Informed by Design

Joseph Lindley, Paul Coulton, Rachel Cooper

Imagination, Lancaster University, United Kingdom

Keywords: Privacy by Design, Data Protection, GDPR, Informed by Design.

1 Introduction

The term the Internet of Things (IoT)is somewhat ambiguous in that its definition varies dependant on who is using it and in what context. Arguably it has gained traction with a wider audience, over other terms such as ubiquitous computing or pervasive computing, in that it is made up of the words 'internet' and 'things' which are more generally accessible terms. However, this also means that it is open interpretation depending upon the meaning an individual associates to these terms resulting in the variety of discourses which invariably help shape how it is perceived, developed and adopted. While this paper is couched in IoT discussions, it is arguably applicable to discourses about technology more generally.

The current (and future) adoption of the IoT has, for some time, stimulated debate about the broader implications for privacy [24], ethics, trust and security that the IoT. Given the IoT's penchant for generating and utilising various (oftentimes somewhat personal) data, the European Union's (EU) forthcoming General Data Protection Regulations (GDPR) will have a significant impact on how the IoT is regulated. As with the term IoT the interpretation of GDPR is generating its own discourses particularly around how wording within the regulation is turned into implementation

The paper begins by critiquing the term *Privacy by Design* (PbD), and an alternate form which appears in article 25 of the GDPR *Data protection by design and default.* We note that these two phrases are in fact part of a broader group which inexhaustively includes: Security by Design, Privacy by Default, Security by Default, Data Protection by Design, Data Protection by Default. Our critique does not concern the sentiments or intentions represented by these phrases, or PbD per se, but highlights ambiguities and potentially misleading interpretations that their invocation promotes. After exploring these potential pitfalls, we go on to discuss design-led research that positions *Informed by Design* as a more fruitful approach to creating IoT devices and services which can more meaningfully respond to concerns about privacy, ethics, trust and security.

Informed by Design draws upon design-led futures research into the adoption and acceptability of the IoT in domestic settings (although the ideas may well be applicable to other contexts too). Using Design Fiction [3,10,20] to explore contemporary philosophies that explore New Materiality [9,36] and Object Oriented Ontology [5,17,19,22], our proposal for an *Informed by Design* approach intends to deliver actionable implementation strategies. In doing so we aim to provide a framework to help designers produce IoT services and devices that meaningfully comply with GDPR, not only to the letter, but with its spirit too.

2 […] by […]

Our initial discussion is relatively broadly scoped, and explores issues across a gamut of related terminology. For example, configuring something *by default* is not the same as creating something in a particular form, *by design.* However, arguably to have something configured a particular way by default means that it must have been designed that way, and if something is set up to, by default, in a particular way then it must have been designed. This issue is exacerbated because 'design' can be used in so many different ways in the English language, e.g. *the designer used her knowledge of design to design the widget, which was then the design.* The original source of this conflation between 'by design' and 'by default' appears to be an influential report which, on several occasions, employs forms of words such as "incorporates *Privacy by Design* principles by default" [8].

These already murky waters may appear completely opaque even before we introduce the further depths of difficult abstractions like privacy and security. To unpack these very quickly: privacy is not the same as security, but in some circumstances privacy may be delivered by security and security may be delivered by privacy. When attempting to make sense of this disciplinary idiosyncrasies may well come into play too. For example, it's plausible that an engineer may interpret *security* in terms of a particular implementation, like access control lists, whereas a psychologist may think of it in terms of a psychological theory, such as Maslow's hierarchy of needs. