PRIVACY AND THE INTERNET OF THINGS

Kieron O'Hara

University of Southampton

Short position paper presented at *Internet of Things Ecosystem – the Next 40 Billion Devices*, NESTA, London, 3rd June, 2014

1. The salience of privacy

The sensors and other devices that constitute the Internet of Things (IoT) will provide an unprecedently rich picture of anything in their purview. As the IoT is projected to include devices embedded in the home, on people's bodies (via instrumenting the body and also wearable computing devices), in communities, on goods and services, and in the environment (Anderson & Rainie 2014, 2), very many previously-hidden activities will be exposed to potential view. The point does not need belabouring – indeed, Sir Mark Walport's opening remark in this event alluded to the privacy issue. Martin Winterkorn, the CEO of Volkswagen, whose company places many digital devices in the cars it builds, recently warned of the dangers of creating a "data monster" (Bacon 2014). Meanwhile security is a genuine worry. Security firm Proofpoint recently uncovered the first known large-scale IoT cyberattack (Proofpoint 2014).

The attack that Proofpoint observed and profiled occurred between December 23, 2013 and January 6, 2014, and featured waves of malicious email, typically sent in bursts of 100,000, three times per day, targeting Enterprises and individuals worldwide. More than 25 percent of the volume was sent by things that were not conventional laptops, desktop computers or mobile devices; instead, the emails were sent by everyday consumer gadgets such as compromised home-networking routers, connected multi-media centers, televisions and at least one refrigerator. No more than 10 emails were initiated from any single IP address, making the attack difficult to block based on location – and in many cases, the devices had not been subject to a sophisticated compromise; instead, misconfiguration and the use of default passwords left the devices completely exposed on public networks, available for takeover and use.

2. How should we understand privacy?

It is tempting to understand privacy as a type of control – one prevents data or information about oneself reaching other people. One sets rules and then enforces them. Yet this badly misrepresents the ways in which we talk about our privacy, or in which privacy concerns us.

I would contend that privacy is better understood as a constant effort to negotiate the boundaries of our selves and our personal space (literally and metaphorically) with our peers. We perceive benefits from being visible to our networks, and to other entities such as our governments or our employers, and at the same time wish to restrict that visibility in order to preserve a private space for action, subversion, relaxation, contemplation, reflection or intimacy. Similar, our networks and governments wish us to be visible in order that they might maximise benefits to themselves and their members, yet also have interests in limiting the amount conveyed, both for reasons of propriety ("too much information!") and to avoid information overload. Real and ideal boundaries change with context, reflecting tensions between our goals (and between our goals and the goals of our networks).

Hence we both push and pull at the boundaries around our selves, and these boundaries are pushed and pulled by our networks. This is a constant process of negotiation and compromise. It will not reach equilibrium, and it cannot be described or delimited by simple sets of rules determining how data may or may not flow. Technology is not an exogenous force which disrupts or reinforces privacy. It is part of the changing context in which boundaries are negotiated (for example, it affects what data can be gathered about an individual without permission, what benefits accrue to the individual, who can potentially get access to a disclosure, who can be effectively held accountable for usage of data, and the relative value of a disclosure to an individual compared to his/her network). This picture of privacy as a dynamic, perpetual process is inspired by the works of writers such as Irwin Altman and Helen Nissenbaum, and the venerable but expressive common law concept of 'reasonable expectations of privacy' (Altman 1975, Palen & Dourish 2003, Nissenbaum 2010).

If we are lucky, our preferences and interests will converge to produce norms, or allow uncontroversial rules to be drafted. This will not necessarily happen, and as I describe later, I would not be optimistic that it will in the case of the IoT.

The lesson, then, of this understanding of privacy is that data protection authorities should not try to micromanage dataflow in order to defuse privacy as an issue. Rather, we need room for our privacy negotiations to take place, and confidence in the process (i.e. that negotiations will be in good faith). As noted by a previous speaker today, one pressing need is to price into markets the current externality of disenfranchisement. Lack or loss of trust is a major cost of badly-adapted systems, but currently it is absorbed by society rather than the institutions which caused it.

3. Who should arbitrate questions of privacy?

This is something of a vexed question. Regulating privacy is a complex matter, and in the particular case of the IoT it is not immediately obvious how that should be done. One important conclusion we draw, however, is that the technology industry, and privacy by design (PbD), must play a prominent role.

This is because there is no completely satisfactory answer to the question. Certainly **government** is not in a good position to preserve confidence through regulation. Post-Snowden, government cannot pose credibly as a disinterested actor in the regulation of personal data. That is not to say that government agencies acted wrongly, necessarily, in their pursuit of security, but only that many arms of government clearly believe they have an interest in the available pool of data.

On the other hand, although predictions in this area are of necessity speculative, it would seem unlikely that **social norms** will easily emerge. Use and understanding of technology are highly fragmented and diverse across demographic groups. Furthermore, IoT technology is developing extremely quickly. It should also be borne in mind that the IoT may be largely invisible to most people, who may not feel much of a pressure to respond to its challenges.

