Adam J. Lee, Rosta Farzan, Apu Kapadia, and Imtiaz Ahmad

Making sense of risk in an increasingly cyber-physical world

Every action carries with it some risk. For instance, driving carries risks related to personal health and wellbeing that may be realised in the event of an accident, investments have financial risks that depend on changing conditions in the markets, and personal and political viewpoints come with reputational risks that individuals must weigh. Our online lives are no different. Risks of identity theft through data breaches increase dramatically with the vast collection and sharing of personal information, online social networks and dating sites are conduits for cyberstalking or cyberbullying situations, and the persistence of self-published online data thought to be ephemeral can cause reputational harm long after information was posted. The changing nature of online media is forcing a shift in the way that we perceive and manage risks in both our analog and digital lives.

As in all contexts, the risks that we face in our digital lives are intrinsically tied to uncertainties. This is particularly true in the era of big data. Even in situations where someone might be aware of the data that they are contributing to some system (e.g., posts on a social media site, clicks within a shopping marketplace), it is not always clear *who* has access to this information. The unintended audience problem associated with social media has been well documented in the popular press and research literature (e.g., see work by Wang et al.[1] and Patil et al.[2]). Here, individuals may post personal information or express strong or off-colour sentiments with the expectation that a small, well-defined audience only would consume them. However, the scale of sharing on these platforms – the average Facebook user has over 300 friends[3] – means that information can easily flow to more distant social ties. This can have physical world implications, as when individuals lose their jobs for inappropriate posts gone viral.

In addition to uncertainty surrounding the consumers of information, there can also be uncertainty surrounding *how* information is used in

online platforms. This is particularly well-documented in the case of systems that make use of individuals' contextual data (e.g., location[4] or activity[5]) both within social media and within modern, connected workplaces. Many workplaces use communication platforms that expose individuals' activity states (working from home, in the office, with a visitor) to facilitate more effective communication amongst colleagues. Even with appropriate access controls in place (e.g., 'Share my location with Alice between 08:00 and 17:00'), such systems often provide a lack of clarity on how information is used. For instance, Alice might track this individual's location every two minutes to build a rich profile of the individual's activities, which counters the putative goal of facilitating effective workplace communication.

The examples above illustrate how uncertainties surrounding *how* and *by whom* data is used can lead to risks that individuals must begin to manage in this increasingly connected age. The story does not end there, however. Given the proliferation of Internet of Things (IoT) and connected devices like smart assistants, it is becoming increasingly unclear *what* information is collected about individuals and *when* this information is being collected, even within traditionally private spaces like the home.

Ubiquitous computing, ubiquitous confusion

We are living in the age of ubiquitous computing. Our smartphones and smartwatches seamlessly track our locations, activities, interactions with others, and a variety of other telemetry. Our online activities synchronise across our devices, allowing browsing or shopping activities on one device to be picked up on another, and even be informed by our presence in certain physical spaces. The IoT age has seen the powerful coordination of many small, embedded devices and media appliances that can draw insights upon data gathered from multiple devices within one home or workplace, and across other spaces. Furthermore, voice control has proven to be a powerful means for manipulating a variety of devices without traditional user interfaces, and cameras have become inexpensive enough to embed throughout physical spaces. As noted by Mark Weiser in his seminal article[6] predicting the emergence of ubiquitous computing, 'the most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it.' We are no doubt approaching this point with IoT technologies.

In his article, Weiser discusses the evolution of motors and mechanical power. In the industrial age, a single centrally located steam engine

typically powered every machine in a factory by using a necessarily visible network of gears and belts. He contrasts this with today, where tens of motors and solenoids can be found within a single car, often hidden from sight. This backgrounding of technology can hide the complexity of everyday life in a way that increases user convenience. This is certainly true in the context of IoT devices as well, but it paints an incomplete picture. Playing an album by saying 'Hey Google, play *Kind of Blue*' is positively convenient: this can be done while using both hands to tend to an upset child, while preparing a meal or washing dishes, or when one simply wishes to avoid the search for a physical album and the manual process of setting it up to play. However, a focus on convenience alone eschews important questions about device behaviour and information flows that are central to risk management in today's connected environments. *When is this device listening to me? Where is it processing the information that it captures? How is this information stored? Who has access to this information?*

The introduction of such high-fidelity sensors thus offers powerful functionality for users but comes with great privacy risks. An individual's private conversations and actions – in high definition – are now potentially streamed to the cloud. This information can then be leveraged by corporations to learn about people's habits, conversations, and their physical possessions. This data can be purchased or legally acquired since cloud-based services cede control of data to the cloud provider; major companies differ in how such requests are handled and do not necessarily defend consumers, according to a recent report by the Electronic Frontier Foundation.[7] Whereas once an individual's physical actions were private, these networked cameras and microphones 'peeking' into the physical world can give rise to new privacy concerns for one's physical behaviours. People now cannot have a reasonable expectation of physical privacy even in their own bedrooms or living rooms. Casual conversations and encounters, once thought to be private and ephemeral, now may be captured and archived digitally for posterity, potentially available to third party entities indefinitely.

