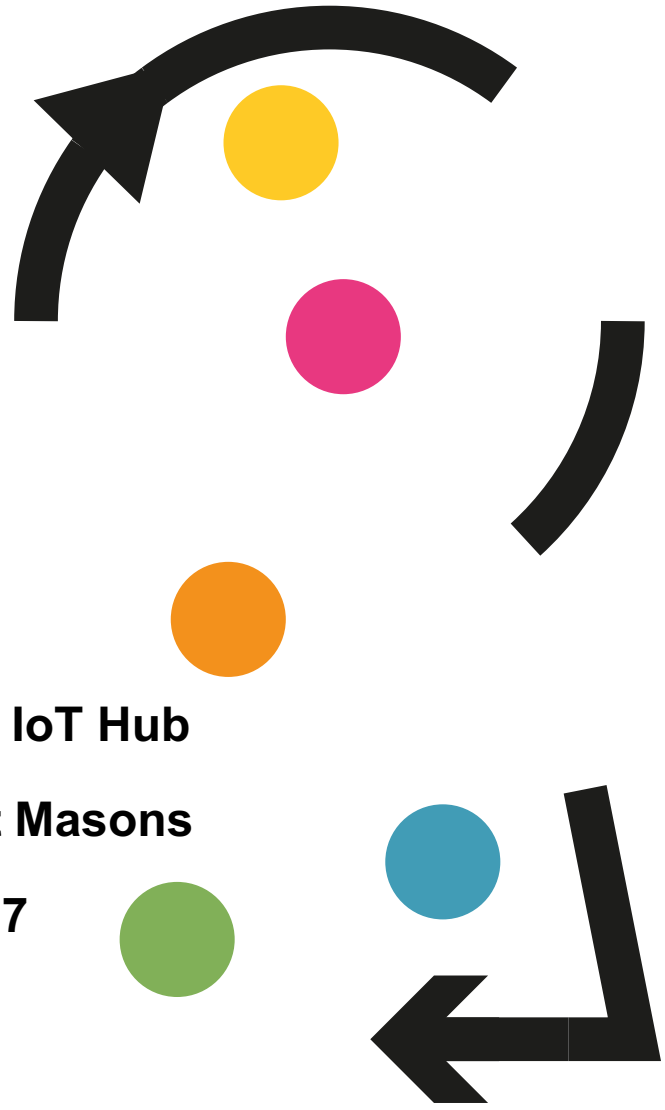
IoT and Its Implications for Informed Consent



PETRAS IoT Hub

Pinsent Masons

25 July 2017

Please cite as:
Tanczer, L., Carr, M., Brass, I., & Blackstock, J. (2017). IoT and Its Implications for Informed Consent. PETRAS IoT Hub, STEaPP: London.

Executive Summary

This report is based on a three-hour long workshop between representatives of the PETRAS IoT Hub, Pinsent Masons, and the HMG Department for Transport. The workshop is part of an ongoing investigation that explores the connections between some of the different dimensions likely to shape conceptions and applications of consent in the emerging Internet of Things (IoT). The impetus for the workshop was the recognition that two significant developments will challenge conventional approaches to online consent. From a technical perspective, the IoT will significantly increase personal data collection, use and re-use. From a regulatory perspective, the General Data Protection Regulation (GDPR) which comes into force in May 2018, will make much higher demands on practices of giving and obtaining consent. Combined, these two factors suggest that consent will be a major issue for all actors in the next five years and it requires some careful analysis now in order to adequately prepare for these developments.

The workshop focused on three particular scenarios, including (a) connected and autonomous vehicles (CAVs); (b) IoT medical health devices; and (c) the built environment. The discussion concentrated on the actors involved in the IoT space and the data sources, purposes, and flows that characterise the IoT data cycle. Predefined scenarios guided the conversation and were used to incentivise a debate amongst workshop attendees. The report summarises the discussion that took place and will guide further research in this realm.

The following are preliminary findings that derived from the workshop:

i. Opaque data cycles and data transfers characterise the current IoT environment, resulting in a **lack of transparency and traceability** of data flows for data subjects;
ii. **Technical and regulatory tools** to resolve consent challenges in the IoT ecosystem are needed to ensure that the right to privacy and data protection principles are upheld.
iii. One of the significant challenges to receiving consent in the IoT will be the difficulty of **informing a data subject** about all the different processes and purposes that are taking place.
iv. Although use cases differ in many ways, there are important parallels between them. There is scope for thinking about these similarities in more depth and applying a **consistent framework** that can be transported into different IoT realms thereby developing a generic **model of nodes**.
v. It is important to differentiate between "**devices**" and "**services**". To understand consent in the IoT realm, where the "device / service" is not created by a single business entity, it is critical to understand these relationships clearly.

This report provides an account of the discussion that took place and concludes with a number of proposed research questions and action points intended to further develop our research agenda into consent in the IoT.

Table of Contents

TABLE OF CONTENTS

Workshop Background

Participants

Leonie Tanczer (University College London)
Madeline Carr (Cardiff University)
Irina Brass (University College London)
Sarah Cameron (Pinsent Masons)
Alex Barnes (Pinsent Masons)
Cerys Wyn Davies (Pinsent Masons)
Stephan Appt (Pinsent Masons)
Carsten Maple (University of Warwick)
m.c. schraefel (University of Southampton)
James Lovesey (Centre for Connected and Autonomous Vehicles, DfT)

Objectives

This workshop explored the connections between some of the different dimensions likely to shape notions of consent in the emerging IoT ecosystem. Within the IoT environment, individuals are, to varying degrees and under different circumstances, less aware of their data-providing interactions than they might be when engaging with the Internet in more conventional ways (Kang, Dabbish, Fruchter, & Kiesler, 2015; Weinberg, Milne, Andonova, & Hajjat, 2015). Thus, we may anticipate that in some circumstances, consent may be neither clearly understood nor explicitly given. In addition, those collecting and handling data may face difficulty in ensuring that consent given at the point of collection is respected all the way through the data cycle. The 2018 implementation of the General Data Protection Regulation (GDPR) will raise the expectations of these consent interactions at the same time that complying in the IoT appears to introduce further complexity. These and many other challenges highlighted to us the need to engage in a closer examination of the changes to the understanding of and practices for giving and obtaining consent.

Some of the unique characteristics of the underlying consent processes, as well as some of the possible implications for future behaviours and decision-making were discussed. These included questions on:

i. degrees of individual **awareness**;
ii. data capture **circumstances**;
iii. data processing **purposes**;
iv. **ownership** of system components;
v. risks to the **individual**;
vi. **risks** to other actors;
vii. **liability**;
viii. **legal mechanisms** and their fit for purpose.

Methodology

The three-hour long workshop was based on a scenario exercise which aimed to:

i. Reflect on whether the principle of consent, defined as "a clear affirmative act establishing a freely given, specific, informed and unambiguous indication of the data subject's agreement to the processing of personal data related to him or her […]" (GDPR, para 32) will be **able to hold in legal practice in the context of IoT, both when the GDPR comes into effect next year and potentially 5-10 years** from now when the IoT becomes more pervasive;
ii. Identify possible **data cycles in different contexts**, i.e. sources, channels, processing nodes, end use destinations;
iii. Identify different **agents involved** in creating and using these data flows;
iv. Identify diverse **agent motivations** throughout the data cycles, i.e. motivations in data processing and aggregation;
v. Identify circumstances in which **gaining consent will be problematic** or cumbersome;
vi. Explore **possible risks** (e.g., liability, privacy) and **wider implications** (e.g. for regulation, legal structures, IoT design, etc.);
vii. Produce content and insight that can be used to construct **illustrative vignettes of issues of IoT and consent** that can be used as engagement devices with other partners / stakeholders.

The workshops did not commence with pre-concluded scenarios (sometimes referred to as 'narratives'). Research shows that at an early exploration stage, a premature scenario development would put at risk any truly novel insights because it can inhibit the contributions of expert participants by being too prescriptive.

