Cyber Security Policy and Research Institute

THE GEORGE WASHINGTON UNIVERSITY

The Internet of (Whose) Things: Business Models, Computer Architectures, and Privacy

Allan Friedman
Lance J. Hoffman
Cyber Security Policy and Research Institute
The George Washington University
Washington, DC

July 8, 2014

Report GW-CSPRI 2014-3

Support for this research was provided through a grant from the Centers & Institutes Facilitating Fund (CIFF) of the Office of the Vice President for Research of the George Washington University

Abstract

The emerging world of diverse, connected smart devices and sensors known as the Internet of Things has the power to transform society, but also introduces or aggravates very real privacy and security risks. While traditional discussions of privacy and security are often linked, we see a marked seperation in the IoT context.Security experts focus on devices while the privacy community has been primarily concerned with aggregated data held on the cloud. We attempt to integrate these two discussions by presenting an architectural perspective of how the different components of end-points, devices and links might fit together.

Comparing this generalized architecture framework to what is currently offered by the market leads to an interesting observation: much of what is currently called an Internet of Things resembles a very simple architecture, the client-server relationship. This one-to-many structure with centralized control makes security easier, but concentrates data for greater potential harms. While there are good reasons for early IoT applications to follow this basic model, how can we understand risk and control of future instantiations of IoT?

We use the architectural framework that emphasizes links in the network to explore control points, where engineers might build in security or privacy tools. Control points can rely on both technical and human-level protections, but their flexibility also introduces too much ambiguity—in an open-ended network structure, control points might be inserted anywhere. This can lead to either over- or under-protection of data and systems. We use a business-case analysis to limit the set of possible network configurations for future IoT applications, and suggest how control points might be used in these cases. The paper concludes with some general rules for engineering security and privacy into IoT.

The Internet of (Whose) Things:
Business Models, Computer Architectures, and Privacy

Allan Friedman
Lance J. Hoffman
Cyber Security Policy and Research Institute
The George Washington University
Washington, DC

Introduction

The Internet has been constantly growing and evolving at in incredible pace over its almost 50 year history. For at least the last 15 years, the next stage in its evolution has been predicted to be a spillover from cyberspace into the world around us: the Internet of Things (IoT). The blurring of the lines between the digital realm and the "real" world is a function of a number of trends. The number of devices connected to the Internet has exploded, with estimates as high as 50 billion by the end of the decade.[1] Sensors are cheap, and can be placed into anything, and connectivity has grown and taken on many new forms, from standardized mobile networks to the potential of newly freed open spectrum. The rise of cloud computing has enabled these new services and applications to take advantage of remote storage and processing. We've made progress in taking advantage of the massive amount of digital data generated everyday by ordinary transactions. When these data comprise information captured by our *stuff* as well, the potential provide consumers with an unprecedented array of smart applications and services seems limitless.

This grand vision of a world of world of networked intelligent objects also brings with it specters of risk. In a world that has yet to understand the risks of 'classic' data collection and analysis, expanding the power and reach of data collection to every corner of our lives have sounded loud, clear alarms among scholars in the academy,[2] activists in the privacy community, and policymakers in Washington.[3] Security is also a significant concern, as data generation and collection becomes even less centralized, and already shaky trust models and security practices such as the certificate system come under even further strain.

A great deal of work has gone into characterizing the nature of the risks this new era of the Internet presents. Much of it has focused on the data itself, the qualitative

[1] Cisco White Paper.
https://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf
[2] A fairly comprehensive survey can be found in Dutton, W., et al. (2013), 'A Roadmap for Interdisciplinary Research on the Internet of Things: Social Sciences', addendum to Internet of Things Special Interest Group, A Roadmap for Interdisciplinary Research on the Internet of Things. London:

TSB.

[3] FTC Public Workshop on Internet of Things – Privacy and Security in a Connected World. November 19, 2013. http://www.ftc.gov/news-events/events-calendar/2013/11/internet-things-privacy-security-connected-world

differences in its source and the potential for transformative impacts following its analysis and use.[4] While these are key questions, they do not tell the full story of IoT. Despite years of hype and investment, we remain quite far from truly ubiquitous computing and connectivity. Yet the fact that we are still in the initial stages of this revolution allows us to seriously consider that most laudable goal of the privacy world: privacy by design.[5] This paper explores the privacy and security concerns surrounding IoT from the perspective of information architecture. We look, not at the data itself, but how it will be generated and flow to understand the security and privacy risks inherent in new forms of networking.

