Mechanisms for Meaningful Consent in Internet of Things

***Susan Y.L. Wakenshaw[1]*, Carsten Maple*, mc* Schraefel[†], Richard Gomer [†] *and* Kevin Ghirardello*,**

**University of Warwick, UK, susan.wakenshaw@warwick.ac.uk[1]*
**University of Warwick, UK, cm@warwick.ac.uk*
[†] University of Southampton, UK, mc@ecs.sotonac.uk
[†] University of Southampton, UK, r.gomer@soton.ac.uk
**University of Warwick, UK,* k.ghirardello.1@warwick.ac.uk

Keywords: privacy protection, meaningful consent, Internet of things (IoT), data activities, apparency-p/s transparency

Abstract

Consent is a key measure for privacy protection. Consent has to be 'meaningful' to give people informational power. Individuals need to be provided with real choices and be empowered to negotiate for meaningful consent. Meaningful consent is becoming increasingly important in IoT as privacy is one of the main factors affecting adoption of IoT. Meaningful consent is becoming increasingly challenging in IoT. It is proposed that "apparency, pragmatic/semantic transparency model" adopted for data management could make consent more meaningful, i.e., visible, controllable and understandable [1]. With meaningful consent embedded in the system, users would trust and have a feeling of control that can enhance information sharing which can further support service provision and exploitation of data. The 'apparency, pragmatic/semantic transparency' model has illustrated the why and what issues regarding data management for potential meaningful consent [1]. In this paper, we focus on the 'how' issue, i.e. how to implement the 'apparency, pragmatic/semantic transparency' model' for meaningful consent in IoT [1]. We discuss the three elements such as apparency (by focusing on the interactions and data actions in the IoT system), pragmatic transparency (by centring on the privacy risks, threats of data actions) and semantic transparency (by focusing on the terms and language used by individuals and the experts). This paper contributes to the research on meaningful consent in IoT. We believe that our discussion would elicit more research on 'apparency, pragmatic/semantic transparency' model' in IoT for meaningful consent.

1 Introduction

Consent is a key measure for privacy protection. Consent is one mechanism in the EU data protection regime. Thus, we need to make it as meaningful as possible so that it can fulfil the role that it is supposed to have.

In order to gives people informational power, consent has to be meaningful, i.e., consents have to be intelligible to, controllable by and visible to [(when, if)] users [2]. It is proposed that "apparency and semantic/pragmatic transparency" model could be adopted for data management ("*apparency reflects how an activity is signalled. Semantic transparency addresses whether we know that the terms of the apparent activity (data activity) are and mean; pragmatic transparency reflects the degree to which we know what these data actions actually do or entail*" [1]. This model would enable meaningful consent to be embedded in the system.

In order to have meaningful consent, we need to understand (1) how to make data activity more apparent; (2) how to make user understand and be aware of the risks and implications of these activities and what their consent means/entails; (3) how to make the terms more readable, understandable with standardised, useable and accessible; and (4) due to the scale and speed of data actions in IoT [3], giving end-users real choices and power of negotiation of consent terms and reducing the cognitive and attention burden of consent on the user, through appropriate use of automation or even make the consent automated by learning the users' privacy preferences through the application of AI.

The application of the 'apparency-p/s transparency' model for potential meaningful consent is even more challenging, in particular in IoT. This paper addresses the issues regarding the implementation of the apparency-pragmatic/semantic model for meaningful consent in IoT. We suggest that in order to achieve apparency and pragmatic transparency, we could centre on mapping scenarios of IoT interactions and data flows across multiple systems and between devices; and modelling users' understanding about these systems and the associated risks and the options. In order to enhance semantic transparency, we could use ontology method to develop ontology/corpus representing the language/terms used by different groups and to make the consent terms more understandable, usable and accessible.

In this paper, we would focus on discussing the three elements in 'apparency-semantic/pragmatic transparency model' by using smart home as an example. We believe that the implementation of the model would enable us to develop a

framework for potential meaningful consent for smart devices in smart home environment.

2. Privacy and privacy protection

The concept of privacy is elusive. Many disciplines must deal with the notion of privacy: anthropology, architecture, behavioural psychology, law, sociology, as well as computer science [4]. A taxonomy for privacy was developed by classifying privacy as being 'person-centred' and 'place-centred' (person-environment interactions': (1) private/public dichotomy, (2) an attribute of places and people, (3) as an interpersonal process, (4) a need, right and freedom, (5) an balancing act (balance between social interaction and the risks; risk/rewards as an economic decision) [5]. This privacy taxonomy would enable us to reveal data interactions and data actions and to understand what they really entail and their implications for privacy in IoT. In our paper, we focus on informational privacy but also touch on other conceptions of privacy by considering the taxonomy of privacy [5]. All these notions all ultimately relate to "the boundaries between public and private" [20]. Information privacy has primarily centred on individuals control of acquisition, uses and disclosure of his or her data (a good review see [6]).