Therefore, a rough scenario outline was provided which was based on a graphical depiction of potential data flows (see: Figure 2, Figure 3, Figure 4, and Appendix A). Figures were substituted with clear guiding questions to structure the conversation, with some illustrative starting points included to maximise efficiency and use of participants' time. The scenarios consequently encouraged a structured reflection on IoT's implications for consent and included:

i. **Mobility - autonomous vehicles**: Where the data subject is representative of users with average level of familiarity with IoT;
ii. **Home - medical devices**: Where the data subject is representative of users who are (short-term or long-term) highly vulnerable as a result of personal circumstances and in need of decision-making support;
iii. **Infrastructure – built environment**: Where the data subject is representative of high engagement but low understanding or awareness of data and identity protection considerations.

Definitions

The workshop kicked off with a necessary discussion about foundational definitions and conceptual terminologies.

Consent

While consent may be broadly understood as the permission for something to happen, because of our interest in regulatory analysis, participants intentionally framed the workshop in terms of the guidelines set out in the General Data Protection Regulation (GDPR).

The GDPR (para 32) defines **consent** as "a clear affirmative act establishing a freely given, specific, informed and unambiguous indication of the data subject's agreement to the processing of personal data relating to him or her, such as by a written statement, including by electronic means, or an oral statement"[1].

The regulation also specifies the conditions in which consent can be given and includes "ticking a box when visiting an internet website, choosing technical settings for information society services or another statement or conduct which clearly indicates in this context the data subject's acceptance of the proposed processing of his or her personal data". According to the GDPR, silence, pre-ticked boxes, or inactivity should therefore not constitute consent (para 32).

Further important criteria set out in the GDPR, include that:

i. Consent should cover all processing activities. Thus, "when the processing has multiple purposes, consent should be given for all of them" (para 32);
ii. If consent is to be given by electronic means, "the request must be clear, concise and not unnecessarily disruptive to the use of the service for which it is provided" (para 32);
iii. Consent "should not be regarded as freely given if the data subject has no genuine or free choice or is unable to refuse or withdraw consent without detriment" (para 42);
iv. In cases where there might be an imbalance between the data subject and the controller (e.g. controller is a public authority), consent "is presumed not to be freely given if it does not allow separate consent to be given to different personal data processing operations despite it being appropriate in the individual case, or if the performance of a contract, including the provision of a service, is

[1] It should be noted that consent is currently defined as "any freely given specific and informed indication of [the data subject's] wishes by which the data subject signifies his agreement to personal data relating to him being processes" (Directive 95/46/EC, Art 2(h)).

dependent on the consent despite such consent not being necessary for such performance" (para 43).

It is important to note that consent is only one of six 'conditions' under which **lawful processing** of personal data can take place. Art 6(1) GDPR refers to these conditions, which apply when:

i. the data subject has given **consent** to the processing of his or her personal data for one or more specific purposes (a);
ii. processing is necessary for the performance of a **contract** to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract;
iii. processing is necessary for **compliance with a legal obligation** to which the controller is subject;
iv. processing is necessary in order to protect the **vital interests** of the data subject or of another natural person;
v. processing is necessary for the performance of a task carried out in the **public interest** or in the exercise of official authority vested in the controller;
vi. processing is necessary for the purposes of the **legitimate interests**[2] pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child.

For the purpose of the workshop and this report, emphasis was given to consent, rather than other lawful bases for the processing of personal data.

Data Subject

The notion of **data subject** (i.e., a term used in the GDPR) was another point that required clarification. The Information Commissioner's Office (ICO; 2017) defines data subject as "an individual who is the subject of personal data". The workshop participants agreed on the understanding that it encompasses "an individual whose data is being collected and processes and who has rights and should not be deprived of the protection of personal data".

Personal / Identifiable Data

Personal data is any information that one can relate to an individual who can be identified. Art 4(a) (GDPR) defines personal data as "any information relating to an

[2] Legitimate interest shall not apply to processing carried out by public authorities in the performance of their tasks.

identified or identifiable natural person ('data subject')". This includes patterns of their behaviour or any other personal characteristics.

Under the Data Protection Act 1998 the definition of what constitutes **identifiable data** was rather narrow. A person was considered to be identifiable dependent on the information in the possession of the data controller. Under the GDPR, the scope of personal, and thus identifiable data, has expanded.

Art 4(1) GDPR clarifies this, conceptualising an identifiable natural person as "one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, **location data**, an **online identifier** or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person."

Due to this definition, one will have to take into account information within the remit of anyone (i.e., not just the data controller) that could identify a person. Significantly then, personal data is no longer restricted to particular data types.

As an example, Alex referred to the implications of an employee ID. A person outside of the company would not be able to match the employee ID to a particular individual. An employee ID would consequently only be personal data for the data controller – which in this example, is the employer. However, according to the GDPR, any data that *can be combined with other data sources* and may then lead to the identification of individuals is to be regarded as personal data. This applies to unique identifiers such as employee ID numbers, telephone numbers or social identity numbers as well as common data sources such as age.

Cerys gave another example of how information about the drainage system of a house could become personal data because it constitutes information about the way an individual uses a property. Hence, even building and utilities information can fall into this category when it is linked with the information about a person's use and habits.

Any of these data sources could, through the association with other information, conceivably lead to identification. The concepts of personal and identifiable data are therefore inherently interlinked with the **aggregation and correlation of data**.

Scenario 1: Mobility - Autonomous Vehicles

Following this brief conceptual discussion, participants moved on to the scenario exercise on connected and autonomous vehicles (CAVs).

Initial Discussion

Overview by Stephan Appt

The mobility scenario was accompanied by a presentation from Stephan Appt. Stephan is attorney-at-law for Pinsent Masons in Munich and focuses on the automotive and IT sector. He outlined some of the data and consent issues he has encountered in the course of his work with industry actors.

Personal Data

According to Stephan, German original equipment manufacturers (OEMs) struggle with the definition of "**what constitutes personal data?**". German data protection authorities take a very broad understanding of personal data (see discussion above), considering that there would always be a way to identify the owner of the car (i.e., through the aggregation of datasets and metadata and due to the fact that most data sets collected from cars carry the vehicle identification number ("VIN")). This broad understanding triggers data protection laws that OEMs now have to comply with and for many use cases outside the framework of legitimate interest places a responsibility on manufactures to obtain consent for the processing of data.

Access to the Data Subjects

A concrete challenge OEMs therefore encounter is that **they often have no practical relationship with the end-customers / data subjects** (unless they establish such relationship through online customer portals like "MercedesMe" etc. where users register and determine the settings for connected car functionalities etc., see below). Instead, it is the car dealers that have a direct relationship with the data subjects. Yet, according to Stephan, many dealers are not willing to share customer information with OEMs or at least demand a financial reward for asking the customers for consent on behalf of the OEMs.

What would help OEMs is that platforms and portals have been established where customers can sign up to customer accounts, enabling manufactures to develop their own relationships with end-users. A direct line to the customer has consequently been established and this allows OEMs, where possible, to ask data subjects for consent. These portals also permit data subjects to receive an overview of data flows and amend functionalities. A recent example is BMW's CarData portal: With BMW CarData, a car owner can view the key vehicle data for his/her BMW whenever he/she wants and

share them with third parties, if required. The user alone decides who he/she wants to share the data with and maintains control of his/her telematics data at all times and benefits from sharing data with third parties to use tailor-made service packages that were not previously accessible. One mouse click is all it takes to allow or refuse data sharing with the individual third party.

There are numerous technical solutions proposed for the access to in-vehicle data and resources (e.g. on board application platform, in-vehicle interface, various data server approaches) which all come with various issues with regard to obtaining consent – further detail being provided in EC Final report (2017a) on "Access to In-vehicle Data and Resources".

Although these communication channels may be a mechanism for giving and obtaining consent between the OEMs and the car owners, one issue that continues to be difficult is the acquiring of consent needed from **independent third-parties who engage in the road traffic**. These include actors such as passengers in the vehicle or pedestrians outside the vehicle who will still remain unapproachable (and possibly unaware) under this model.