We begin by clarifying the scope of our inquiry, not re-defining IoT, but highlighting what is genuinely new about the issue. We then introduce an architectural perspective of IoT, building a generalized model of how data can be collected, transmitted, and used. We then use this architecture to understand the importance and challenges in controlling the flow of data. While our model is general, the future is not one of infinite potential paths. The shape of the network will be driven by business models *and* architecture, since it is inextricably bound into how the technology will be used, and who will be trying to use it. We conclude with a series of observations about how different business models will enable different types of control over the flow of data, which will, in turn, enable different responses to privacy and security risks.

Defining and Scoping the Issue

There have been a number of attempts to define the Internet of Things, for a variety of reasons, in a number of contexts.[6] For our purposes, the actual definition is less important than the acknowledgement that we are still very far in real life from any previous vision of ubiquitous computing and connectivity, however it might be conceived. There are some interesting devices and applications that are embedded in our lives today, but they are relatively small, existing in narrow verticals that quickly veer back into the digital realm.

We also want to set aside two key issues as outside the scope of this inquiry. Locational information, and the privacy and security concerns it brings, have been a key focus of many of the discussions around IoT. While it is true that locational data can be very useful, and introduce a great many risks, we are already well on our way to a world with sharable location information *without* starting up the steep S-curve of IoT diffusion. The mobile revolution is further along, and prognostications about it should be treated as distinct (although not completely independent) from those about the future of IoT. Similarly, just adding a GPS transponder to, say, a car raises a set of issues that are still somewhat distinct from an open ended set of data

---

[4] Dutton et al.

[5] Langheinrich, Marc. "Privacy by design—principles of privacy-aware ubiquitous systems." Ubicomp 2001: Ubiquitous Computing. Springer Berlin Heidelberg, 2001.

[6] See, e.g. Atzori, Luigi, Antonio Iera, and Giacomo Morabito. "The internet of things: A survey." Computer networks 54.15 (2010): 2787-2805.

collection tools and communication pathways. For the purposes of this paper, we are looking past location information.

Second, we carve out for exclusion the paradigm of the "Industrial Internet,"[7] although we briefly revisit it at the end of the paper. While this is a marketing term coined by GE, it captures the distinction between an open, generative vision of IoT, and a no less powerful but much more controlled vision of smarter manufacturing and business operations. This paradigm exists within a relatively closed vertical, where the value from the investment will be captured by the firm or firms making the investment. And while we pick up this question briefly below, most of the Things in the Industrial Internet will remain within the confines of corporations and suppliers, rather than the consumer.

A useful analogy might be the recent evolution of Software-Defined Networking (SDN).[8] SDN is a relatively new, powerful innovation in networking that essentially virtualizes network management and resources, doing for the switch and the NOC what server virtualization did for cloud computing and data centers. Invented in 2008, this technology has begun to revolutionize management of data centers and large networks. It could also, eventually, support the management of individual traffic flows all the way down the consumer. By making network management a software issue, it could potentially enable perfect discrimination down to not only the individual but the individual and the application. This could potentially pose serious issues to those concerned with network neutrality and Internet freedom. Yet before activists storm Stanford's networking research group and the R&D departments of networking manufacturers, we have to understand what has to happen before this is a real concern. The technology has been deployed upstream but is far from reaching the consumer. Costs will have to be brought down dramatically, and technology and business models will need to be developed before this technology will touch the consumer space.

Comparing the evolution of IoT to the evolution of Internet technologies actually helps us understand where we are at the moment, and why we need a systemic, architectural perspective. The IoT world today exists as a handful of small islands, and the connectivity between them requires traversal of proprietary networks, different data standards, and domains with different expectations of behavior and control. In many ways, this is akin to the point in Internet history when the majority of traffic was inside specific applications on local area networks, with only a small amount of data spilling out into the Internet.

Just as the Internet is a network of networks, the transformative vision of IoT rests on interactivity and interoperability. Widespread interoperability of different types

---

[7] The "Industrial Internet" is a marketing term coined by GE that addresses the integration of sensors and processing power with traditional industrial equipment, supported by data analytics. https://www.ge.com/stories/industrial-internet

[8] https://www.opennetworking.org/sdn-resources/sdn-definition

of smart devices and sensors will lead *emergence*.[9] Applications not possible in extant distinct silos can be created by bringing different pieces of the network together. While we don't claim to be able to predict the future of IoT innovation, we can look at the raw components available today to try to understand the building blocks to this new world, and how they could fit together.

Architectures of Things

What is an architectural understanding of IoT? It is an abstraction that focuses on the flow of information. We return to the 'net' in the Internet of Things to consider the issue as the flow of data between nodes. We thus need to characterize what these nodes are, and how they are related. By abstracting past specific implementations, we can gain a more systematic understanding of the ecosystem. Once we fully understand this, we can begin to fill in many of the other properties also essential to understanding technology policy, such as why people want information and what will be done with it. But we argue that a key step to building out long term understanding and solutions lies in seeing IoT as a network, enabling a holistic view rather than focusing immediately on specific questions of privacy or security.