With technological advancement, data needs to be internalised and turned into business insights and/or useful information to improve individuals' lives [7]. Moreover, in many situations of everyday life, people need to and want to share information with others. Instead of keep information from accessibility (security) and for secrecy (confidentiality), the focus of privacy work has shifted to how to empower people with choice and informed consent so that they can share the right information (type and amount), with the right people and services, in the sight situations/contexts for their benefits. Privacy after all, entails much more than just control over a data trail, or even a set of data. Privacy could be perceived as a dialectic and dynamic boundary regulation process between the individual (data subject/self), the others (firms and other individuals), and data/information (premise) in contexts [8, 9, 10]. As a dialectic process, privacy could be regulated in situations/contexts such as our own expectations/experiences, those of others with whom we interact and social norms (cultural, social) and regulations (legal). As a dynamic process, privacy could be viewed as being under continuous negotiation and management of (1) disclosure boundary: what (type and amount) information could be disclosed in this context; (2) identity boundary: how much identity related information would be displayed and maintained in this context; (3) temporality boundary: boundaries associated with time, that is, the disclosure and identity boundary depending upon the interpretations of contexts for the past, present and past. Indeed, privacy could be a fluid and malleable notion with a range of trust levels and needs. The boundary regulation could enable the privacy management between the self, others to be appropriate and fair by meeting the expectations, following the rules/norms in time frame to create zones of intimacy and inclusion that shape the relationships with each other.

If privacy is deemed as a boundary regulation process, in which they have to make privacy decisions in terms of information disclosure, i.e., whether and what personal data could be disclosed for the optimal utility. Individual has to make tradeoffs. In this research area, research has centred on relationship between perceived control over personal information and willingness to disclose by taking into account of the benefits, the costs, and the risks. Indeed, privacy decisions are the result of trade-offs. With technological development, privacy-related trade-offs are increasingly difficult to see and resolves. It is suggested that 'control' and 'notices' are used as instruments for privacy protection. However, these instruments might not be sufficient. For example, control may backfire because it may lead individuals to reveal more information in risky situations. Notice may not be effective enough to communicate the risks associated with the information disclosure. Privacy notice (privacy policies) can be too long and complex to be comprehensible for the average users. Notification mechanisms do not consider the user limitations and biases and therefore are not effective [11]. Therefore, research on how to communicate/present the risks, the policies and develop effective notification mechanisms is urgently needed in order to enable users to make effective privacy decisions.

3. Meaningful consent

Consent is one mechanism to protect user's privacy. Consent is a mechanism that, notionally at least, ensures that a data subject is aware of, and agrees to, data processing – and privacy is one concern (among others) that might influence that decision. Consent could be claimed on the basis of information disclosure made through privacy policies, cookie notices and terms and conditions (Ts&Cs) on the Internet. However, consent in the EU for data protection purposes is legally distinct from Ts&Cs which invoke a weaker concept of consent taken from contract law. In order to have consent, we need to give end users better, readable, understandable with standardised, useable and accessible presentations; empower end users by giving them real choices, negotiability; friend's choices or crowd choices; keep it transparent, through automated term analysis; giving expert advice or third party certification; auto-consent via preference model (Recognize standard term packages; keep track of what I accepted before; tool to show only what is different); Mandatory user protection (Par Lannero of Common Terms.org. 2015). Consents have to be intelligible to, controllable by and visible to [(when, if)] users [12].

In addition to the data for identified person such as employee to employer, student to the school, data can be collected for directly and indirectly identifiable person through devices and software by a variety of mechanisms such as browser cookies or fingerprinting [13], the information about the users' web

browsing history, other information resources such as social networking profiles [12]. In the GDPR, these types of data are the categories of personal data which the organisations process. Organisations would process these data under the following conditions: consent, contract between the organisation and the individual ('data subject'), contract between the individual and someone requiring the organisation to process the personal data; A law or obligation, when someone's personal interests are at stake, in public interests or when acting under official public authority, and in "legitimate interests" of the organisation ("Data Controller") or the individual (http://missinfogeek.net/gdpr-consent/). Consent is a legal basis, and in many ways is the basis of "last resort" - organisations will rely on legitimate interest wherever possible. The only area where consent is required in relation to data types is in the special categories (sensitive data in UK terms) – like biometrics, religious beliefs etc. However, individual users could be tracked in an adversarial context. Moreover, information generated in these devices could be left in the data pipe, which might remain there for decades. Users are not even aware of these data and the privacy concerns related to these data. In addition, users could give consent to these data they shared with firms. However, how about the historical data they have shared with these firms. It is warned that most users are not able to infer the consequences of data collection and processes by service providers … sometime what was entailed by the practice themselves…" [14, 12]. In IoT, these issues become even more acute. In IoT, privacy protection entails not only the data-oriented privacy protection but also context-oriented privacy protection. The former centres on protecting the privacy of data content. The latter focuses on protecting contextual information such as the location, timing information of traffic transmitted in the network [15].