A growing body of work seeks to address privacy in IoT devices and sensors. The majority of techniques assume a trusted infrastructure where access control policies and mechanisms control (through software) *who* can see *what* data and under *what circumstances*. IoT devices can communicate with an individual's personal devices (e.g., their smartphones), provide information about what is sensed, and how users may control the sensing (e.g., see work by Das et al.[8] Korayem et al.[9] Roesner et al.[10] and Templeman et al.[11]). However, software access controls cannot ensure compliance by the device, and individuals must

therefore trust the implementation that enforces these controls. Recognising the low assurance provided by software solutions, another body of work builds upon the assumption that the sensing infrastructure cannot be trusted (e.g., by not supporting policy-based controls, or by being overtly malicious) and proposes techniques to thwart sensors. These works demonstrate techniques that may derail the cameras either by using additional accessories[12] or tools.[13] Recent work shows that users seek and exhibit alternative privacy-enhancing behaviours such as sticking tape or stamps over their laptop cameras, and senior citizens adapating their movements around home-monitoring cameras.[14]

Thwarting sensors via overt action is a reactionary approach in the absence of such trust. A more holistic approach is easily motivated by this lack of trust in opaque devices. More pointedly, we argue that device designers and developers can and should aspire to remove the responsibility of engaging in extreme or inconvenient privacy-enhancing behaviours by creating devices that better engage with the social theories around personal privacy and provide more transparent mechanisms for understanding, awareness, and control.

Towards a more tangible notion of privacy

Physical privacy in the absence of electronic devices is typically well understood: by closing the door prior to having a sensitive conversation, or lowering the blinds in a bedroom, we feel confident that we have eliminated privacy risks. These assurances need not be strictly binary, however. Distancing oneself from third parties during a conversation can make it more difficult – though not impossible – for eavesdropping to occur; and frosted glass windows, walls, or doors can hide certain information about the goings-on within a space while exposing others. The observation in all of these cases is that our spaces can provide *tangible* feedback that reflects what can be observed about an individual; this, in turn, can inform individuals' actions.

D.A. Norman[15] posited that feedback to individuals about the internal operations of machines or systems is essential for awareness, reassurance, and anticipation of further actions. This tangible feedback and control is noticeably missing in many IoT devices, as individuals often are not able to discern when and how they are being recorded. It is oftentimes unclear whether an audio or video sensor is, indeed, in the 'off' state, even upon visual inspection. For example, certain types of connected cameras can be set to stop recording when an individual is at home, as determined by their phone's location sensor, but there is no physical indication or assurance that the camera is indeed off. In some

of these cameras, LED indicators of recording state exist, but the use of this indicator is optional and controlled by software. Even for situations where a camera's LED recording state indicators are controlled (ostensibly) through hardware, there have been demonstrated attacks that are able to disable the LED indicator for some laptops.[16] Without clear assurances of device activities, an individual's ability to assess, respond to, and manage risk is severely hindered.

As computing becomes ever more inextricable from everyday life, addressing this source of uncertainty – and therefore risk – is crucial. The exploration of sensor designs with tangible affordances offers a rich space for interplay among the humanities, social sciences, and computer and information sciences.

Risk management in a cyber-physical world

Existing theories related to social interactions and privacy management can serve as a basis for understanding privacy within the context of IoT devices in private and in social environments. Although much progress has been made in theories of privacy to keep up with the information age – for example, work by Helen Nissenbaum[17] and Sandra Petronio[18] – Irwin Altman's theory of privacy regulation has remained a seminal work with respect to regulating social interactions in physical spaces.[19] Altman conceived of privacy as 'an interpersonal boundary process by which a person or group regulates interaction with others' by altering the degree of openness of self to others.[20] Altman's model of privacy regulation starts with an individual's desire to achieve a certain level of privacy, which is derived from a combination of personal, interpersonal, and situational factors. Then, through an iterative process, individuals use different control mechanisms to move towards their desired level of privacy, and in each iteration, they assess the effectiveness of their control mechanisms by comparing their *desired* privacy with their *actual* privacy. If their attempt to control their privacy is insufficient (i.e., their actual privacy is less than their desired privacy), they can experience a feeling of *crowding* in their space. If their attempt at managing their privacy is more than sufficient (i.e., their actual privacy is more than their desired privacy), they can experience *social isolation*. Privacy control mechanisms involve verbal and nonverbal behaviours (such as body language) and consideration of personal space and territories. For instance, lowering one's voice to make a private comment, or leaving a crowded space to avoid observation, would be examples of privacy control mechanisms.