Communicating Consent to Data Subjects

In addition to these problems of establishing a relationship between data subjects and data controllers, an additional concern is the comprehensible transmission of information to data subjects. There is potentially **too much information that needs to be communicated to individuals** (e.g., what data is being collected; how is data transferred; who has access to the data; how long is data stored etc.). In the scenario of CAVs, OEMs are currently restricted to the limited space of a car dashboard / human machine interface screen (HMI). It remains therefore unclear how this kind of information will be exchanged satisfying legal requirements for transparency. OEMs also express concern that due to the limited display, they may even be open to accusations of hiding important information by not presenting it in adequate detail.

Video Cameras

Stephan explained that a particular challenge to German OEMs is the use of video cameras in CAVs. Video cameras are integral to the design and operation of CAVs because they feed back traffic and environmental data essential to navigation. However, in Germany, **the monitoring of public spaces is permitted** only under very strict conditions. As video cameras are tools to monitor the surrounding environment and as an individual unconnected to the vehicle (a pedestrian, for example) cannot verify who is monitoring the area (i.e., who the car/camera owner is), some data protection authorities argue that the use of video cameras in CAVs is strictly speaking not in line with current legal requirements.

It has been suggested that one option to tackle this problem would be to abstain from relying on live camera footage, but collecting **pixelated footage** instead. Nonetheless, there are some that say this is not possible as:

i. the car needs to learn from the data set in order to identify what particular objects look like; and
ii. live camera footage will be required for evidence purposes in the instance of a car accident (especially where the accident occurred in a jurisdiction where court proceedings require a jury to consider provided evidence which typically is much more responsive to live footage).

Stephan said that there was the hope that the video camera issue might be resolved with the implementation of the GDPR, where video usage is not explicitly being addressed. However, new German data protection law that has been enacted for local legislation surprisingly provides a provision on camera usage in public spaces. Thus, for Stephan this German law is arguably challengeable.

Many of the privacy issues the automotive industry are facing need to be reviewed against the background of the overarching question **on how to effectively anonymize data** in the IoT context.

For OEMs, it remains therefore unclear:

i. Which data and data flows require consent in a CAV?
ii. Who are all the actors that consent must be obtained from?
iii. Through which mechanisms can they manage consent and user preferences?
iv. How can access to in-vehicle data be managed and will OEMs eventually be forced to share data based on the current EU data economy / free flow of data discussion (see: European Commission, 2017b)?[3]

Data ID / Passport

At the end of Stephan's presentation, he suggested the use of a form of "**passport**" or "**Data ID**". Data subjects could use this identification tool to connect with their car and any other IoT device in order to determine what data can and cannot be shared and processed.

Stephan therefore encouraged the attendees to think about a universal approach of framing consent issues in the IoT ecosystem, promoting the development of a tool that would allow data subjects to carry their settings with them. Otherwise, Stephan

[3] The EU data economy debate is focussing on non-personal data but a data producer right is proposed which would, if applicable, ultimately require obtaining some form of consent as well.

suggested, the sheer volume of consent agreements required under these two conditions of IoT and GDPR could lead to people being overwhelmed by requests for consent.

Meaningful Consent Project

Leading on Stephan's presentation and in particular, his calls for a 'data passport' solution, mc introduced the **Meaningful Consent in the Digital Economy** project. This project is developing a technical solution in the form of a fully automated tool for managing consent processes (see for example: Baarslag et al., 2017; schraefel & Gerding, 2013). In their conception, consent is designed to become a separate process independent to the usage of IoT devices and services. An individual validates certain attributes and privacy preferences, which then become part of an automated process that learns to make decisions on behalf of the user based on previously indicated settings.

The Meaningful Consent project situates itself within a body of research that shows that users have an aversion to being interrupted in the course of their interaction with technologies. Studies highlight that requesting consent from data subjects through terms and conditions in-between a person's attempt to conduct a primary task (e.g., uploading pictures to social media) is an inappropriate time to ask an individual for consent. Consent becomes a secondary task; it stands between the user and their intended action. The "I Agree" button that is so essential for businesses and their compliance to data protection laws, has, she suggests, become the "Go Away" button for users. Thus, the button serves two different purposes which have become conflated; giving consent and clearing the way for the user to move to their primary task.

The project also addresses the fact that terms and conditions are currently **non-negotiable**. Current consent solutions do not account for deviations and remain fixed in binary "yes / no" terms but this 'conditionality of consent' will no longer be possible under the GDPR and the **Privacy and Electronic Communications Regulations (PECR)**. The Meaningful Consent Project will provide the capacity to negotiate terms and conditions and thus allow for the expectations set out in the GDPR to be implemented.

In order for consent to become meaningful and informed, the project aims to manage consent in such a way that it enables two properties:

i. Consent becomes an **autonomous process** that can be managed for a person by an 'agent' on their behalf;
ii. Consent becomes the **basis of a negotiation** and is no longer dependent on the binary understanding of approval and denial.

The tool they are developing provides individuals the opportunity to audit consent preferences at a time when it is not a critical task (i.e., not at the point when an

individual wants to upload photos). It ensures that consent becomes an activity of its own and receives the attention it requires, with **validation and verification** becoming important concepts as opposed to identification.

Workshop participants observed that this approach does not conflict with the notion of "clear, affirmative" consent as necessitated in the GDPR. Clear, affirmative consent does not apply to a particular timing and can consequently be given at an earlier time. In addition, the GDPR does not specify that consent cannot be given by devolving your responsibility to an agent and may therefore also be carried out by algorithms.

The project can be situated in a growing body of work that studies dynamic consent mechanisms. Neise et al. (2015) showed how users could give consent by selecting profiles associated to pseudonyms when subscribing to services and checking-in smart spaces. Their model draws on the "**Event-Condition-Action**" (ECA) structure, where when an event is triggered, the condition (C) is evaluated and an action (A) thereafter executed. And in an earlier example of this type of thinking, the EnCoRe project developed the idea of **sticky policies** which would allow users to set privacy preferences and have them travel across multiple parties and services (Pearson & Casassa-Mont, 2011).

All of these studies further attempt to accommodate the fact that data preferences can change dynamically and any mechanism for managing consent therefore has to be adaptable. Indeed, mc argued that their research found that users are **more willing to share data if they perceive that their preferences are being addressed**.

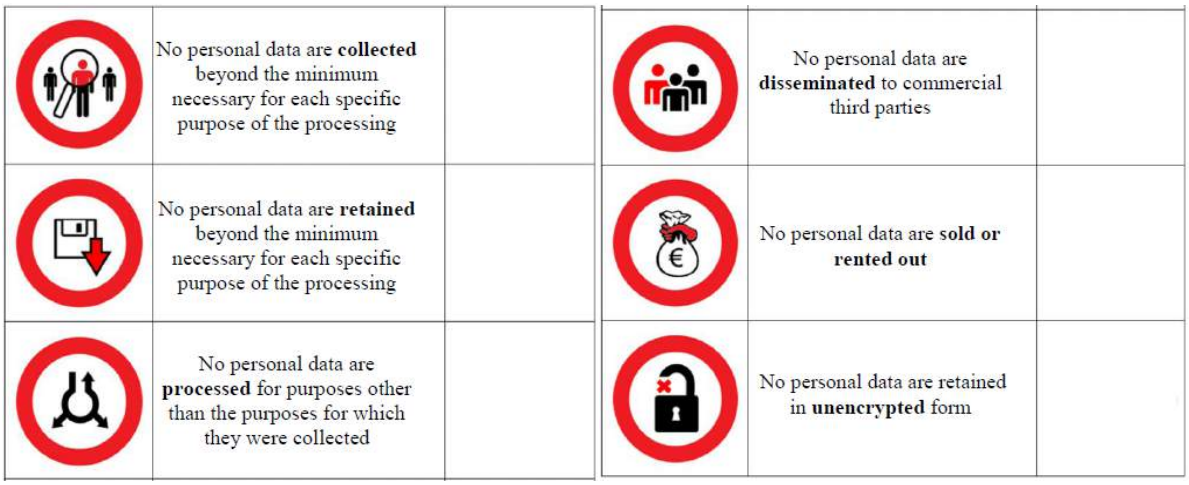
Standardised Icons

The conversation now shifted to the usage of standardised privacy policy icons which were initially suggested and developed by Alexander Alvaro, former Vice-President of the European Parliament (ENISA, 2013; Fischer-Hübner, Angulo, & Pulls, 2013).

The icons shown in Figure 1 are intended to simplify the communication of data flows and address the legal transparency requirement set out in the GDPR. They are meant to disclose to data subjects how data controllers are collecting and storing data. The European Data Protection Board could be delegated to approve certain icons in order to enhance the public's understanding about what particular institutions are doing with their information.

Attendees noticed the resemblance to road signs, and criticised the culturally narrow, gender biased, and dated (e.g., floppy disk) depiction. It was felt that the icons would lack tangibility with participants struggling to comprehend their meaning without an explanatory description. mc therefore critically remarked "that sometimes one word is worth a thousand icons".

Figure 1: Proposed Icons



Note: Icons suggested by Alexander Alvaro (ENISA, 2013) and graphic adapted from Data Science Innovation (2015).

Data Cycle Discussion

The workshop shifted its focus to the scenario exercise. Attendees were asked to discuss the content outlined in Figure 2. As explained in the methodology, the exercise intentionally did not include all potential data points and data flows. The lack of graphical detail was meant to enhance the debate amongst attendees and guide the conversation rather than constraining it. As intended, participants therefore began the activity by inquiring about the position and connection of actors and other agents.

Scope of Exercise

Carsten felt that the data cycle shown on the poster was very restrictive to roads. He therefore sought to clarify the **scope of the exercise**, pointing to the difference between CAVs / vehicles – which is a broader term that includes drones, planes, trains, lorries – and cars. The latter was agreed to be the core focus of this exercise.

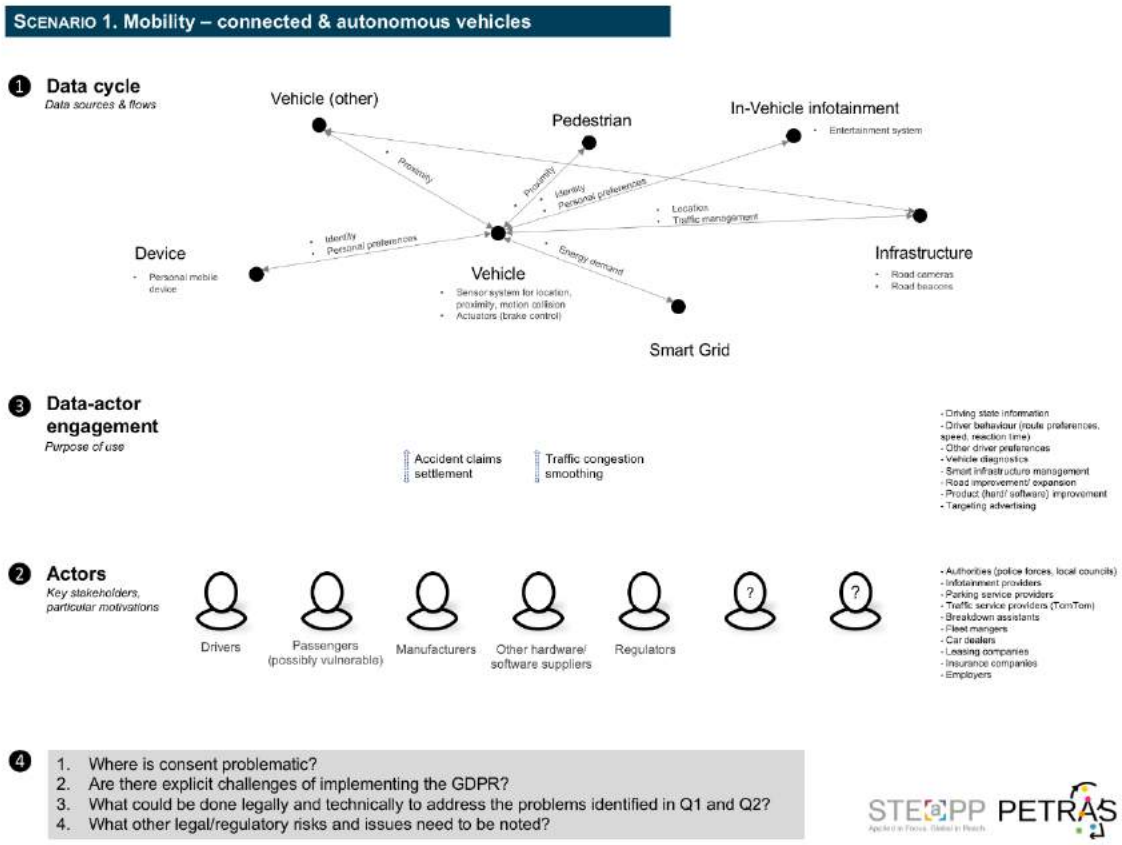
Drivers and Passengers

Additionally, attendees felt that the data point "vehicle" should incorporate **both drivers and passengers**, which would also apply to the data point "vehicles (other)". This differentiation would be necessary as passengers might not give the same form and extent of consent as their drivers.

Participants also perceived the need to acknowledge "other actors" who are not necessarily pedestrians (e.g., cyclists; and less or more mobile actors) in the graphic. This could extend and/or complement the link **vehicle-to-pedestrian** (**V2P)**. Thus, an

updated data cycle design would need to add different types of actors that embody specific roles and characteristics in the road traffic.

Figure 2: Scenario Exercise



Drivers and Passengers

Additionally, attendees felt that the data point “vehicle” should incorporate **both drivers and passengers**, which would also apply to the data point “vehicles (other)”. This differentiation would be necessary as passengers might not give the same form and extent of consent as their drivers.

Participants also perceived the need to acknowledge “other actors” who are not necessarily pedestrians (e.g., cyclists; and less or more mobile actors) in the graphic. This could extend and/or complement the link **vehicle-to-pedestrian** (**V2P)**. Thus, an updated data cycle design would need to add different types of actors that embody specific roles and characteristics in the road traffic.

Consent Issues Raised

Opaqueness versus Transparency

Carsten pointed out that when driving a car, one passes a lot of different infrastructures with different data processors and collectors. This creates **opaque data cycles and transfers**. At present, both the knowledge of and the graphical interface for these to be displayed to the data subject in a transparent way are not available.

To highlight this opacity, Carsten referred to his experience with Uber. The transportation service unexpectedly audio recorded and photographed him when entering a car. When he became aware of this, he consulted the Uber driver who pointed him to the terms and conditions. While Carsten had consented to using the service, he had not been fully informed about the recording practice. The Uber driver was in charge of the collected data and how he/she would use it. The storage of the gathered information consequently posed an additional attack surface, which in the IoT ecosystem is heightened due to the sheer length of the data flow and the lack of traceability.

Irina linked the opacity of data flows in the IoT back to the GDPR. The GDPR sets out clear guidelines on the processing of data for multiple purposes. Hence, when information is collected for the purpose of traffic management and also targeted advertising for example, the data subject is expected to consent to each of these purposes. This is expected to increase the **transparency** of these data transfer transactions.

Based on this understanding, Cerys suggested that the bigger challenge to receiving consent may actually be the difficulty of informing the data subject about all the different processes and purposes that are taking place.

Nested Services

A subsequent discussion that emerged around the topic of opaqueness was the idea of nested services. While not specific to CAVs, the **reliance on software provided from third-party services** might create additional, unintended data flows that users are unaware of and have not consented to.

