded in the mid-1990s by the Internet Privacy Working Group. P3P was designed “to let people define privacy preferences for themselves and embed those preferences in a browser or software agent that can communicate directly with Websites and other electronic partners.”

Fundamentally, P3P addressed the need for more standardized and cogent notice, hypothetically allowing data subjects to understand clearly and succinctly how the data practices of a given website both gelled with their preferences and compared with the practices of other websites. In its initial version, the system also included a feature that would have acted as a privacy gatekeeper, negotiating on behalf of the user about what content and websites they could access according to their preferences. The initial abandonment of this feature speaks to an underlying concern that privacy choices made by proxy—even with their initial consent—contributes to a soft paternalism that works to undermine the very autonomy that the notice and choice consent regime seeks to uphold. Today, however, advances in machine learning have provided privacy engineers with a broader toolkit for addressing this paradox.

Other forms of improved notice include those that seek to provide easily recognizable, uniform visual explanations of data collection practices. The multi-layer privacy notice project proposed an easily adaptable template for privacy policies featuring a concise, standardized cover sheet, which would convey the policy’s most salient details.

Users desiring further explanation could then proceed to the relevant area in the full policy that followed. Similarly, “privacy nutrition labels,” modeled on the clear, uniform visual language of consumer warning labels, offered an even simpler means to synthesize the complicated language of privacy policies in one straightforward package. Both of these approaches, however, fall short in terms of circumventing the fundamental problem with privacy policies: despite shortened length and more accessible visuals, these policies continue to overwhelm users, rendering them useless. Further attempts to convey notice in a more standardized and immediate fashion—a “privacy meter” rating the information practices of particular websites, or icons similar to those used by Creative Commons to signify these details—may communicate this content more succinctly, but they also may leave data subjects wanting further clarification.
Widespread adoption of the abovementioned systems has been largely unsuccessful, due in part to a lack of implementation incentives from the web host's perspective. Why, for example, would a company consent to rejecting potential visitors whose interest in their services might actually supersede their privacy concerns? Or grapple with the challenge of ensuring that their data policies comply with the requirements of multiple user agents? This dilemma illustrates why the conversation around notice has largely shifted away from simplifying notice and toward how data subjects experience it.

Calo suggests that notice might be enhanced via experiential design considerations at the point of implementation. “Visceral notice,” as Calo deems it, is fundamentally similar to its written or symbolic counterpart: “The goal is basically the same: to alert one or more individuals to a specific danger within a particular context.” The “alerts” embedded in privacy policies are intended to cultivate a user’s expectations, and subsequent behaviors, for a particular website or service. But where written and symbolic privacy policies often fail short in this regard, visceral notice achieves this objective by leveraging the psychological effects of experiential design. For example, a user might notice an animated magnifying glass appear when they begin entering information into a search form. Hovering over the icon would inform the user that their information is now being collected for third-party purposes designated in the website’s privacy policy. When reading a privacy policy, it is nearly impossible to anticipate how data collection practices will make you feel and behave when they are implemented. Experiencing them, as in the example above, may leave its mark. It’s not unlike being asked to pay for a service with cash in lieu of a credit card; handing over one hundred dollars in bills is figuratively no different than signing it over, but it feels different, and that experience may inform future decisions.

Although seemingly an effective method for educating data subjects about their behavior online, this proposed method is not without drawbacks. As Calo acknowledges, privacy notices that have a dramatic effect on user behavior may ultimately overstep both their intended purpose and the agency of the user. In the aforementioned example, the magnifying glass icon might instill a sense of distrust in the user, with the user ultimately forfeiting their use of the service out of discomfort. But had the user not opted out, they would’ve received more immediate and specific search results based on their search history. In this case, a well-intentioned effort to be more transparent results in a potentially negative behavior. The purpose and design of visceral privacy notices, then, should be “to create awareness of data collection and other relevant issues and realities, rather than to stop consumers from disclosing per se.”

In constructing informed consent, notice is just half the equation. Even if users were to fully understand data policies and practice, the dearth of options for objecting or agreeing to these uses might effectively nullify their consent. In other words, users must be able to choose whether they accept the terms and conditions of a particular data practice. Whether intentional or not, the complex jargon and lack of transparency in many privacy policies obfuscate a user’s choices as well. Developments in choice architecture seek to address this problem in a number of ways. First, a recent study demonstrated how supervised machine learning could be utilized to isolate user choices enounced within a conventional privacy policy. Complemented by enhanced notice mechanisms, this system would theoretically provide users with a more straightforward approach to self-management. In addition, others have suggested that another potential improvement to choice architecture might be building more refined controls into system design, subverting the binary choices typically proffered by these policies (that is, “opt in” or “opt out”).