Professor mc Schraefel and her colleagues such as Richard Gomer at the University of Southampton have been working in the domain of meaningful consent. They proposed that "consent is a state of mind in which somebody decided they are ok with X happening; X need to match what will actually happen (transparency); X should include known risks or side effects of what is being proposed; consent is signalled to someone-often by pressing a button and the party who relies on that signal uses it as evidence of consent (the mental state in another person)". They argue that meaningful consent must move towards apparency and semantic/pragmatic transparency regarding how data is managed in order to have meaningful consent [1]. For them, "apparency reflects how an activity is signalled. Semantic transparency addresses whether we know that the terms of the apparent activity (data activity) are and mean; pragmatic transparency reflects the degree to which we know what these data actions actually do or entail" [1]. Thus, Apparency entails making the data processes 'apparent'. Apparency can be achieved by 'signalling' the data activity [1]. However, apparency for properties for consent decisions can be variable, dynamic and identification and designing of the effective signalling is very challenging. For example, a project called Web Mirror (http://mirror.websci.net/) could mirror back to students their browsing history and browsing activity (what they browse) is their personal data. Apparency also seeks to make the connection clear in data activity such as between how and what data is used) and why it is used and make it traceable toward meaningful transparency [1]. Pragmatic transparency entails what these data actions actually do and entail (p.29). Indeed, pragmatic transparency entails the implications of these data actions. Semantic transparency entails what these terms for describing these actions really mean.

4. "Apparency-pragmatic/semantic transparency" model for meaningful consent in IoT -smart home as an example

Due to the complexity of data actions, privacy risks and implications of consent in IoT, meaningful consent is even more challenging. These challenges include: how to make data activity more apparent, traceable and better signalled; how to make the connection between data actions (what, how and why) transparent (apparency); how to present what these data actions entail and mean and how to make the user be aware of the risks and implications of these activities and what their consent means/entails (apparency/pragmatic transparency); how to make the terms used to describe actions, connections and implications more readable, understandable with standardised, useable and accessible (apparency/semantic transparency). Moreover, due to the scale and speed of data actions in IoT, it is crucial to understand the sensitive point where people really want to give the opportunity to say yes or no. It is also crucial to have the default privacy setting which needs to come pre-configured in a way that people are happy with most of the time. Individuals need to give the real choices and be empowered to negotiate the terms of consent with firms in the IoT network.

The Apparancy-P/S transparency model illustrated why and how issues regarding data management and thus would potentially make meaningful consent. However, with the increased importance and complexity of meaningful consent in IoT, it is urgent needed to address the 'how' issue, i.e., how to implement this model theoretically and empirically. In this paper, we attempt to address this issue by focusing on the three elements of the model theoretically. We would use the smart home as an example to illustrate some of the viewpoints.

4.1 Apparency

In IoT, transparency usually refers to the fact that the terms and conditions of use of a service, the privacy policy of how data may be used are explicitly stated. Based on these terms, we can consent to engage with a device/service. We know from copious related work that this is a kind of false transparency as few of us read the T&C and fewer of us understand them. And even if we did put in this effort T&'C's do not make clear what is happening with our data and how it might be used by third parties in particular. We have referred to this level of interrogation as Apparency. We

need to enable users to understand the interactions and data actions within an IoT ecosystem. Based on this, the T&Ss would be more apparent and meaningful for users.

In order to achieve this, we need to design scenarios enabling us to model users' IoT interactions and data actions within these interactions. These could include the interactions between the user and the device, the interactions between devices, and interactions between the service provider and the devices. Data actions involved in these interactions can be modelled such as data flow across devices and data processing. Without these flows and interactions being apparent, the users cannot truly consent to the data flowing into the pool of these ecosystems [1]. To have consent, we also need greater apparency of how data is being used as a result of our consent [1] (p.33). The autonomous agents could keep the users more informed about these interactions with IoT systems and signals the sensitive points for consent and also signals the adversary/abnormal data actions. Thus, apparancy entails (1) making interactions in the IoT ecosystem apparent, (2) making data flow across devices apparent, (3) data actions (processing and usage) apparent, and (4) making signalling these interactions, flows and data actions apparent.

Interactions in the IoT ecosystem

In order to address issues regarding interactions and data actions, we need to understand the interactions in the environment of smart home. Smart home was not a new phenomenon. A decade ago, Augusto and McCullagh (2007) discussed smart home as an application of Ambient Intelligence (AmI) [16]. It is suggested that AmI is the combination of all these resources such as networks, sensors, Human Computer Interfaces, Pervasive Ubiquitous Computing and Artificial Intelligence (AI) to provide flexible and intelligent services to uses acting in their environments [13]. Maple (2017) argued that "the internet of things (IoT) is a technology that has the capacity to revolutionise the way that we live" [3, p.155]. Indeed, it is argued that IoT is "a technological phenomenon originating from innovative developments and concepts in information and communication technology associated with: Ubiquitous Communication/Connectivity, Pervasive Computing and Ambient Intelligence" [17] (Dohr et al, 2010, p.804). Thus, we can argue that the basic idea for ambient assisted living is shared across AmI and IoT: using technology such as sensors and devices interconnected in a system for taking decisions or enhancing decisions to benefits the users based on real-time information gathered and historical data accumulated [13]. Thus, we suggest that the design principle of AmI (Brooks, 2003) [18] would enhance our understanding of interactions within the IoT system at smart home.