Madeline shared an anecdote of a colleague having developed an app to explain how magnets work to school children. When attempting to download and rate the app, it asked for a wide range of access permissions (e.g., photos, audio recordings, videos). She consequently inquired about its purpose, with her colleague responding that the collection of this data was never his intention. He was dependent on a free, third-party development tool when building the app. This widely used development tool had the requests pre-installed and deprived him of the opportunity to engage in **data minimisation**.

This echoes a study conducted at the International Computer Science Institute where researchers tested 5,000 of the most popular children's apps for their compliance with the United States Children's Online Privacy Protection Act (COPPA). The researchers identified that many developers whose apps fail to protect data often do so unknowingly. Many fail to configure their software properly or neglect to scrutinise ready-to-use code from different third-parties that they are relying on (Egelman, 2017).

In this regard, research and regulators will have to engage with software engineers to examine their practices, cultures, and tools in more depth. While **data minimisation is discussed in the GDPR on a regulatory level, the actual implementation on a practical level lacks certainty**.

It is therefore crucial to develop and provide software frameworks that are compliant with data protection laws and do not require software engineers to continually reassess the tools they use. Hence, it will be important to look carefully at **existing software libraries and platforms** and ensure that the principle of data minimisation and a sufficient level of consent is upheld.

The Value of Data

Data subjects also need to be informed about the value of their own data. Because there is currently no mechanism for users to measure this, most users cannot **comprehend its financial worth**. Carsten anticipates that when people realise how much revenue their data could generate, they may be more willing to pay for services.

Conversely to the data subjects' lack of knowledge, OEMs and other IoT manufacturers are very aware of the potential to monetize digitally generated information. One participant therefore argued that there is a "**land grab for data**", even though many businesses do not yet know how much these records will add to their revenue.

Responsibility and Liability

Building on the discussion of the monetization of data, workshop attendees noted that commercial actors often lack understanding about the **responsibility, liability and cost** that comes with the collection of this data.

Cerys highlighted that clients have only recently started to realise that they are faced with costs associated with storing and protecting large quantities of historic data that they have collected over the years. With the GDPR coming into force, they now have the opportunity to **either delete or anonymise the data** if they wish to continue saving and using it. The GDPR therefore, seems to be incentivising a reconsideration of economics of data capture and storage amongst some of Pinsent Masons' clients.

Reasonable Expectations

A final point that was brought up in the course of this exercise was the notion that data collection must not only happen with the consent of individuals but within people's reasonable expectations.

Cerys explained that some charities have been fined as a consequence of the way they have been collecting personal data from public sources, mining it for information that would help them to understand individuals' wealth, and using it to identify likely donors. The ICO has argued that this practice would **not be within people's reasonable expectations** and is therefore prohibited. This has implications for many organisations that may have previously searched online for publically available personal data in order to supplement their records.

Participants observed that the judgment might be useful for people who struggle to control their data and those who lack a full understanding of the implications of their data being publicly available.

Cerys added that the **decision might also be subject to change**. Certain sectors might argue that this practice is within people's reasonable expectations in these certain instances or that it is dependent on **where the data is coming from** e.g., from an individual or a separate agent.

Scenario 2: Home - Medical Devices

Following the extensive discussion that arose from the CAV exercise, workshop attendees discussed Figure 3 that was centred on IoT medical devices. The conversation began with an analysis of data flow similarities and differences between this case and the CAVs case, in particular in relation to the scope of actors involved in the data cycle. The conversation then moved on to consent considerations that emerge from particular actors and their respective interests in this domain.

Data Cycle Discussion

Carsten stressed the disparity between the data cycle presented in Figure 3 compared to Figure 2. Although there would be many **similarities** to the mobility example, these were not sufficiently expressed in the graphic. These parallels include, for example, cloud providers and insurers who also apply to the CAV scenario but were absent in Figure 2. He suggested that it would be worth thinking about these similarities in more depth and applying a **consistent framework** that can be transported into different IoT realms. This would allow for the development of a generic **model of nodes** (e.g., node1 = person, node2 = piece of kit, node3 = service etc.) for the IoT ecosystem.

The data flow graphic should also differentiate between "**devices**" and "**services**". In the IoT-realm, products like health monitors, but cars in particular, are systems of systems. They encompass many different services. Carsten referred to a paper that explained how many stakeholders were receiving information from various fitness devices. The average number of services involved in the information exchanged equalled 5 and rose up to 27 in certain circumstances. These figures can be extrapolated for complex systems such as vehicles, where the number could be significantly higher and where the "device / service" is not created by a single business entity. Thus, it is extremely important to convey these relationships through the graphical display of the data flow.

Carers

Cerys directed the conversation to **carers**. She provided the example of clients working in the health realm who are expected to engage with carers as much as with patients. In these contexts, carers become authorised agents that are able to share personal data in the IoT ecosystem. Carers communicate with patients and might send information to third-parties. Conversely, they may also be the person receiving information from these third-parties (e.g., medical practitioners). Hence a graphical display has to account for carers being situated between those two agents.

Relatives

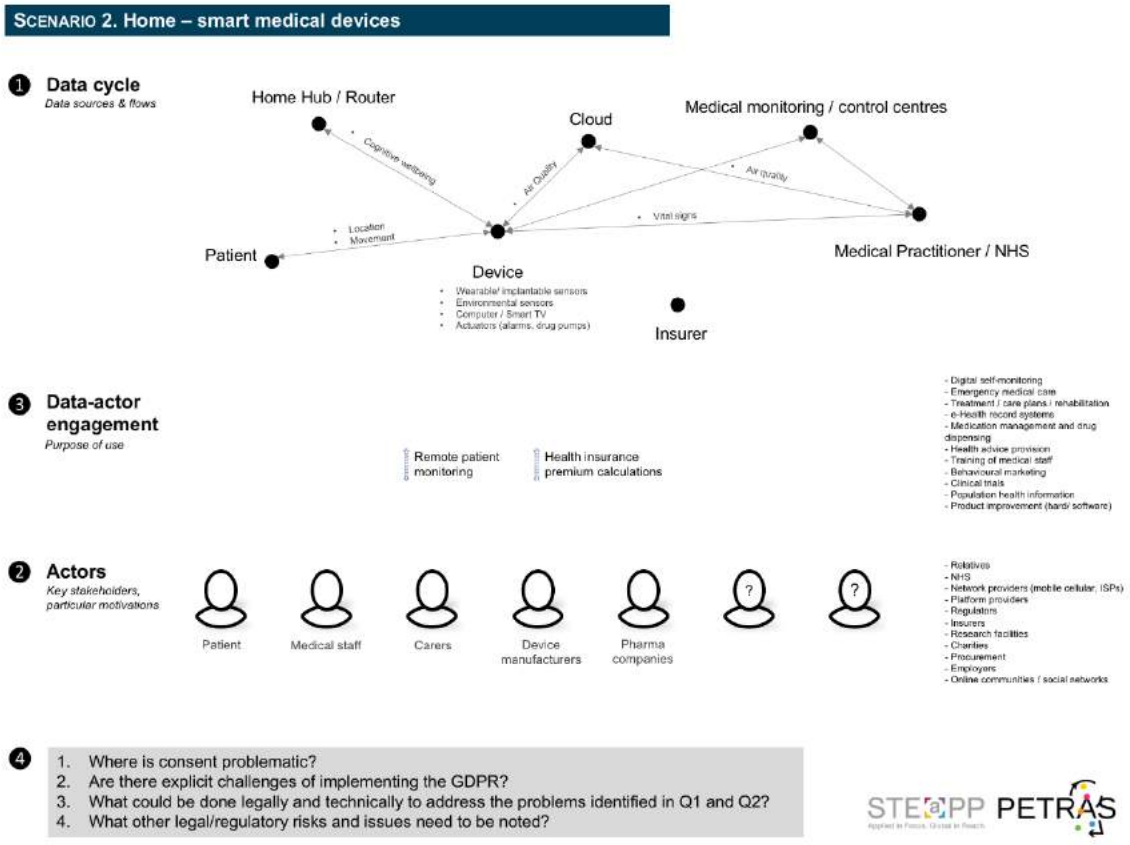
Alex provided the example of having worked on a privacy policy for an insurer who created a product for employees to monitor their **relatives**, primarily those affected by dementia. Alex had to work out the data flows between all actors involved. The assessment and consent considerations were complicated by the fact that a lot of the affected patients could not be considered compos mentis.