These strides in notice and choice innovation are fundamentally designed with self-management in mind. But within the greater context and scope of data collection, this framework leaves something to be desired in terms of empowering data subjects to give their consent. Under these circumstances, the onus remains on the data subject to understand the myriad uses of their data across overlapping, often opaque series of networks.

One potential solution to this problem is AI-supported notice. Leveraging the benefits of machine learning, these systems build on the design inferences of the aforementioned notice innovations to endow data subjects with relevant information, informed recommendations, and trust in data-collecting services. With regard to the latter, a recent study employed machine learning and natural language processing to evaluate the behavior of nearly 18,000 free Android apps with regard to their respective privacy policies. The results expose widespread inconsistency between stated privacy policies and app behavior—an average of 1.83 potential inconsistencies per app. A major cause of these discrepancies seems to be a lack of privacy education at the design stage—app developers are not often privacy scholars, and vice versa—but the result is the same: users cannot willingly consent to a system that is at odds with itself. This research serves as the potential basis for a much-needed privacy policy analysis system, and illustrates the mechanism, machine learning, by which this system could be scaled and implemented across the expansive field of data services. This system could prove beneficial from both designers’ and users’ perspectives—the former enabled to detect any potential liabilities in their product before launch, the latter
instilled with confidence that their information will not diverge from its intended use.8

But these improvements in regulatory oversight achieve little in the way of preventing privacy violations that occur at the hands of unwitting users themselves. Beyond consenting to opaque data collection and usage requests, users also put themselves at risk by inadvertently exposing sensitive information in plain sight.9 Although common sense would suggest snapping a photo of a new credit card and sharing it on Instagram might have serious privacy ramifications, a study out of the Max Planck Institute for Informatics sheds light on the recent trend of voluntarily sharing other sensitive information—home addresses, usernames, license plate numbers—via public images. This is largely attributable to the fact that users often fail to understand how this visible information might represent a risk to their privacy. This is particularly salient given that this oversight appears to exist even when the user has expressed an explicit aversion to sharing this kind of information publicly within their social media privacy settings.9 Here, machine learning can also be used to help users understand how to limit the exposure of their private information—at least to the extent that this behavior aligns with their previous privacy preferences. Trained on a dataset of 22,000 images annotated with 68 privacy attributes, the Visual Privacy Advisor monitors a user’s social media output for potentially incriminating data and evaluates these disclosures against the user’s privacy settings, ultimately providing users with a “privacy score” and evaluating these disclosures against the user’s privacy preferences. This system would not address images shared by friends that exposed this information, for example. In other words, a user’s adherence to her personal privacy preference is not fully within her control—it may hinge on the behavior and privacy preferences of other people. Whereas one-time, initial consent can be improved by these efforts, digital consent remains weak for three reasons. First, widespread user adoption or standardization of these tools has not been achieved. Second, the tools focus on providing upfront information to the user, which remains limited for emerging or otherwise unpredictable AI behavior. Finally, the tools cannot help users understand future risks and uses that may be impossible to glean at collection, and as such, risks would remain speculative for both the system and the user.

More to our point, these tools do not foster a partnership between parties—a vital component of consent. They instead empower the user to better understand notification of what a powerful actor is willing to state they intend to do with personal information and choose to go to another service. Digital systems should be responsive to user questions about data collection and easily engage with a user who is not interested in sharing in certain ways and contexts.

Automating Consent
AI may be able to do more than improve notice and choice to support user consent. AI may be able to technically grant consent as an agent of the data subject by predicting what a particular user finds unobjectionable. The Usable Privacy Policy Project contemplates and explores a number of approaches for enhancing the notice and choice consent regime with the help of machine learning. In addition to an automated privacy policy analysis tool similar to that mentioned above, the group has also begun to use artificial intelligence to offer insight into the predictability of user preferences. Despite significant variances across their sample in terms of the range of privacy preferences, the results illustrate that most people fit comfortably within a small group of “privacy profiles.” Layering these profiles onto the detected policies, AI was able to predict user behavior with nearly 80 percent accuracy. When users’ settings were changed to match these recommendations, only 5.6 percent were ultimately changed back.10

Beyond enhancing the user’s ability to make informed choices via recommendations, it is possible to see how these profiles could be put to use in effectuating consent on behalf of the user. This is the aim of the Personalized Privacy Assistant Project, an outgrowth of the Usable Privacy Project. According to the project’s facilitators, “We envision personalized privacy assistants as intelligent agents capable of learning the privacy preferences of their users over time, semi-automatically configuring many settings, and making many privacy decisions on their behalf” (www.privacyassistant.org). Such a system alleviates the user burden of constantly

Such a system alleviates the user burden of constantly configuring individual privacy settings, while keeping the “human in the loop” by alerting the user to any deviations from their predicted preferences.
configuring individual privacy settings, while keeping the “human in the loop” by alerting the user to any deviations from their predicted preferences, prompting a decision from the user, and refining preferences based on these choices. The more the user interacts with her privacy assistant, the greater its accuracy in predicting her preferences. In practice, then, a privacy assistant would serve as the user’s proxy for granting consent to data-collecting bodies. This might be particularly useful in the context of the IoT, in which unconventional and screenless interfaces complicate the notice and choice consent regime.