According to Brooks (2003) [18] and Augusto and McCullagh (2007) [16], in order to design a smart home, we need to understand how the user interacts with the environment in the daily life by exploring the five "W" including: who, where, when, what and why. First, we need to know who use the system and the role of the user in the system in relation to other users. Other users here not only include the human but also other elements and objects/devices such pets, robots and objects of interest within the environment [16]. We need to track where the user and others are located in the environment at each moment during the system operation. We need to know when and duration of interactions and association of activities (changing location) take place to model the dynamic and evolution of these interactions and activities. We need to understand what activities and tasks users are performing temporally and spatially. We need to know why these activities and tasks are performed (intentions and goals). With these understanding, we could provide appropriate help in a sensible way to assist the users' life [16]. Indeed, through this analysis, we could understand the interactions between the user and the objects, activities and tasks performed, intentions and goals and services needed in the environment. We can model the scenarios and mapping these interactions in the smart home system. We can also understand what data is needed and how data is used and for what purposes to provide the support/help. When we mapped these interactions, we can signal the user in terms of data actions and the services. At smart home, On one side the IoT-enabled assisted living aims to act as a passive human assistant by observing activities and inferring situations and user needs to help users when (and only if required). On the other hand, users may directly interact with the devices and system to indicate their needs and preferences. These interactions between the user and these autonomous agents/assistant and the interactions between the user and the system can be signalled to the user.

Data flow across devices

In order to address the second issue, we need to map the flow of data in the smart home system.

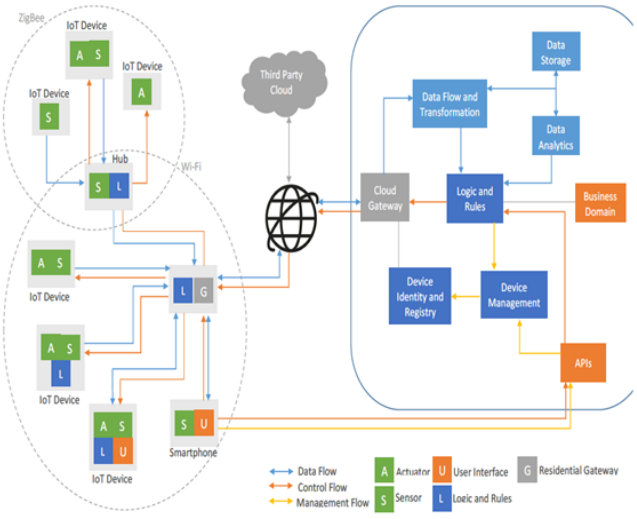

Figure 1: Example of smart home architecture.

In this architecture, a series of different IoT Devices connected to a router directly through Wi-Fi or indirectly through an IoT hub. Each device with a sensor produces data as it observes its environment and uses its connection to a router to send it to the remote cloud servers. Depending on the information and its configurations, devices with the Logic and Rules modules may make certain control decisions locally, at which point specific actions are taken by actuators. Otherwise, the residential gateway collects data and prepares to forward it to the Cloud, at which point the data is transferred from the Cloud Gateway to the Storage, Analytics or Logic and Rules. This last module may receive data which is "raw" or processed by the Analytics module, and consequently determines the appropriate command. Once the Logic and Rules module creates a device actuation command, it is forwarded to the Cloud Gateway, which then connects to the Residential Gateway and sends the command to an Actuator. In case the created command involves a device or service not managed by the same cloud platform, which is often the case in smart home environments, then the Logic and Rules module can make use of APIs to connect to the appropriate Third-Party Cloud. Furthermore, device interaction may also happen at a local level through certain IoT hubs. This is one example of the smart home architecture. It could enable us to map the data flow and data processing and the points of data flow and data processing can be signalled to the user for their consent.

4.2 Pragmatic transparency

In IoT, based on the mapping of interactions and data actions across multiple systems, we need to explore what these interactions and actions really entail and their implications for the user if they give consent to these data actions and data uses. We can call them values-apparency. We also need to think about if attack/adversary actions take place, what privacy harm can be done to the individual with these consent. To have consent, we need greater transparency of the implications/risks/potential harms as a result of our consent. In IoT, the challenge is how to signal these new properties for greater apparency. Issues of privacy threats and privacy protection in IoT have attracted much attention in research. Privacy threats and protection could be data-oriented and context oriented [15, 19].

We suggest that Parkerian Hexad's six fundamental attributes of information can be considered when we examine threats and adversary actions. These attributes include:

- Confidentiality ensures that data is not made available to unauthorized individuals, entities or programs);
- Integrity guards against improper information modification or destruction maintaining integrity of a data and systems;
- Availability is a property which ensures that the data and security ecosystem is fully available when required;
- Authenticity refers to the assurance that a message, transaction, or other exchange of information is from the source it claims to be from;
- Utility refers to how useful the data is to the user;
- Possession or control refers to the physical disposition of the hardware in which the data is stored.