Workshop participants therefore recognised that the status of both carers and relatives is **dependent on their legal status** in this scenario. Given that relatives can be (and very often are) simultaneously carers, Cerys suggested to display both actors as "carers / relatives".

National Healthcare Service

Attendees also emphasised that there is **scope for differentiating the medical practitioner in the data cycle**. An updated Figure 3 should account for different entities that are engaged in this realm, with the NHS being a distinct data controller. NHS services have to be separated from medical professionals who have a private relationship with the patient.

Figure 3: Scenario Exercise



The "Invisible" Third-Party

A further actor that is not represented in the data cycle are "invisible" third parties. Madeline considers these actors to be **stakeholders that do not own a device**. In the IoT environment their data is being collected, but due to the potential inability of engaging with the IoT ecosystem, they may be deprived from giving informed and active consent.

Carsten gave the example of services that require users to own a phone or electronic contact details such as an email address. This can exclude a set of people that for various reasons (e.g., socio-economic, age) are not active consumers of such devices or services. It led mc to argue that individuals are basically becoming "invisible" when not participating in these technological developments.

Consent Issues Raised

After the debate on the scope of actors prevalent in the health IoT space, participants shifted their focus to particular consent concerns that are prevalent in the data cycle of medical devices and services. Here, they made reference to insurers, questions of public and vital interest, and the cost of privacy.

Insurance

Leonie highlighted issues around consent in the context of **insurance**. The data cycle between insurers, patients, and third-parties (e.g., medical practitioners) would remain unclear.

Cerys clarified the arrangement, emphasising that there generally had to be a direct agreement between the patient and the insurer for data exchanges to take place. The arrangement is also country-specific: In the UK, individuals can prevent the insurer from getting involved in the data exchange. It is therefore up to the **personal preference** of the data subject to determine the information flow. This might not be the same for patients in the US or other places.

Cerys referred to a **real-life example** in the UK. An insurer tried to directly access medical data from GPs. The insurer based the procedure on what they saw as the consent of the individual. The insurer wanted to rely on the same subject-access route that patients can use to receive information collected by medical professionals. The insurer consequently assumed to be able to authorise this route and thereafter receive access to health information directly.

The ICO however rejected the claim that those patients who gave consent to the sharing of this data entirely understood the implications of their decision. The ICO argued that they were unlikely to be aware of the full extent of data, breadth of material, and level of granularity of information that they would be exchanging. The ICO

consequently refused the request on the basis that the data subject's approval would not equal "informed" consent.

Cerys elaborated that the insurer was subsequently trying to identify an alternative data transfer model. One idea was that data had to first be approved by the individual, allowing them to screen the data, before forwarding it on to the insurer. However, the insurer realised that this was not practical on multiple levels (e.g., security concerns).

Based on this example, Cerys said that there are issues of consent connected to the question of insurance. Yet, it **would be an additional concern** that can be considered apart from the hidden actors that are part of the data transaction between the patient and the device.

Vital Interest

A further aspect that was discussed was the question of consent and vital interest.

Irina highlighted that the health example is distinct, insofar as the GDPR requires **explicit consent** when sensitive data such as medical information is being collected and processed. The level of criticality of this data is not the same as for data generated in other fields, where one might even argue that data collection is potentially in the **public's interest**.

Thus, processing is permitted if it is necessary for the performance of a task carried out in the public interest or carried out by an official authority. General Practitioners, for instance, can collect personal data on the basis of "public interest". Art 6 (GDPR) specifies means for achieving such "lawfulness of processing".

Art 6.1(e) states that lawful processing occurs when "processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller". However, there is also the concern that there might be "imbalances between the data subject and the controller, in particular where the controller is a public authority" (para 43, GDPR). In such a case, the data controller (public authority) may be in a disproportionately more powerful position than the data subject.

Cerys further noted that consent in the health scenario can be overruled by the **vital interest exemption**. Data controllers can override a data subject's consent in instances where the sharing of information is protecting the vital interest of the patient (e.g., the sharing of information with ambulance services, relatives etc). In such cases, the minimum level of data required for a necessary action to be taken is all that can be shared.

The Cost of Privacy

A final point that was raised in this exercise was the question of cost and benefits when consenting to the sharing of data.

mc suggested that data subjects should be incentivised to release parts of their data by, for example, receiving **tax benefits**. This, however, requires users to fully comprehend the implications of their consent and an understanding of the ramifications of the openness of their data.

Differential pricing was a further concept that was discussed. Cerys underlined that the earlier mentioned insurance example was based upon a differential pricing model. The insurer argued that the ability to have direct access to health data would bring benefits to people as premiums could be dynamically adjusted.

Despite the prospective benefits of this model, participants critically noted its hidden **discriminatory potential**. A person might not only pay a lower premium for being at less risk, but also for having less privacy. A data subject who does not feel comfortable sharing his/her data and withdraws his/her consent might consequently end up **paying for privacy**.

Hence, although researchers are increasingly trying to determining the value of information and the cost of privacy (Bölöni & Turgut, 2017; Turgut & Bölöni, 2017), more knowledge about the possible ethical failures that could arise from these market dynamics are needed and should guide policy making.

Scenario 3: Infrastructure – Built Environment

Data Cycle Discussion

The third and final exercise in this workshop focused on IoT in the built environment. The latter was conceptualised in broad terms, but with a particular focus on the external ecosystem (e.g., smart city) rather than the domestic realm (e.g., smart home).

Based on the content displayed in Figure 4, participants began with an evaluation of the comprehensiveness of the listed actors and data points. The conversation expanded thereafter and focused on two specific consent issues prevalent in the built environment, including “lack of personal data” and “legitimate interest”. Before moving onto that part of the discussion, we clarified the scope and actors involved in this case.

Data Points and Their Meaning

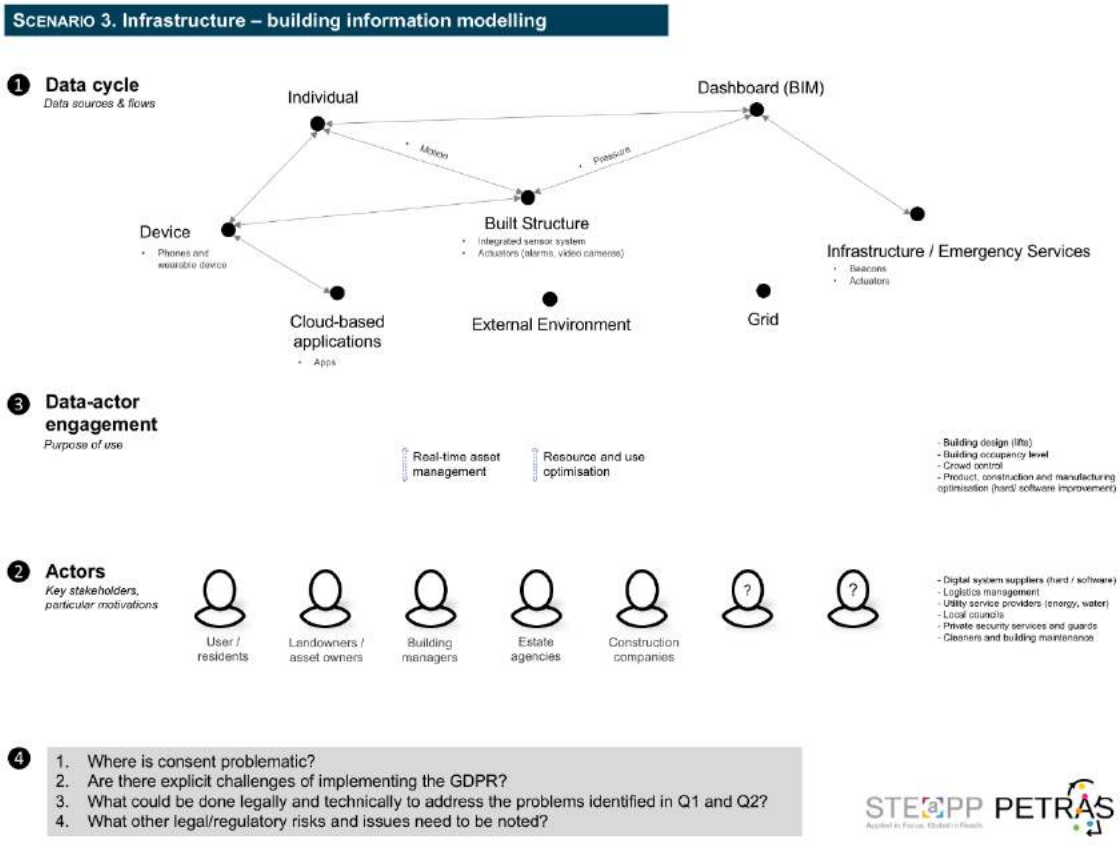
The data point **"built structure"** was understood to include the smart home but also critical infrastructures such as bridges or public and private spaces such as parks and shopping centres.