What are the merits of this approach? First, it forces us to think about the Things in terms of information collection, processing, and action. We can apply more nuanced components such as personal preferences and data sensitivity later—the key is to understand how many different generic types of Things we might have to deal with. How is a Nest[10] like a Fitbit[11]? How is it different? We can define node properties, focusing on their technical capabilities to deal with information. By focusing on information flow, we abstract to a level higher than counting gates on an integrated circuits, or calculating power requirements. Instead, we can define a series of technical capabilities we're interested in—permanent memory, connectivity, broadcast range, etc. This focus on object capacity will also be useful in the next section, where we explore how to manage this flow of information through insert control points.

An architectural perspective also allows us to make strong statements about the links between objects. Is the connectivity direct, mediated through a local device, or does it pass through a larger network. In many visions of IoT, everything may be able to talk to everything else, but there are no guarantees that this connection will be direct. A small thermocouple, or temperature sensor, with a tiny power source and radio antenna probably won't have the technical capacity to manage a TCP/IP stack, so the smaller devices will depend on intermediary devices.

---

[9] Palfrey, John Gorham, and Urs Gasser. Interop: The promise and perils of highly interconnected systems. Basic Books, 2012.

[10] https://nest.com/thermostat/

[11] http://www.fitbit.com/

A Generalized Architecture

By sketching out the set of potential relationships, we can thus define the superset of all network relationship types, a generalized architecture. The applications discussed above are thus instantiations of some of the elements of this architecture. The generalized architecture will allow us to compare different application or paradigm-specific architectures. But we first hope to define all potential relationships between different types of nodes, and their potential connections.

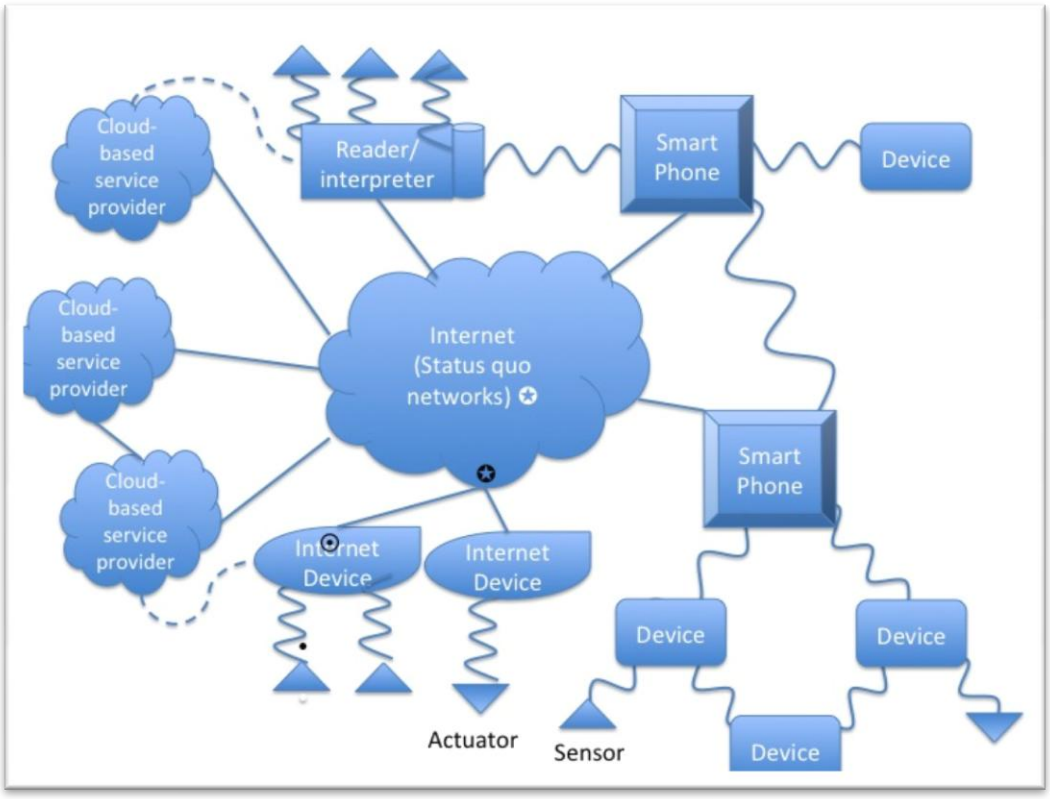

Figure 1 - A Generalized Architecture and communication links networks in the Internet of Things

Figure 1 captures this generalized architectural view of IoT. The nexus of this model is the Internet as we currently understand it: a decentralized, public packet-switched network. This is certainly not necessarily the focal point of every IoT application or instantiation, but we imagine it will play a role in many architectures.