Smart devices

The first aspect of pragmatic transparency is to do with the smart devices. To be apparent about what entails in the data actions with these devices is crucial for users to understand what their consent would lead to. For example, smart medical devices could be one of the smart devices at smart home. These smart medical devices have great potential to enable patients and their doctors to monitor the patients' health. But there are also potential privacy-related threats to these devices. For example, Kotz (2011) [19] identified three threats to users' privacy (defined as 'individuals' rights to control the acquisition, uses and disclosure of his or her identifiable data'). A threat is defined as 'the possibility that his/her right to control his personal data is weakened or eliminated due to erroneous or malicious actions' [19]. These threats include:

- Identity threats: lose or share their identity credentials, enabling others to access to their personal health data and personal health record). This threat entails misuse patients' identities by the insiders and outsiders.
- Access threats: unauthorised access to personal health information (in the medical network or the personal health record). Personal health information can be modified for insurance fraud or malice by the insiders and outsiders.
- Disclosure threats: unauthorised disclosure of personal health information including data at rest or data in transit. This is an example of the potential threats at the device level.

At smart home, there are many devices, the privacy threats and risks need to be examined and signalled to the users for these smart devices to enable them to have meaningful consent.

IoT system

The second aspect of pragmatic transparency is to do with (1) what the smart home system really entail, (2) what are the threats and risk to the system, (3) what are the privacy threats/risks if attacks on these systems take place, and (4) the implications of these attacks. Figure 1 has illustrated the data flow in the IoT system for smart home. We can see that there are many attack surface exposed to adversary actions. These attacks could be:

- Spoofing – impersonation of device or network;

- Tampering – manipulation of certain parameters over the data being sent;
- Repudiation - ability of entities/ users to deny the actions performed;
- Information Disclosure - An information disclosure attack results in an application revealing sensitive information to the attacker;
- Denial of Service – attacker seeks to make a machine or network resource unavailable to its intended users by temporarily or indefinitely disrupting services;
- Elevation of Privilege – gaining elevated access to resources that are protected from an application / user.

The potential attacks on the attack surfaces need to be analysed the potential risks could be signalled to the user if we want to give them choices for meaningful consent.

Privacy harms of data actions

The third aspect of the pragmatic transparency is to do with the privacy in general. Solove (2005)'s seminal paper on the taxonomy of privacy would provide a guidance to explore what data actions in particular adversary actions in IoT really impinge upon privacy. Solove's taxonomy identified and summarized the problematic activities in information collection, information processing, information dissemination, and invasion [20].

- Information collection

One problematic form of information collection is surveillance. According to Solove (2005) [20], direct awareness of surveillance makes people feel uncomfortable and cause that person to change her/his behavior. "Surveillance can lead to self-censorship and inhibition. [ ] and thus could be a tool of social control" (p.493). Too much social control can adversely affect freedom, creativity and self-development (p.494). The data collected through surveillance is significantly beyond any originally sought. If lasting long enough individuals might be caught in some form of illegal or immoral activity, which can be used to discredit or blackmail him/her (p.495). It is suggested that "in United States v. Karo, the Court concluded that a tracking device that monitored a person's movements within his home implicated that person's reasonable expectation of privacy" (p.496). Surveillance is harmful in all settings. Thus, in the smart home, the user needs to understand how much his behavior is tracked and he/she can make decisions about it.

- Information processing

Information processing entails the use, storage and manipulating the data. Five forms of information processing were discussed in [19] including: aggregation, (2) identification, (3) insecurity, (4) secondary use and (5) exclusion. Alongside benefits, these forms of data processing can be problematic. For example, aggregation can cause dignitary harms. People give out a bit of information in different settings. But the aggregator would acquire much greater knowledge about the person's life by consolidating these pieces of information (p.507). When data are collected they are disconnected from their original contexts and also data are reductive and incomplete, this could lead to distortion. It is described that "some courts have recognized that aggregation as violating a privacy interests" (p.508). Identification is connecting the information to the individual. Identification attached information baggage to people. Identification can inhibit one's ability to remain anonymous or pseudonymous (p.513). Insecurity is a problem caused by the way our information is handled and protected. Insecurity is related to data aggregation issues, identification issues and identity theft issues. Distortion is related to insecurity- the dissemination of false information about a person (p.477). Second use can cause problems. It causes dignitary harms because the data is used in ways in which the person does not consent and might not find preferable (p.477). Second uses generates fear and uncertainty over how one's information will be used in the future and create sense of powerless and vulnerability. Removed from the original context and consented purpose for use, data could be misunderstood (p.477). Exclusion refers to 'the failure to provide individuals with notice and input about their records" (p.521). People should be provided with notices about the use of their personal data and give them rights to access and correct it (p.521). Exclusion can cause a sense of uncertainty and vulnerability in the individuals, powerlessness and frustration (p.521). Exclusion breaches confidentiality (p.522).

- Information dissemination

The forms of information dissemination could cause privacy harms including (1) breach of confidentiality, (2) disclosure, (3) exposure, (4) increased accessibility, (5) blackmail, (6) appropriation, and (7) distortion. The harm caused by breaching confidentiality includes information disclosure and victim being betrayed (p.525). Disclosure entails the reveal of the true information about a person to others. Disclosure could damage the reputation of the individual when the information is disseminated (p.529). Disclosure could threat people's security. It is argued that "people want to protect their information that make them vulnerable or that can be used by others to harm them physically, emotionally, financially and reputationally" (p.530). Disclosure can make other judgement of a person distorted (p. 530). Exposure refers to "the exposing to others certain physical and emotional attributes about a person" (p.533). Exposure of these attributes could create embarrassment and humiliation. Exposure involves revealing some attributes we have been socialized into concealing these activities that we possibly find animal-like or disgusting (p.534). Exposure could strip people of their dignity (p.535). Increased accessibility could cause problems of disclosure (p.537). It is suggested that "Blackmail involves coercing an individual by threatening to expose her personal secrets if she/he does not accede to the demands of the blackmailer" (p.539). The harm of blackmail is due to the control exercised by the one who make the