"Individual" was understood as any **active or passive** human being, who is potentially vulnerable, and engages with the build environment.

The data point "device" was conceptualised as a **node or 'thing' that collects, stores, and transmits data for one or more services**. A device might be integrated in the built infrastructure, but equally, it may also be completely separate from it.

"External environment" was included to highlight that IoT devices and services might interact with **extraneous surroundings**, for example, to assess the air quality, air temperature or weather conditions.

Figure 4: Scenario Exercise



The "**dashboard"** represents a centralised hub which could be the control system of a house or that of a building or part of the critical infrastructure. The dashboard acts as unified coordination centre. It controls aspects such as the air conditioning and energy or water usage of a building.

The graphic also accounted for data flows to **larger sets of "infrastructure providers**", including emergency services such as law enforcement or the fire brigade.

Participants highlighted that based on the **actors and their specific interests**, the data cycle might change and need to be amended.

Consent Issues Raised

Due to time constraints, the discussion of Figure 4 had to be shortened. However, two insights that emerge from the exercise point to differences in the data flow and consent structure that diverted from the scenarios on CAVs and IoT medical devices. For one, participants acknowledged that less 'personal' data is being collected; for another, questions on legitimate interest might be more prevalent in the build environment.

Lack of Personal Data

Participants noted that the public space is a unique scenario to look at. Data subjects are part of a transitional engagement. People are coming and going and, in many instances,, are not even required to **actively and personally** engage with the surrounding ecosystem.

Nonetheless, data subjects still need to receive notification about the types of data that are being collected, and here Cerys emphasized that there needs to be **transparency** about this process taking place (i.e., akin to the current CCTV usage, where the UK Surveillance Camera Commissioner (2015) specifies that operators have to provide a notice informing people that recording is taking place).

Cerys also noted that not each of the data flows in Figure 4 involve the collection of personal data. For example, the collection of environmental data that is subsequently used and transported between IoT devices does not count as personal data being exchanged.

Similarly, information collected in the public space may be gathered on an anonymous, **non-personal basis**. For instance, IoT devices may transmit non-descriptive knowledge such as 'an individual has passed by', rather than forwarding specific details on who exactly this individual was.

Going back to the earlier discussion, this does not mean that the collected data cannot become identifiable personal data when it is related to other information. However, it does narrow the level of immediate identifiability, which may be relevant for consent considerations.

Legitimate Interest

A second factor that was discussed centred on the collection and processing of data based on legitimate interest. The GDPR (47) defines a legal basis for a **legitimate purpose conditions**, but expects data controllers to also consider the interests or the fundamental rights and freedoms of their data subjects.

Claims about the existence of a legitimate interest will require **careful assessment.** A data controller will have to specify the conditions for such a legitimate purpose condition **in advance** (to the data subject?) and provide a **compelling reason for its necessity**. A **paper trail** of rationalisation as well as a balance of the interests of the data controller and the data subject has to be demonstrated.

Besides, data subjects have to be made aware of what the data controller is doing and on what basis. If this process lacks transparency, a person may challenge its legitimacy, with Cerys noting that “with consent, a person can withdraw it; with legitimate interest, a person can challenge it”.

Conclusion and Further Action Points

Working through these three scenarios allowed us to explore a number of issues around consent that will or may arise as a result of the combined changes that the implementation of the GDPR and further development of the IoT will bring about. The starting point for the workshop was that consent issues may be less problematic than anticipated in some ways and more problematic in others but either way, there was a need for much more clarity around this. The workshop helped us to identify some of the challenges and articulate some useful research questions. They included:

i. **Considering that the GDPR introduces new concepts for understanding consent including "clear affirmative act" and "unambiguous indication":** What differentiates the current definition of consent as stated in the Data Protection Act 1998 to the new GDPR definition and what are the implications for the implementation of consent in the IoT?
ii. **Considering the momentum towards developing a universal tool and approach for managing consent in the IoT ecosystem**: What possible challenges might this introduce for the a) integrity of consent; b) data subject flexibility to tailor their consent preferences, withdraw? c) 'invisible' data subject who does not own/carry a device with which to grant consent?
iii. **Considering that the GDPR does not preclude devolving responsibility for consent to an 'agent' which may be interpreted as an 'algorithm'**: What ethical problems might such tools for managing consent generate or encounter? Can the intention of the GDPR be fully realized in such a model?
iv. **Considering that the proposed standardised privacy policy icons attracted some criticism**: To what extent can the use of standardised icons meaningfully depict data flows, storage, etc.?
v. **Considering the concerns in Germany around the use of video cameras in CAVs**: How will the GDPR obligations around consent be applied to those data subjects peripheral to the car? How does the vehicle-to-pedestrian link need to be reconsidered and what might this tell us about peripheral consent in other IoT application areas?
vi. **Given the complexity of nested services and data transfers in the IoT**: How can software/application developers be supported to ensure they comply with the GDPR all through the development process?

Based on these questions and overarching findings, we also propose below some further action points that some of the workshop participants (or others who engage with this report) my wish to pursue:

i. Update, expand and compare the data flows and generate a generic **framework model** that can be applied to diverse IoT scenarios;

ii. Engage with PETRAS research projects as well as other **IoT UK testbeds** to identify what data flow issues they have encountered that have consent implications.
iii. Examine existing **libraries, platforms and tools** that software engineers engage with and assess them for their compliance with the GDPR and its associated principles;
iv. Assess both **regulatory and technical solutions** to evaluate how they will promote or ensure informed consent in the IoT ecosystem;
v. Conduct more research on **data subjects'** understanding of IoT's data flows and their right to consent.
vi. Systematically trace the data flows involved in a single use case (connected autonomous cars) to establish points at which consent will be required.

Acknowledgements

The workshop's structure and methodology was designed by Dr Ine Steenmans, Research Associate in 'Foresight and Futures' at STEaPP. Her support and guidance was invaluable for the delivery of this workshop and the insights that were generated.