We can then define a set of nodes. We first differentiate between technical nodes—Things—and human nodes—users or organizations. There will be some flows that will enable machine-to-machine communication which characterizes much of the discussion around IoT,[12] and others where we have to consider the potential for an explicit user or organization as a link in the chain.

[12] Atzori, Luigi, Antonio Iera, and Giacomo Morabito. "The internet of things: A survey." Computer networks 54.15 (2010): 2787-2805.

The technical nodes can be further delineated by their role in the infoSmation flow. Simple Sensors are the leaves of our networked tree—they can only collect data from the environment as inputs to the network. Examples of simple sensors include camera, chemical detectors, and accelerometers. But there are other leaves on this tree that also only input information, without processing or routing. Simple RFID tags, for example, are not architecturally different from a thermometer, except the data returned is preset or algorithmically determined, a function of Man rather than Nature. *The defining characteristic is the insertion of data from the physical world into the digital world*.

We differentiate these endpoint Sensors from Devices, which are characterized by their inputs and outputs, as well as other features. Devices allow us a small amount of flexibility to be technology neutral—we can think of them as applications or physical devices with abstracted software, depending on the context. Sometimes, it is necessary to treat different pieces of software running on a shared operating system as (ideally!) separate, while there are other times it is useful to consider the hardware and the application as a single unit. Devices can be further characterized by their technical attributes, such as processing capacity, memory, and battery. Although the rapid pace of technical innovation can make it tempting for us to treat these, to a first approximation, as unlimited resources, this is not always the case. Low power medical devices, for example, monitor sensors via simple algorithms, but must be distinguished from the smart phones we carry in our pockets. Phones may have the capacity to do this as well, but bring the costs, complexities, and risks of a general purpose computing device.

At the other end of information flows are information sinks that use data without generating anything further. Actuators receive information and translate this back into the physical world. We draw a distinction between operations that impact the physical world as an endpoint in the network without further flow and those that enable future action. Entering bits on a drive, or displaying textual output should be thought of as a flow, rather than an output.

Finally, we depart from a purely technical architecture to explicitly incorporate the human element. The human user is critical, of course, but we focus on modeling this individual merely as an agent, rather than assuming explicit roles of "owner" or "data subject." Similarly, we define a class of Third Parties quite broadly to avoid building normative values into the information flow. Either data can be routed to a third party, or it can't. From a network flow perspective, it's important to draw a distinction between two types of Third Parties. A device may send data to another device or web service that it knows, a counter party. If this counter party sends data to another node without an explicit relationship, that is a routed third party. Third Parties can also be a single hop away by intercepting the signals directly.

Control Points – How we get security and privacy

The generalized IoT architecture in Figure 1 illustrates the myriad of ways that data can flow. Different networks can layer on top of each other, and interact or operate independently. From a generativity perspective, this seems to present the opportunity for practically limitless innovation.[13] We are just at the very beginning of this technical era, and experimentation and flexibility could enable the next Big Thing. But the completely open nature reflects a world where the flow of data is determined by whichever node controls it at any given time. Yet this vision is both unrealistic and dangerous. The Internet as we know it today is not such a rule-free place. We have developed a range of mechanisms to help check the flow of data based on predictable security policies.

In the generalized architecture in Figure 2, we superimpose Control Points to ensure that data only flows as it is supposed to. By continuing to use the generalized form, we do not yet have to specify what "supposed to" means, only to understand the potential for control. How can control be exercised on the flow of information? They take the form of functions such as

a. authentication of individuals or programs (or things)

These include passwords, passphrases[14], biometrics[15], physical tokens, and computer identification numbers. (Note that a computer can be as small as a mobile phone.)

b. authorization of individuals or programs (or things)

This is typically done by maintaining one or more tables of authorized entities and the items they are authorized to access (and what types of access are permitted). But other methods have been used in the past including the sharing of capabilities among users, including, for example, software or hardware protection rings of layers of a(n operating) system,.

c. identification of individuals or programs

This is typically a user name, but can also be a number and in some cases a biometric.

d. encryption of data flowing among devices or between users or on the Internet or a local network

---

[13] Thierer, Adam. "Permissionless Innovation." Mercatus Center, 2014

[14] Neil J. Rubenking, "Forget Passwords, use passphrases for extra security" PC Magazine, May 23, 2013, http://www.pcmag.com/article2/0,2817,2419274,00.asp

[15] Ross J. Micheals, Kevin Mangold, Matt Aronoff and Kayee KwongSpecification for WS-Biometric Devices (WS-BD) Version 1: Recommendations of the National Institute of Standards, (Jun 30, 2012)

There are numerous encryption tools and standards available[16]. They have to be used properly (an example of where this was not done is the recent Open SSL breach[17]) and usually one has to rely on a standards body such as the National Institute of Standards and Technology or other trusted source to vouch for the (relative) security of the methods used.

e. logging some or all transactions that take place

This allows for later replay and analysis, signature (code) analysis[18], (near) real-time deep packet inspection[19], and even social network analysis[20], a double-edged sword, both a privacy (-invading) mechanism and a security mechanism, depending on whose eyeglasses one is looking through.