threats over the data subject. Blackmail related to disclosure, exposure, breach of confidentiality (p.540). It is described that "appropriation is the use of one's identity or personality for the purposes and goals of another" and "appropriation involves the way an individual desires to present him/herself to the society" (p.543). The harm of appropriation involves an interference of freedom and self-development (p.544). Distortion is 'the manipulation of the way a person is perceived and judged by others, and involves the victim being inaccurately exposed to the public" (p.547). Distortion involves revealing the false and misleading information (p.547).

- Intrusion

A grouping of privacy harms is labelled 'intrusion' (p.548). "Intrusion involves invasions or incursions into one's life. It disturbs the victim's daily activities, alter her routines, destroy her solitude, and often makes her feel uncomfortable and uneasy" (p.549).

Even though these privacy harms described by Solove (2005) [20] are in generic terms and are not specific applied to the IoT context. In the IoT, in particular smart home setting, there are potential risks for these harms. When we model the interactions, data flow and data processing these harms could provide guidance for us to fully grasp what these actions, activities and practices really entail and make it apparent for the users, i.e. risk apparency [1].

4.3 Semantic transparency

The pragmatic transparency discussed above identified the potential threats, risks and harms for privacy in IoT. These threats, risks and harms need to be presented in terms/language which is understandable and accessible. Otherwise, the pragmatic transparency and apparency cannot be achieved. This is crucial for empowering users. Only can they understand can they not only accept them but also are able to make choices and negotiate their terms.

Language for privacy

In order to achieve this, we first need to review what language is used to describe privacy. The language of privacy used in the context of video media was investigated [18]. Privacy was decomposed into "three normative controls for regulating interpersonal boundaries in an embodied dialectic: solitude, confidentiality and autonomy". They have developed vocabulary of terms to describe many interrelated and subtle meanings of privacy. The vocabulary includes the terms for (1) solitude, (2) confidentiality, (3) autonomy, and (4) mechanics for privacy [21].

(1) Vocabulary terms for solitude:

Physical Dimensions	Psychological Dimensions	Presentation Dimensions
i) Interpersonal Distance (1) isolation to crowding ii) Attention (1) focus to periphery	i) Interaction to Withdrawal (1) anonymity and reserve to intimacy ii) Escape (1) refuge (2) fantasy	i) High-level Awareness (1) availability (2) accessibility ii) Distraction (1) relevance (2) salience

(2) Vocabulary terms for CONFIDENTIALITY

Information Channels	Information Characteristics	(Information Operations
i) Medium (1) aural (2) visual (3) numeric (4) textual ii) Processing (1) sampling (2) interpolation (3) aggregation (4) inference iii) Topic (1) information about the self (2) personally identifying information (3) activities (4) whereabouts (5) encounters (6) utterances (7) actions (8) relationships	i) Basic Characteristics (1) sensitivity (2) persistence (3) transitivity ii) Fidelity (1) precision (2) accuracy (3)misinformation (4) disinformation iii) Certainty (1) plausible deniability (2) ambiguity	i) Basic Operations (1) capture (2) archival (3) edit ii) Intention / Use (1) accountability (2) misappropriation (3) misuse iii) Scrutiny (1) surreptitious surveillance (2) analysis

(3) Vocabulary terms for AUTONOMY

Social Constructions of the Self	Social Environment
Social Constructions of the Self i) Front (1) identity (2) digital persona (3) appearance (4) impression (5) personal space ii) Back (1) flaws (2) deviance* (3) idealisations iii) Signifiers* (1) territory (2) props (3) costumes iv) Harms (1) aesthetic (2) strategic	i) Social relationships (1) roles (2) power (3) obligations (4) status divisions (5) trust ii) Norms (1) expectations (2) preferences (3) social acceptability (4) conformance (5) deviance (6) place

(4) Vocabulary terms for MECHANICS for PRIVACY

Boundaries	Process Characteristics	Violations	Behavioural and Cognitive Phenomena	Environmental Support
i)disclosure ii)temporal iii) spatial iv) identity	i) dialectic ii) dynamic iii) regulation iv)cooperation	i) risk ii)possibility iii)probability iv) severity v) threat	i) self-appropriation ii) genres of disclosure iii) policing iv) reprimand v) reward vi) risk/reward trade-off vii) disclosure boundary tension viii) disinformation* ix) reserve* x) Signifiers* (1) implicit (2) explicit	i) situated action ii) reflexive interpretability of action iii) constraints iv) transitions v) choice vi) reciprocity vii) liberty viii) refuge* ix) Embodiments (1) rich to impoverished x) Cues (1) feedback (2) feed-through

(Source: Boyle et al, 2009) [21].

Boyle et al (2009) provided the language for privacy. We also need to develop the language for other privacy related issues such as privacy harms, risks, attacks, and threats [18].