Meanwhile, **market solutions to privacy** are also unlikely to provide a quick win. Data has been commoditised and monetised so successfully that its value now dramatically outweighs consumers' market power.

This leaves **PbD** solutions to play a leading role. However, even this is not perfect, as they will add complexity and possibly undo the business model of connecting many very simple devices. For example, consider a smart meter. One of its functions is to convey billing information to the utility. The live stream of data would be extremely disclosive, as many electrical devices have clear signatures that enable conclusions to be drawn about their use (Lisovich et al 2010). There is of course an obvious solution – the utility does not need the live stream, so the output of several days or weeks can be aggregated and sent periodically (Wicker & Thomas 2011). But aggregation implies storage, which implies the need for extra security measures, which implies greater complexity. Security imposes a cost, which may be a particular problem in an area where business models often depend on keeping costs down.

4. What specific problems does the IoT pose for privacy?

A final question to consider is what particular issues the IoT raises for privacy. Clearly, digital technology generally has challenged our current norms and concepts of privacy, and this has been a subject of an immense literature. As a digital technology, this will apply to the IoT, but on top of this, the architecture of the IoT also poses complex questions of its own. My point is not that these are unanswerable, but merely that they have to be faced.

(i) The IoT strings together many devices that are (a) small, cheap and simple, and (b) highly heterogeneous. Can we ensure that these devices are clever enough to implement decent security? If not, where in the architecture can security be safely located without creating vulnerabilities?

(ii) In many IoT scenarios, devices will be in place for several years (e.g. embedded in white goods). So security needs to be implemented not only for now, but needs to be futureproofed as well. We can't withdraw billions of devices if hackers undo today's security arrangements.

(iii) IoT systems are highly distributed, so data needs to be moved around a lot. This raises a problem for privacy regulation. Not only is there an increased risk of interception, but more importantly it is harder to trace where data is, and what it is used for. This will make it much more difficult to hold people and institutions to account for their use of data.

(iv) Many IoT devices are very simple, and the data they produce will probably not be personal data. However, when aggregated with other datasets, personal data will be created. This situation, where the data collected are not personal but the aggregate is, is unusual, and data protection practice may not be fully geared up to it at scale. The issue of "when does this become personal data?" does need to be addressed.

(v) There is a serious issue of consent. In many cases (e.g. wearables), the wearer's consent is implied. But implied consent is not a tool that can solve every problem in this space. If our normal environment is instrumented, then the implied consent model for IoT risks making life impossible for those who do not consent to their data being used. For a discussion of these issues, see (Luger & Rodden 2013).

(vi) Finally, there is a strong and tight connection between people and things in the IoT, which will complicate privacy issues and expose many vulnerabilities in human engineering. For instance, it is significant that the Proofpoint discovery (Proofpoint 2014) quoted above, was of an attack dated over the Christmas period. In many cases, it would seem possible that devices were Christmas presents, turned on and tested by their recipients, with a short period before passwords were changed from their defaults. I speculate here, but this may have provided a brief window of vulnerability for the attackers. I set out the example to show how Sisyphean it would be for security systems to attempt to cover all vulnerabilities.

References

Irwin Altman (1975). *The Environment and Social Behavior: Privacy, Personal Space, Territory, Crowding*, Monterey, CA, Brooks/Cole Publishing Co.

Janna Anderson & Lee Rainie (2014). *Digital Life in 2025: The Internet of Things Will Thrive by 2025*, Pew Research Center, http://www.pewinternet.org/2014/05/14/internet-of-things/.

Jonathan Bacon (2014). 'All brands should heed VW's "data monster" warning', *Marketing Week*, 11th Mar, 2014, http://www.marketingweek.co.uk/disciplines/data-/-crm-/-loyalty/all-brands-should-heed-vws-data-monster-warning/4009745.article.

Mikhail Lisovich, Deirdre Mulligan & Stephen B. Wicker (2010). 'Inferring personal information from demand-response systems', *IEEE Security and Privacy* 8(1), 11-20.

Ewa Luger & Tom Rodden (2013). 'An informed view on consent for ubicomp', *UbiComp 2013*, Zurich.

Helen Nissenbaum (2010). *Privacy in Context: Technology, Policy and the Integrity of Social Life*, Stanford: Stanford University Press.

Leysia Palen & Paul Dourish (2003). 'Unpacking "privacy" for a networked world', *CHI 2003*, Ft Lauderdale, FL.

Proofpoint (2014). *Proofpoint Uncovers Internet of Things (IoT) Cyberattack*, press release, 16th Jan, 2014, http://www.proofpoint.com/uk/about-us/01162014.

Stephen Wicker & Robert Thomas (2011). 'A privacy-aware architecture for demand response systems', *44th Hawaii International International Conference on Systems Science (HICSS-44 2011)*, Koloa, Kauai, HI.