All these mechanisms can be tools in the service of security checks instituted at Control Points after, ideally, a complete risk analysis[21] of the system to be protected is carried out. Many such mechanisms have been around for many years, but are often not put into place since users or administrators haven't experienced such an attack and don't think it will happen to them, and also haven't bothered to do a risk analysis that will in many cases (but not all) make a business case for instituting them. Also, the cost to change systems in place, rather than to limp along with insecure systems, can be quite high. Still, that cost has to be weighed against the legal, administrative, technical, and reputational cost of responding to a data security incident.

Control Point Challenges

Of course, it's not enough to simply lay out a set of tools and declare that they could go anywhere on the network. The right tool has to be used for the right Control Point, and different mechanisms bring their own challenges. For example, authentication and authorization can be conflated which could lead to a control failure if the authorization step occurs before authentication. Other mechanisms offer their own challenges. Encryption is a very powerful tool, but introduces added complexity. If two nodes wish to securely exchange information without having an

---

[16] IEEE Security & Privacy, Special issue on Key Trends in Cryptography, January/February 2015, to appear.

[17] Nicole Perlroth, Experts Find a Door Ajar in an Internet Security Method Thought Safe, New York Times, April 8, 2014,

[18] "Signature Discovery and Threat Detection", Pacific Northwest National Laboratory, http://www.pnl.gov/nationalsecurity/leadership/cybersecurity/feature.stm

[19] "What is Deep Packet Inspection?", PC World, http://www.pcworld.com/article/249137/what_is_deep_packet_inspection_.html

[20] Hanneman, Robert A. and Mark Riddle. 2005. Introduction to social network methods. Riverside, CA: University of California, Riverside (published in digital form at http://faculty.ucr.edu/~hanneman/)

[21] Teodor Sommestad , Mathias Ekstedt , Pontus Johnson, "A probabilistic relational model for security risk analysis", http://dx.doi.org/10.1016/j.cose.2010.02.002, Computers & Security, Volume 29, Issue 6, September 2010, Pages 659–679

established trusted relationship, they will need to use a mechanism like public key cryptography, which imposes an added challenge of key management. As one security expert notes, "Encryption is Easy. Key Management is Hard."[22] Similarly, logging data flows is fairly useless if the auditing mechanism isn't efficient at distinguishing between legitimate and illegitimate data behavior.

The right mechanism has to be used in the right place. What helps determine effective placement? Questions such as cost and risk modeling are not unique to IoT, but there are some issues that are. Many of the Things in our ecosystem don't have much in the way of power, computation, or user interface. A full cryptographic handshake could be out of the question, as could any form of user-generated authenticating password. If a component cannot assume to be connected to the network all the time, then an attacker must be assumed to have offline attack capability.

Take, for example, the challenge of pairing a relatively unsophisticated device with a smart phone. This link may require authentication (the device knows that it is talking to this phone, and not that one) by transmitting a code, or just use the proximity of the two devices, and rely on the lower power of the transmission to limit range, and the user to detect some one attempting to eavesdrop physically.

The choice and location of control points should reflect the relative importance of dynamic vs. static policies. Will anyone need to change the control policies at some point down the road. If there is an evolution in the demands of the network or user preferences, can the control points evolve? If there is a mistake in the code, can the devices be patched? If they can be patched, who's responsibility is it to make sure that components in the network have the updated secure code, and how will counterparties know? Even in the smartphone market today, there is a misalignment in responsibilities between handset manufacturers, software developers and the telephone carriers, leading far too many Americans with phones that are not supported and patched.[23]

But all evolution in policy doesn't necessarily require security and control to be added. As technology (and business models) evolve, it may make more sense to shift control to another point in the architecture, and the network could operate more efficiently if control was loosened at the edge.

This highlights the biggest challenge in mapping control points to an IoT architecture. We do not want too little control, but we also do not want too much. The actual amount of desired control is, in part, a political question about society's

---

[22] Anton Chuvakin and Branden Williams. PCI Compliance. Elsevier Press 2012. P. 127.
[23] Peter Singer and Allan Friedman. *Cybersecurity and Cyberwar: What Everyone Needs to Know.* Oxford Press, 2014. p. 218

tolerance for risk.[24] We also have to consider how the decision to introduce control points is made. A laissez-faire world would make that decision in a purely decentralized fashion. This local view may not allow even a well-intentioned decision-maker to understand the importance of their role in the global flow of data.