This iterative process of performing controlling actions and assessing their resultant (actual) privacy as compared to their desired privacy is the key feature of Altman's model. As noted, in the case of privacy regulation of interpersonal interactions in the physical world, the assessment of the actual level of privacy yields a high level of certainty. However, as noted by Mitew,[21] IoT devices can act more like social agents whose primary focus is not only computing data, but also actively shaping the social environment. In this new socio-technical landscape, people must navigate their personal boundaries of privacy not only with other people, but also with devices that can monitor, record, and share their physical interactions.

When interacting with IoT devices, individuals can feel a similar level of crowding and desire to achieve a certain level of privacy. However, in these cases, the lack of clear, tangible feedback from the devices makes the assessment of privacy level ambiguous and inaccurate: an individual's *perceived* privacy can differ greatly from their *actual* privacy. Thus, we argue that in order to achieve an individual's desired state of privacy, social theories for privacy in the context of IoT devices must concretely account for tangible privacy feedback – or lack thereof – in the context of IoT devices.

Understanding the path forward

High-fidelity sensors now capture rich information, transmit this information to remote servers for storage and processing by algorithms (and possibly human analysts), and provide little in the way of insight as to what information is being gathered, or when it is being collected or transmitted. Similarly, they do not account for who has access to it. This leaves consumers unable to make informed risk-management decisions. What is less clear is the best path forward. Addressing the vast issues surrounding this space requires an interdisciplinary, socio-technical perspective to rethink cyber-physical risk management for the IoT age, and an understanding that social theory should inform system design at the same time that system functionality plays into these social theories. As we design, deploy, and use sensors with tangible affordances, there are a variety of questions that deserve consideration.

What is this device doing? The idea behind tangible privacy is the grounding of a device's sensing state in a physically observable configuration. Although a green LED *might* accurately indicate whether a camera is recording, there is no physical link between the indicator LED and the camera's status. On the other hand, the observation that a microphone's cable is physically disconnected from the sensing platform provides a clear indication that the microphone is not powered and cannot

be recording or transmitting audio data. Similarly, frosted or opaque shutters over camera lenses convey exactly what the camera is capable of capturing at the moment. Tangible sensors in the 'off' state should convey the same privacy assurances to individuals regarding *electronic* observation that closed doors and windows do with respect to *physical* observation. However, the design of these tangible affordances cannot be carried out as a purely engineering effort; the messages conveyed by tangible affordances must be consistent with prevailing social norms and unambiguous to the individuals using these devices.

How tangible should this device be? It is possible to design sensors with a variety of types of tangible affordances; however, it is important to keep devices simple and to address the most pressing needs of their users. Understanding the privacy and functionality desires of users and balancing the tensions that may arise within this space is necessary for enabling the types of iterative control that is typical of risk management in social settings. For instance, one could imagine the distinction between tangible *control* and tangible *feedback* being important for different classes of individuals. Someone who identifies as a 'privacy fundamentalist'[22] might insist upon tangible control of a device (e.g., manually adjusting camera shutters or microphone effects) to ensure that the configuration of sensors within their space cannot be surreptitiously altered without their knowledge. However, such manual control would probably chip away at one of the key draws of IoT devices: convenience. The difficulty of adjusting the configuration of cameras mounted in inconvenient locations would probably discourage the use of tangible controls, for instance. On the other hand, the tangible feedback conveyed by an opaque lens cover might offer reassurance of a camera's current state, even if it could be enabled and controlled electronically (e.g., via a cloud-connected smartphone app). 'Privacy pragmatists' might find enormous value in such a configuration, even if there is the possibility for surreptitious changes to sensor configurations. Human-centred approaches that balance technological possibilities with social needs should play a driving role in the design of sensors embodying tangible privacy.

Can I trust the assessments that I am making? Risk management is an iterative process that requires making judgements of one's exposure, taking an adaptive action, and reassessing. Unless an individual's exposure assessments are accurate, the process will not converge to a desirable state. In particular, researchers have shown that risks should be communicated to users in a way that leverages their existing 'mental models' for them to manage their privacy more effectively.[23] Close cooperation between social scientists and technologists, as well as social theory and sensing platform co-design are necessary conditions for success within this space.

A bright horizon

The popular press and research literature are full of instances where individuals have been caught off-guard by privacy violations in networked systems resulting from a lack of clarity about *what* information is being captured by a system, *when* that information is being used, *who* is using that information, and the *purpose* for which that information is being used. This trend shows no signs of slowing, and the landscape is being made ever more complex by the increasing degree to which computing and data collection are being woven into even our most intimate physical spaces. In this essay, we have put forward the argument that the disconnect between social theories of privacy regulation and the social presence of ubiquitous sensing platforms implies that privacy gains in this space will be hard fought. However, the exploration of sensing platforms that expose tangible affordances to users both in terms of *feedback* describing their current state and *control* over their state changes stands poised to close this gap. Making progress in this space will require cooperation among humanists, social scientists, and technologists – and sooner, rather than later.