In a similar vein, Mario Pascalev offers an alternative approach to AI-as-proxy consent. Rather than relying on a small number of collective profiles, Pascalev’s system of “privacy exchange authorities” would confer a user’s individual privacy profile at the point of initiation. These intermediaries would then juxtapose the user’s preferences against the salient details of a website or app’s privacy policy. If the user’s preferences align with that of the service, the agreement would automatically commence. However, if there is any discrepancy between these preferences, the user will then be given a coherent breakdown of these differences and the choice to accept regardless or ultimately decline use of the service. From a user perspective, the ease of implementation and level of transparency might render such an approach an attractive improvement of the current privacy notice and choice consent regime.

These privacy assistants are potentially useful tools for managing the expected dataflows that exist within apps and websites, but they are dependent on a visible interface that allows users to interact with the intelligent agent or quickly configure their privacy settings at a glance. The IoT presents a number of challenges to this system; these devices are often screenless and connected to a wider network of information collection and sharing. In other words, it may be difficult for a user to even recognize when these devices are collecting information about her, let alone manage these preferences. Pappachan and colleagues address this problem by proposing a system of integrated technologies that would allow users to better effectuate consent in this environment. First, an IoT Resource Registry (IRR) would make the privacy policies of local IoT devices easily accessible to the user. Next, an IoT Assistant (like the privacy assistants mentioned above) would interact with the IRR and convey relevant information to the user while automatically configuring the IoT’s settings relative to the user’s privacy preferences. The research team envisions this system as particularly applicable to “privacy-aware smart buildings”—spaces equipped with IoT-enabled devices and sensors that interact with often unwitting inhabitants—but these technologies could be readily applied to any public space where users interact with data-collecting bodies, especially without their knowledge.

Although smart devices like Mattel’s Hello Barbie are not currently able to understand questions and directions regarding information practices, they could be designed to foster a more fruitful form of notification and participation between data subjects and controllers. We currently have the ability to ask Siri or Alexa all kinds of questions about our world. AI is getting better at understanding and responding to natural language questions and offering answers as needed, including the ability to adjust or clarify their initial answer if needed by the user. In this context, there seems to be nothing that prevents AI from engaging in a kind of cooperative partnership with users who can ask about how their information is being used at any point. As machines get smarter, the ability to model key aspects of a cooperative transaction may be possible. Much like the initial context in bioethics where the doctor is expected to explain the proposed procedure or treatment to the patient in understandable terms, AI could be used to explain key aspects of how the system is using or sharing a person’s data. The smart system could then, in theory, respond with additional explanations or easily adjust settings to tailor the collection of data to fit the user’s comfort level. In short, AI could be designed to better model the interactive and cooperative nature of consent between two human beings—potentially fixing some of the problems that plague digital consent when faced with one-sided transactions.

Building on this work, AI systems may eventually be able to more accurately and consistently understand an informational context, as used by Nissenbaum. Such a system in the future may be able to determine when consent is not required (collection and use are within expected information flows) and when consent is required (the information flow would violate contextual integrity). Wijesekera and colleagues suggest that privacy assistants fail because “they often do not take the context surrounding privacy decisions into account.” More specifically, these systems do not understand when context has changed. In this regard, a proposed system based on contextual integrity would closely resemble information sharing in real life. We have certain expectations of how the information we share with our doctors, for example, will be used. These normative expectations hold true across the medical field but differ from how third-party advertisers might want to use this information. We would expect to be consulted and our consent requested to authorize third-party use because this is a relevantly different context than collection of information for medical purposes. A semi-autonomous agent could play this intermediary role, taking action on
behalf of the user when she historically conforms to or deviates from contextual norms and requesting further clarification from the user when her preferences are unclear.\textsuperscript{14}

Research has already been done to explore the feasibility of such a system. A version proposed by Gomer and colleagues promotes semi-autonomous consent by training an agent on a user’s initial preferences toward contextual norms, permitting the agent to approve or deny consent requests based on these preferences, and then refining the model based on the user’s feedback.\textsuperscript{15} Norval and Henderson recommend a similar system with a significant caveat: whereas Gomer’s system retains the users’ ability to make choices with regard to how their data is used, it imposes a user burden similar to that of a privacy assistant, which requires users to more frequently interact with their privacy preferences.\textsuperscript{16} Their approach lessens the user burden by training the agent on a corpus of collective historical data of users’ preferences toward appropriate and inappropriate dataflows.\textsuperscript{16} Finally, a study by Wijesekera and colleagues merges these two approaches, suggesting that the contextual agent might be trained on users’ past decisions toward their data usage as well as supplementary information from their privacy history at large (for example, if they use a passcode lock on their phone).\textsuperscript{14}

As we’ve seen, much of the practice of consent can be automated, but how much of it can be automated ethically?