Ontology for privacy

The vocabulary for privacy [21] includes the terms used by the academic community and used by experts. The vocabulary would be different from the vocabulary used by the users of the IoT. We need to develop corpus/vocabulary which could represent different groups. We suggest that we could use ontology engineering method [22] to develop a corpus /ontology to represent the terms of these two groups. Neches et al. (1991) define ontologies as "the basic terms and relations comprising the vocabulary of a topic area" (p.40) [23]. Ontology is defined as "a formal representation of knowledge as a set of concepts within a domain, and the relationships between those concepts" [24]. In addition to these formal representations, within a domain, it is argue that knowledge of different groups needs to be represented for information exchange and coordination [22]. This could be achieved by building a large corpus of related concepts, i.e. a large collection of possible related terms [22]. These concepts should represent knowledge of different groups in the domain, from experts to ordinary people. In order to construct such a corpus, the first step is to identify the key word sets (the basic terms comprising the vocabulary of a topic area/domain). A Delphi approach is used to collect a small number of words for a subject area from domain experts. The second step is using the key word sets as seeding words to produce more related concepts for a large corpus construction. At this stage, the seeding words are paired and linked to knowledge bases such as existing ontology and Google search engine to generate semantically-related terms from the initial seeding words. The purpose of the Google search is to derive terms representing the knowledge at social and cultural levels not limited to the domain experts. The methodology for ontology construction includes: (1) data source selection; (2) seeding word configuration; (3) seeding word selection; (4) corpus construction. Any reader interested please reads the paper on the methodology [22].

By building the ontology of privacy and ontology for privacy related concepts such as privacy harms, privacy risks, and threats by using the ontological engineering method, we could develop the vocabulary which is more understandable, accessible for both the experts and the users.

The framework for understanding data management is illustrated in Figure 2, which can potentially enable users to have for meaningful consent in IoT.

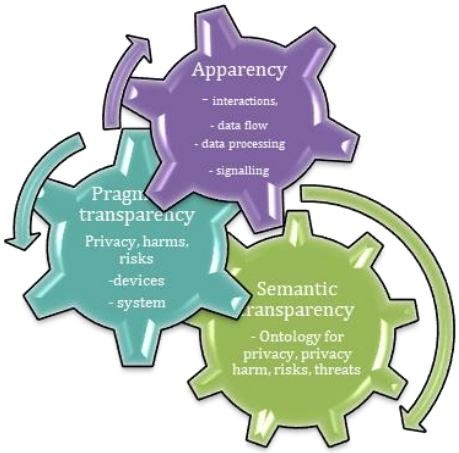

Figure 2: Framework of Apparency-P/S transparency for meaningful consent

5. Conclusion

The apparency-pragmatic/semantic transparency model of how data is actually being used could enable users to consent in a meaningful way [1]. "Having strong, clear apparency to real semantic and pragmatic transparency as a backbone to meaningful consent will help clarify risks within data flows of large-scale heterogeneous IoT infrastructure, from homes to cities to national infrastructure" [1] (p.33). This model provides the why and what need to be modelled to provide choices in understandable accessible way to have a meaningful consent in IoT. However, how to apply this model in IoT is very challenging. In our paper, we focused on the HOW issues by discussing about how to address these three components in the model. We believe that our discussion would further elicit more research on these issues and on meaningful consent. By addressing these issues, meaningful consent is highly possible. When meaningful consent is achieved and become part of the IoT system, the customers would be empowered to make choices for data sharing and data would be used in a consented way. The value of data would be leveraged in IoT.

Meaningful consent entails the decisions to protect or surrender privacy. However, individuals are very likely to be uncertain about how much information to share [25]. Meaningful consent mechanisms illustrated by the Apparency-P/S Transparency model in our paper would potentially make the consequences of privacy behaviour (related to information sharing) more tangible. This would greatly influence privacy behaviour. However, the privacy decision making is "only in part of the result of a rational 'calculus of costs and benefits [26, 27]" [25]. These tradeoffs can be affected by many factors such as "(mis)-perceptions of those costs and benefits, as well as social norms, emotions and heuristics" (p.510) [25]. Moreover risks could also be

personal and contextual. For instance, it is suggested that "present-bias can cause even privacy-conscious to engage in risk revelations of information, if the immediate gratification from disclosure trumps the delayed and hence discounted, future consequences" (p.510) [25]. Therefore, there is an inherent tension with the risks identified by techniques and the users' perceived risks in contexts. These factors exacerbate the difficulty of ascertaining the potential consequences of privacy decisions. These factors could make meaningful consent very difficult. These issues need to be addressed in future research on meaningful consent.

Acknowledgements

The work informing this paper is supported by the Engineering and Physical Science Research Council UK projects: ComPaTrIoTs Research Hub (PETRAS); Control and Trust as Moderating Mechanisms in addressing Vulnerability for the Design of Business and Economic Models (ConTriVE)

References

[1] Schraefel, m.c., Gomer, R., Alan, A., Gerding, E. and Maple, C. "The internet of things: interaction challenges to meaningful consent at scale". *interactions*, *24*(6), pp.26-33, (2017).