References

Baarslag, T., Alan, A. T., Gomer, R., Alam, M., Perera, C., Gerding, E. H., & schraefel, m. c. (2017). An Automated Negotiation Agent for Permission Management. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems (AAMAS 2017)* (pp. 380–390). São Paulo, Brazil: International Foundation for Autonomous Agents and Multiagent Systems. Retrieved from http://dl.acm.org/citation.cfm?id=3091125.3091184

Bölöni, L., & Turgut, D. (2017). Value of information based scheduling of cloud computing resources. *Future Generation Computer Systems*, *71*, 212–220. https://doi.org/10.1016/j.future.2016.10.024

Data Science Innovation. (2015). Data Science and Privacy Regulations - A Storm on the Horizon. Retrieved July 29, 2017, from http://dsianalytics.com/data-science-and-privacy-regulations-a-storm-on-the-horizon-full/

Egelman, S. (2017, July 27). We tested apps for children. Half failed to protect their data. *Washington Post*. Retrieved from https://www.washingtonpost.com/news/the-switch/wp/2017/07/27/we-tested-apps-for-children-half-failed-to-protect-their-data/

ENISA. (2013). *On the security, privacy and usability of online seals. An overview* (pp. 1–32). Heraklion, Greece: European Union Agency for Network and Information Security. Retrieved from https://www.enisa.europa.eu/publications/on-the-security-privacy-and-usability-of-online-seals

European Commission. (2017a). *Access to In-vehicle Data and Resources. Final Report* (pp. 1–259). Brussels: European Commission. Retrieved from

https://ec.europa.eu/transport/sites/transport/files/2017-05-access-to-in-vehicle-data-and-resources.pdf

European Commission. (2017b, October 9). Building a European data economy. Retrieved October 9, 2017, from https://ec.europa.eu/digital-single-market/en/policies/building-european-data-economy

Fischer-Hübner, S., Angulo, J., & Pulls, T. (2013). How can Cloud Users be Supported in Deciding on, Tracking and Controlling How their Data are Used? In *Privacy and Identity Management for Emerging Services and Technologies* (pp. 77–92). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-55137-6_6

Information Commissioner's Information Commissioner's Office. (2017, June 13). Key definitions of the Data Protection Act. Retrieved August 8, 2017, from https://ico.org.uk/for-organisations/guide-to-data-protection/key-definitions/

Kang, R., Dabbish, L., Fruchter, N., & Kiesler, S. (2015). "My data just goes everywhere": user mental models of the internet and implications for privacy and security. In *Symposium on Usable Privacy and Security (SOUPS)* (pp. 39–52). USENIX Association Berkeley, CA. Retrieved from https://www.usenix.org/system/files/conference/soups2015/soups15-paper-kang.pdf

Neisse, R., Baldini, G., Steri, G., Miyake, Y., Kiyomoto, S., & Biswas, A. R. (2015). An agent-based framework for Informed Consent in the internet of things. In *2015 IEEE 2nd World Forum on Internet of Things (WF-IoT)* (pp. 789–794). Milan, Italy: IEEE. https://doi.org/10.1109/WF-IoT.2015.7389154

Pearson, S., & Casassa-Mont, M. (2011). Sticky Policies: An Approach for Managing Privacy across Multiple Parties. *Computer*, *44*(9), 60–68. https://doi.org/10.1109/MC.2011.225

schraefel, m. c., & Gerding, E. (2013, 2017). Meaningful Consent in the Digital Economy. Retrieved July 29, 2017, from http://www.meaningfulconsent.org/

Surveillance Camera Commissioner. (2015, December 8). Domestic CCTV: using CCTV systems on your property - GOV.UK. Retrieved August 8, 2017, from https://www.gov.uk/government/publications/domestic-cctv-using-cctv-systems-on-your-property/domestic-cctv-using-cctv-systems-on-your-property

Turgut, D., & Bölöni, L. (2017). Value of Information and Cost of Privacy in the Internet of Things. *IEEE Communications Magazine*.

Weinberg, B. D., Milne, G. R., Andonova, Y. G., & Hajjat, F. M. (2015). Internet of Things: Convenience vs. privacy and secrecy. *Business Horizons*, *58*(6), 615–624. https://doi.org/10.1016/j.bushor.2015.06.005

Appendix A: Scenario 'Worksheets'

Context	Data Subject	Data points (where data is being collected)	Data flows (transfer and processing of data)	Agents (all actors involved)	Data purposes (what is the data being used for)
Mobility – connected and autonomous vehicles	Driver / passenger	- Personal mobile device (smart phone) - Sensor systems for vehicle location, proximity, motion, collision (GPS, tyre pressure monitoring system, engine control, light detection, camera and infrared systems etc.) - Actuators (brake control) - On-board diagnostic system (dashboard) - Entertainment system - Bluetooth - USB ports - Cryptographic keys - Road cameras, road beacons	- V2I (vehicle-to-Infrastructure e.g., road, sensors etc.) - V2V (vehicle-to-vehicle) - V2P (vehicle-to-Pedestrian) - V2D (vehicle-to-device) - V2G (vehicle-to-grid) - IVI (in-vehicle infotainment	- Passengers (potentially vulnerable) - Other drivers / passengers - Vehicle manufacturers - Additional hard/ software suppliers - Network carriers (mobile telecom providers) - Certificate authority responsible for public key distribution - Infrastructure managers (Highways England/ Transport NI) - Vehicle performance monitors (e.g., MOT testing centres) - Regulators - Authorities (police forces, local councils) - Infotainment providers - Parking service providers - Traffic service providers (TomTom) - Breakdown assistants - Fleet mangers - Car dealers - Leasing companies - Insurance companies - Employers	- Driving state information (speed, real-time location, journey data, crash notification, tire pressure) - Driver behaviour (route preferences, speed, reaction time → profiling) - Other driver preferences (music, car interior) - Vehicle diagnostics (damages, malfunctions) - Smart infrastructure management - Road improvement/ expansion - Traffic management (congestion, traffic flow, parking management) - Product (hard/ software) improvement - Car insurance lawsuit settlement - Targeting advertising

Context	Data Subject	Data points (where data is being collected)	Data flows (transfer and processing of data)	Agents (all actors involved)	Data purposes (what is the data being used for)
Home – smart medical devices	Patient	- Personal mobile device - Wearable/ implantable sensors to collect systemic data (blood pressure, heart rate, glucose levels) - Environmental sensors (smoking exposure, air quality, motion etc.) - Computer/ Smart TV/ home hub (games for cognitive wellbeing) - Home routers - Actuators (alarms, video cameras, drug pumps, defibrillators)	- Patient to device - Device to home hub/ router - Device to medical monitoring/ control centres - Device/ home hub to cloud service provider (CSP) - Device/ home hub to medical practitioner/ NHS - CSP to medical practitioner/ NHS/ researchers - Patient to insurer (dynamic insurance)	- Carers - Relatives - NHS - Medical staff (physicians, hospitals, private medical providers) - Device manufacturers - Network providers (WLAN, mobile cellular, ISPs) - Platform providers - Regulators - Insurers - Pharmaceutical companies (personalised medicine) - Research facilities - Charities - Procurement - Employers - Online communities / social networks (asthma online communities that share experiences and data)	- Digital self-monitoring - Remote patient monitoring (vital signs, fall prevention, cognitive decline) - Emergency medical care - Treatment/ care plans/ rehabilitation (quality of life enhancement) - e-Health record systems - Medication management and drug dispensing - Health advice provision - Training of medical staff - Behavioural marketing - Clinical trials - Population health information - Product (hard/ software) improvement - Health Insurance premium calculations

Context	Data Subject	Data points (where data is being collected)	Data flows (transfer and processing of data)	Agents (all actors involved)	Data purposes (what is the data being used for)
Infrastructure – Building information modelling	?*	- Various types of integrated sensors (ultrasonic, proximity, motion, temperature, pressure, energy, lighting, smoke and gas sensors) - Actuators (alarms, video cameras) - Phones and wearable devices - Building information systems - Cloud-based applications (apps)	- Individual to sensor network - External environment to sensor network - Sensor network to dashboard (i.e., BIM) - BIM to physical infrastructure (actuators) - BIM to infrastructure managers and emergency services (British Gas, fire station, police etc.)	- User / residents - Landowners and asset owners - Building managers - Estate agencies - Construction companies - Digital system suppliers (hard/ software) - Logistics management - Utility service providers (energy, water, telecoms) - Local councils - Private security services and guards - Cleaners and building maintenance	- Building design (lifts etc.) - Real-time asset management (humidity levels, air quality, ventilation, cooling/heating systems, energy conservation, security and safety) - Building occupancy level - Crowd control - Resource and use optimisation (energy consumption) - Product, construction and manufacturing optimisation (hard/ software improvement)

* To be discussed.