Notes

This material is based upon work supported by the National Science Foundation under Grant No. CNS-1252697, CNS-1253204, CNS-1814513, and CNS-1814866. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

1 Y. Wang, G. Norcie, S. Komanduri, A. Acquisti, P.G. Leon, L.F. Cranor, '"I regretted the minute I pressed share": a qualitative study of regrets on Facebook', in Proceedings of the Seventh Symposium on Usable Privacy and Security (SOUPS), July 2011.
2 S. Patil, G. Norcie, A. Kapadia, and A.J. Lee, 'Reasons, Rewards, Regrets: Privacy Considerations in Location Sharing as an Interactive Practice', in Proceedings of the Eighth Symposium on Usable Privacy and Security (SOUPS), July 2012.
3 See A. Smith, 'What people like and dislike about Facebook', February 2014; https://www.pewresearch.org/fact-tank/2014/02/03/what-people-like-dislike-about-facebook/
4 R. Schlegel, A. Kapadia, and A.J. Lee, 'Eyeing your Exposure: Quantifying and Controlling Information Sharing for Improved Privacy', in Proceedings of the Seventh Symposium on Usable Privacy and Security (SOUPS), July 2011.
5 A.J. Lee, J.T. Biehl, and C. Curry. 'Sensing or Watching?: Balancing Utility and Privacy in Sensing Systems via Collection and Enforcement Mechanisms', in Proceedings of the ACM Symposium on Access Control Models and Technologies (SACMAT), June 2018.
6 M. Weiser, 'The Computer for the 21st Century', *Scientific American*, September 1991.

7 See 'Who Has Your Back? Government Data Requests 2017', July 2017; https://www.eff.org/who-has-your-back-2017#govt-requests.
8 A. Das, M. Degeling, X. Wang, J. Wang, N. Sadeh, and M. Satyanarayanan, 'Assisting users in a world full of cameras: A privacy-aware infrastructure for computer vision applications', in 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 1387–96, 2017.
9 M. Korayem, R. Templeman, D. Chen, D. Crandall, and A. Kapadia, 'Enhancing lifelogging privacy by detecting screens', in Proceedings of the Conference on Human Factors in Computing Systems (CHI '16), 4309–14, May 2016.
10 F. Roesner, D. Molnar, A. Moshchuk, T. Kohno, and H.J. Wang, 'World-driven access control for continuous sensing', in Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, 1169–81. ACM, 2014.
11 R. Templeman, Z. Rahman, D. Crandall, and A. Kapadia, 'PlaceRaider: Virtual theft in physical spaces with smartphones', in Proceedings of the 20th Annual Network and Distributed System Security Symposium (NDSS), February 2013.
12 M. Sharif, S. Bhagavatula, L. Bauer, and M.K. Reiter, 'Accessorize to a crime: Real and stealthy attacks on state-of-the-art face recognition', in Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, CCS '16, 1528–40. ACM, 2016.
13 S.N. Patel, J.W. Summet, and K.N. Truong, 'BlindSpot: Creating Capture-Resistant Spaces', in *Protecting Privacy in Video Surveillance*, ed. A. Senior (London: Springer-Verlag, 2009), 185–201.
14 K. Caine, S. Sabanovic, and M. Carter, 'The effect of monitoring by cameras and robots on the privacy enhancing behaviors of older adults', in 2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 343–50, March 2012.
15 D.A. Norman, *Turn Signals Are the Facial Expressions of Automobiles* (New York: Basic Books, 1992).
16 M. Brocker and S. Checkoway, 'iSeeYou: Disabling the MacBook Webcam Indicator LED', USENIX Security Symposium 2014, 337–52.
17 H. Nissenbaum, 'Privacy as contextual integrity', *Washington Law Review*, 79:1 (2004).
18 S. Petronio, 'Communication privacy management theory: What do we know about family privacy regulation?', *Journal of Family Theory & Review*, 2:3 (2010), 175–96.
19 Irwin Altman, *The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding* (Monterey CA: Brooks/Cole, 1975).
20 Ibid.
21 Teodor Mitew, 'Do objects dream of an Internet of Things?', *Fibreculture Journal*, 23 (2014).
22 'Privacy On and Off the Internet: What Consumers Want'. Technical Report 15229, Harris Interactive, February 2002; http://www.ijsselsteijn.nl/slides/Harris.pdf.
23 Farzaneh Asgapour, Debin Liu and L. Jean Camp, 'Mental Models of Computer Security Risks', in *Financial Cryptography and Data Security*, ed. Sven Dietrich and Rachna Dhamija (Berlin: Springer-Verlag, 2007), 367–77.