Some control points lend themselves to aiding in a global view of data flow. In an electronic medical record environment, for example, with many devices feeding into records, we may wish to limit who has access to most data, even inside a hospital. But given the complexities of roles and responsibilities in a modern healthcare organization with thousands of employees, allowing data to be accessed in a break-glass model with an audit is widely recognized as superior to any form of delayed or forbidden access.[25] On the other hand, a control mechanism in a consumer's smart grid solar array that determines how power flows into or out of the public grid should probably be protected. We might even argue in favor of redundant protection, with strong authentication around the controller, and clear boundary-condition checking at the actuator level to make sure that no action could overwhelm the physical layer.

Having laid out the generalized architecture above and discussed the challenges of general-purpose controls for privacy and security, we now consider how to apply them to specific instantiations. There exist different approaches, visions, or applications of IoT. The flow of information with Nest, for example, is different than a series of small wearable devices that communicate through a smart phone.

Business Models as Constraints

Above, we describe a number of different potential instantiations of the Internet of Things. Which might actually emerge? We can derive some information from the early forays into the space but, as we describe above, the more complete vision has not yet arrived. Different approaches will be tried, and experimentation has been going on in the research community for decades. Yet society-wide integration of an IoT ecosystem will require massive investment. We focus on the incentives for this investment, and argue that the business model is one of the most important things to consider.

First, these architectures are expensive. Compared to transformative web services that have defined the initial web explosion and the rise of an interactive Internet

---

[24] Compare, for example, to recent public comments to the FTC on IoT Privacy, one strongly urging regulatory vigilance and protection (http://www.ftc.gov/sites/default/files/documents/public_comments/2014/01/00016-88256.pdf) and the other demanding a lighter touch in regulation (http://www.ftc.gov/sites/default/files/documents/public_comments/2014/01/00010-88247.pdf).

[25] Break-glass: An approach to granting emergency access to healthcare systems. White paper, Joint NEMA/COCIR/JIRA Security and Privacy Committee (SPC), 2004.

over the past decade, the IoT revolution will be more expensive. These earlier revolutions have not been free, of course. Infrastructure such as servers and data centers and innovations in virtualization and novel programming practices weren't free. But from a marginal cost perspective, two key features will drive cost as a close-to-linear function of the adoption rate. The "Things" discussed above will have real per-device costs to manufacture and distribute. Sensors maybe getting cheaper all the time, but still have a non-trivial cost. The heroic assumptions about the drastic reductions in RFID tags, for example, have not come to complete fruition. At the same time, some sensors might be integrated into existing devices. The binding constraint for inclusion into a smart phone, for example, might be one of space and power, since the hardware will be relatively cheap compared to the entire device. Who will pay for the Things? Once we have a world of things, data has to get into and out of these things. The long term dream may be of ubiquitous data connectivity, but we cannot assume that this will be free without also making other assumptions. Moreover, even if we have devices and connectivity, the challenges of routing require some coordination. A system that assumes that the expanded IPv6 or IPv8 address space will suffice, for example, limits us to an IoT with the primary link through the global internet.

Against these costs, we have to sketch out the space of value. As discussed above, we've seen the greatest recent advances in development and deployment in the "Industrial Internet" space, where the value can be captured by the industrial vertical. There have been instances where the value accrues downstream to the investing party—the emergence of the UPC barcode is a notable example—but that required substantial coordination and cross-subsidization. The two-dimensional bar code, such as a QR code, has been developed but remained on the relative fringes of our information society. (Although a handful of countries, such as Japan, make more use of it than others.) The technical investment is a one-time cost, but the value has to match the costs of deployed infrastructure.

These benefits could all flow to one actor, a series of actors competing in a given space, or the value could accrue to different actors at different points in the architecture. In a mobile device, for example, connectivity has traditionally been seen as a separate value service than the device itself, but this isn't always the case. Amazon subsidized limited 3G connectivity to some of their Kindle ebook readers to drive traffic to their store. One of the most discussed (and feared) IoT value propositions is the commercial value of very rich, detailed data about individuals or Things collected automatically, and refined into useful information, or passed along to third parties who can extract value.