Although AI can perhaps be used to predict a model of information norms and privacy violations, the normative conversations around risk, marginalization, power, and autonomy should supplement these rules. Normative conversations can and should be used to edit AI-formalized norms and expectations learned by looking at user behavior and preferences, providing an opportunity to address the challenges to privacy self-management mentioned above. In other words, contextual integrity should be established based on explicit normative reflections on information norms and expectations. Explicit user consent would then need to be sought when data collection or use goes beyond expected information flows. Automating when consent is required should not be done without many humans in the loop.

When consent is required, a personal assistant system could reject, block, or opt out based on user preferences, but automating the signification of this user consent should be limited. Automating the roles of the parties (consentee and consenter) limits their investment in the partnership that consent is intended to build. As we’ve seen, consent transforms the moral landscape between two parties, and it is unlikely to be morally valid consent if one of the parties is not even aware of the transformation because her consent has been automatically provided. Consent makes possible important aspects of social cooperation, allowing us to engage in cooperative ventures each party can mutually accept.\textsuperscript{1} Over-automation of the terms of the transaction threatens to undermine the basic trust and control that are essential to any theory of consent.

It must also be noted that these examples of improving or automating consent do little to offer continual consent. They do not attempt to address consent that may need to be adjusted or revoked as risks are reassessed or discovered. AI-effectuated consent does not solve the problems of unforeseeable informational associations or potential uses, nor does it mitigate problems with poor risk assessment that plague digital consent. What is offered is still initial, one-time conclusive consent in an information environment that demands continual consent. Such ongoing participation should be sought, as it is an important component of shifting consent to a more cooperative model.

\begin{quote}
When we automate aspects of consent, we should ask ourselves whether we are enabling or simulating consent.
\end{quote}

A I systems place increased burdens on an already strained system of digital consent, but also offer new possibilities for reform. Significant efforts have been made to improve, through technological means, pieces of the consent puzzle by automating certain portions of notice and choice selection in digital consent. AI systems present unique challenges to digital consent, but also an opportunity to move beyond consent as simply a procedural practice toward a model of continual and cooperative consent. While AI can provide more tools for data management, automating aspects of consent important to the roles and relationship between consentee and consenter may undermine the framework needed for cooperative partnership and mutual trust between parties. When we automate aspects of consent, we should ask ourselves whether we are enabling or simulating consent, and the impact on partnership development between parties. \textsuperscript{•}
Acknowledgments
The authors gratefully acknowledge Georgetown University for its support of this project, part of the Digital Consent Project, through the Complex Moral Problem grant, as well as the Kennedy Institute of Ethics for further intellectual and logistical support.

References

Meg Leta Jones is an assistant professor in the Communication, Culture and Technology Department at Georgetown University. Contact at ma1318@georgetown.edu.

Ellen Kaufman is a second-year MA student in Georgetown University’s Communication, Culture and Technology program. Contact at ek792@georgetown.edu.

Elizabeth Edenberg is senior ethicist for Ethics Lab at Georgetown University’s Kennedy Institute of Ethics. Contact at elizabeth.edenberg@georgetown.edu.

IEEE Software seeks practical, readable articles that will appeal to experts and nonexperts alike. The magazine aims to deliver reliable information to software developers and managers to help them stay on top of rapid technology change. Submissions must be original and no more than 4,700 words, including 250 words for each table and figure.

Call for Articles

IEEE Software seeks practical, readable articles that will appeal to experts and nonexperts alike. The magazine aims to deliver reliable information to software developers and managers to help them stay on top of rapid technology change. Submissions must be original and no more than 4,700 words, including 250 words for each table and figure.

Meg Leta Jones is an assistant professor in the Communication, Culture and Technology Department at Georgetown University. Contact at ma1318@georgetown.edu.

Ellen Kaufman is a second-year MA student in Georgetown University’s Communication, Culture and Technology program. Contact at ek792@georgetown.edu.

Elizabeth Edenberg is senior ethicist for Ethics Lab at Georgetown University’s Kennedy Institute of Ethics. Contact at elizabeth.edenberg@georgetown.edu.

myCS Read your subscriptions through the myCS publications portal at http://mycs.computer.org