[2] Baarslag, T., Alan, A.T., Gomer, R.C., Liccardi, I., Marreiros, H. and Gerding, E.H., mc schraefel, "*Negotiation as an interaction mechanism for deciding app permissions," in ACM SIGCHI Conference Extended Abstract on Human Factors in Computing Systems*. (2016)

[3] Maple, C., 2017. Security and privacy in the Internet of Things. *Journal of Cyber Policy*, *2*(2), pp.155-184.

[4] Boyle, M., & Greenberg, S. "The language of privacy: Learning from video media space analysis and design". *ACM Transactions on Computer-Human Interaction (TOCHI)*, *12*(2), 328-370, (2005).

[5] Newell, P.B. "Perspectives on privacy". *Journal of environmental psychology*, *15*(2), pp.87-104, (1995).

[6] Smith HJ, Dinev T, Xu H. "Information privacy research: an interdisciplinary review". *MIS quarterly*. 35(4):989-1016, (2011).

[7] Ng, I.C. and Wakenshaw, S.Y., 2017. The Internet-of-Things: Review and research directions. *International Journal of Research in Marketing*, *34*(1), pp.3-21.

[8] Altman, I. "Privacy regulation: Culturally universal or culturally specific?". *Journal of social issues*, *33*(3), pp.66-84, (1997).

[9] Petronio, S. "*Boundaries of privacy: Dialectics of Disclosure"*. Suny Press. (2012).

[10] Perera, C., Wakenshaw, S.Y., Baarslag, T., Haddadi, H., Bandara, A.K., Mortier, R., Crabtree, A., Ng, I.C., McAuley, D. and Crowcroft, J., 2017. Valorising the IoT databox: creating value for everyone. *Transactions on Emerging Telecommunications Technologies*, *28*(1).

[11] Acquisti, A., Adjerid, I. and Brandimarte, L. "Gone in 15 seconds: The limits of privacy transparency and control". *IEEE Security & Privacy*, *11*(4), pp.72-74, (2013).

[12] Gomer, R. "Consentful Surveillance: Supporting User Understanding and Control", (2016).

[13] Mayer, J.R. and Mitchell, J.C. "Third-Party Web Tracking : Policy and Technology". *Proceedings of the 2012 IEEE Symposium on Security and Privacy*, 413–427, (2012).

[14] Marreiros, H., Gomer, R., and Tonin, M. "Exploring user perceptions of online privacy disclosures". Proceedings of 14th International Conference on WWW/INTERNET 2015, IADIS (2015).

[15] Li, N., Zhang, N., Das, S.K. and Thuraisingham, B., "Privacy preservation in wireless sensor networks: A state-of-the-art survey". *Ad Hoc Networks*, *7*(8), pp.1501-1514, 2009.

[16] Augusto, J.C. and McCullagh, P. "Ambient intelligence: Concepts and applications".*Computer Science and Information Systems*, *4*(1), pp.1-27, (2007).

[17] Dohr, A., Modre-Opsrian, R., Drobics, M., Hayn, D. and Schreier, G., 2010, April. "The internet of things for ambient assisted living". In *Information Technology: New Generations (ITNG), 2010 Seventh International Conference on* (pp. 804-809). Ieee (2010).

[18] K. Brooks. "The context quintet: narrative elements applied to context awareness". In *Proceedings of the International Conference on Human Computer Interaction (HCI 2003)*. Erlbaum Associates, Inc., 2003.

[19] Kotz, D. "A threat taxonomy for mHealth Privacy". In *Communication Systems and Networks (COMSNETS), 2011 Third International Conference on* (pp. 1-6). IEEE, (2011).

[20] Solove, D.J. "A taxonomy of privacy". *U. Pa. L. Rev.*, *154*, p.477, (2005).

[21] Boyle, M., Neustaedter, C. and Greenberg, S."Privacy factors in video-based media spaces". *Media Space 20+ Years of Mediated Life*, pp.97-122, (2009).

[22] Ma, X., Bal, J. & Issa, A. “A fast and economic ontology engineering approach towards improving capability matching: Application to an online engineering collaborative platform”. *Computers in Industry*, *65*(9), pp.1264-1275, (2014).

[23] Neches, R., Fikes, R.E., Finin, T., Gruber, T., Patil, R., Senator, T. & Swartout, W.R. “Enabling technology for knowledge sharing”. *AI magazine*, 12(3), pp.36, (1991).

[24] Fensel, D., Van Harmelen, F., Horrocks, I., McGuinness, D.L. & Patel-Schneider, P.F., “OIL: An ontology infrastructure for the semantic web”. *IEEE intelligent systems*, *16*(2), pp.38-45, (2001).

[25] Acquisti, A., Brandimarte, L. & Loewenstein, G., “Privacy and human behavior in the age of information”. *Science*, 347(6221), pp.509-514, (2015).

[26] Laufer, R.S. & Wolfe, M., “Privacy as a concept and a social issue: A multidimensional developmental theory”. *Journal of social Issues*, *33*(3), pp.22-42, (1977).

[27] Klopfer, P.H. & Rubenstein, D.I.. “The concept privacy and its biological basis”. *Journal of social Issues*, *33*(3), pp.52-65, (1977).