Still, we can make some observations about how benefits accrue in networked systems. The above discussion on the importance of interoperability stresses the importance of standards. In digital ecosystems, particularly in networked ecosystems, the value is a function of how many others can use it. This has led to a

relative rise of winner-take-all competition in the digital economy.[26] This type of competition, in turn, shapes how and when actors will choose to enter the market. When the entry costs are more expensive than the web-world described above, actors may delay, or fight viciously over a space and deter other complementary entries because of uncertainty in the marketplace. Often, order is restored by a standardized platform, and completion can occur above or below these platforms. Microsoft's massive market share in the PC world promoted an explosion of software development on the common operating system platform. The lack of a common standard of usable and reliable Internet identity has held back innovation, and driven the US government to work to encourage private sector cooperation and coordination to address this issue.[27] At the same time, we shouldn't expect a single standard, but should also consider the potential for a series of standardized interfaces across the ecosystem.

So what are these business models? We present four archetypes of business models below, and offer a few examples based on what we see in the market today. The goal here is to illustrate the diversity of potential costs to the firm or firms, a different set of values offered to the user, and different means of revenue. We also describe a pathway to ubiquitous deployment. Each model suggests some means by which early adopters would enter the space, and the IoT ecosystem could grow organically.

How we'll get to the future: IoT Business Models

Stand-Alone Connected Device – We see a growing number of existing devices being transformed by making them a little smarter, and connecting them to the Internet. A VCR with a bit more brains and a lot more storage becomes a DVR. A thermostat that can learn, and allow remote monitoring becomes smart thermostat like Nest, or ecobee. These devices are sold as devices, and can be priced based on their hardware components (the Thermostat) or through service fees. The value for the user tends to emerge from what they *do* rather than the data generated, although innovative use of data to support the core function can be market differentiator. Over time, we may expect to see some cross-integration between different stand-alone devices, but these predictions often feel a little forced: "imagine your refrigerator talking to your toaster!" The path towards adoption will come from the immediate utility of the device, and competition should stay in the application's domain. Thermostats and DVRs compete with themselves for market share, but not with each other.

Vertically Integrated Data Ecosystem – A market actor with an established user base may seek to extend the value of this space into the physical realm. The above example about a smart wearable device integrating face and object recognition

---

[26] Schilling, Melissa A. "Technology success and failure in winner-take-all markets: The impact of learning orientation, timing, and network externalities." Academy of Management Journal 45.2 (2002): 387-398.

[27] Camp, Jean. "Identity Management's Misaligned Incentives." IEEE Security & Privacy 8.6 (2010): 0090-94.

capacity, for example, fits neatly into a possible direction for Google Glass. This plays to the Internet giant's strengths: rapid data queries, processing and online storage. The user already has a relationship with an existing set of tools, and this extends that relationship into a new domain. Value for the customer could initially be built on existing data and applications, with new capacity added over time. Imagine, for example, smart conference badges directly linked to LinkedIn, that can not only track new contacts, but measure social cues and predict success of future relationships.[28] For the firm, value comes from the potential for further growth, as well as the ability to head off disruptive competitors. These are also likely to be environments where data is already being mined to generate value, so new sources can be integrated into that value stream.

Cross-application Pollination – The devices on and in our person have been growing progressively smarter and more capable, but they have largely been developed in isolated silos. A few applications or devices could grow to a large enough size that other devices and applications build on top of them as data platforms or communication links. These popular vendors share APIs to support this innovation. For example, imagine a radio-enabled blood-sugar monitor that could notify a wearable exercise monitor like the Fitbit when exercise might be helpful. The Fitbit could then determine whether it had been a while since you took a walk, and vibrate to remind you to engage in healthy metabolic activity. The Fitbit could also handle the tasks to passing the sensor's data on to an app on a smart phone, which then upload it to that app's cloud-based servers. New devices and applications can be developed more cheaply by piggybacking on existing local devices, services or data, while still demonstrating their unique value to the customer. Being a shared platform confers a market advantage to nexus devices by offering larger suites of complementary devices and services. The nexus may also have access to a larger amount of local data, which could then be exploited for profit. This local area network could be tightly integrated or quite decentralized depending on the context. Pricing would be a function of the devices or data services. It is hard to predict any specific winner in this space, but complementary networked goods tend to follow a tipping point approach.

Industrial Internet Spillover – The Industrial Internet trend discussed above could also be a pathway for diffusion into the broader commercial space. We have assumed the defining aspect of the Industrial Internet is that there are sufficient returns from IoT investment outside the consumer space. Yet the infrastructure could still emerge out beyond the control of the industrial vertical. Imagine RFID tags staying on some products and used downstream, or smart environmental

[28] Wu, L., Waber, B., Aral, S., Brynjolfsson, E., and Pentland, A. (2008) Mining Face-to-Face Interaction Networks using Sociometric Badges: Predicting Productivity in an IT Configuration Task , Proc. Int'l Conf on Information Systems. Paris, France. December 14-17 2008.

sensors being read by public interest organizations as part of an open data movement. Auto fleet transponders could be read by urban planners, and data from a sophisticated RFID lookup database could help consumer protection organizations detect counterfeit products. This creates new stakeholders for existing sensors, for any number of reasons. The relationship between these stakeholders, and the designers of the IoT ecosystem could be good with a positive feedback for innovation and value creation, bad with an iterative game of disruption and adaptation, or non-existent. Since this model is highly dependent on emergent properties of unintentional deployment, building value across specific deployed infrastructure is particularly hard to predict, and seems rather unlikely.

Business Models and Information Architectures

The business models above illustrate the different approaches to value generation, and different paths to a widely adopted ecosystem. Each business model also collects and uses data differently. How tightly coupled are the applications, business models and architectures? We argue that each of the above visions would lead to a different architecture, with different control points.

Stand-alone Devices only demand simple architecture. Each device exists fairly self-sufficiently as a collection of sensors, actuators, storage, and processing, with a connection to a service-specific server through the Internet. This connection may go through a local network or some other first-hop on its route, but the connection is quite simple. Data can be processed locally, or it can all be collected, stored, and accessed in the cloud.

This predictable relationship gives us a fairly straightforward set of control points. We need to worry about the security during the initial path set up. The relationship with the service provider is a key point of trust, but the data collected should be observable and predictable enough to allow a user or regulator to understand the privacy risks involved.

In a *vertically integrated ecosystem*, the data will also flow through relatively predicable paths. Sensors will collect data, and the interpretation will be done locally or remotely, with results sent back to the user.

As in the first case, the data will flow into a server, but now we cannot so easily predict the value of information collected. The Internet links must be secure, and control points may be inserted into directly onto the device. In the Glass example, we might imagine an interactive set of policies about what information might be shared. In the smart name badge example, or with a cheap RFID reader tied to a large lookup dataset, the interface may lack the sophistication, so the control would have to be baked into the device. As far as the risks from information processing, that varies as well. Depending on the size and scope of the data ecosystem, this new data could unlock a vast new set of inferences as data is combined and mined in new ways. Based on the experience of existing information ecosystems, we might expect

this data to be both very useful, and the user to have relatively little control absent some outside intervention.

The dynamics of the information flow in the *Cross-application pollination* scenario are particularly unpredictable, because the local network structure could take so many different forms. One key point of architectural design is the relative centralization or decentralization of each application and device from the smart mobile platform. If it is a wheel-and-spoke design, each sensor, device or app either directly and exclusively connected to the smart platform, or even resident on it. Alternatively, if could be more complex, with devices talking directly to each other without any intermediation. If these are low powered devices with limited processing power, then the routing will be simple and local data processing will be minimal, but it will be harder to impose a defined routing pattern locally while still keeping the environment open for new devices and applications.

A natural control point is the mobile device and its applications that connect the local applications devices to the Internet. But this would be insufficient in some architectures to rely on this interface if the feeds in are not predictable and controllable. This is an architecture where control points may have to be deployed liberally, since the patterns of data sharing are less predictable. On the other hand, because the initial growth hooks on a few killer apps or devices, it might be overly cautious to build controls against all possible data flow. The control patterns could reflect the initial context (healthcare data, or personal habit monitoring, or wearable environmental sensors). While this presents a real risk of the ecosystem evolving and spilling over into other contexts, that transition would probably be visible. This dynamic suggests a good opportunity to implement context-based privacy practices, either voluntarily or through some other accountability mechanism.

Conclusions

There are many useful and important approaches to thinking about IoT privacy and security. If we want to understand how different pieces of data will change what is known about us, for example, then focusing on data collection and analysis is key. This paper took a different approach, following the paradigm of building privacy and security into a nascent technology. To do this, we focused on the flow of data through the architecture of networks, sensors, devices, and servers.

This focus on the flow of data led to an appreciation of the importance of control points, and the very real challenges of understanding what types of controls to put where in the network. Yet the analysis above illustrates a key point: that we don't have to solve the general formula for every conceivable architecture. The IoT future will not emerge out of whole cloth. Different, potentially parallel versions will have to evolve. This evolutionary consideration is necessary, since an IoT network would be too expensive to deploy without value to early adopters today. Evolution and diffusion of technology, as in biology, follows a series of fitness functions. With the

ability to anticipate some aspects of the network architecture, we can derive insights into the placement and attributes of control points.

This exercise combined a technical analysis of network architecture with the organizational and economic analysis around the pathways to technical deployment and diffusion. The technical analysis informs the business side, and vice versa. Together, they inform the broader governance questions of responsibility and control. Understanding how Things interact in a network of networks requires understanding whose things they are, and how they will fit together.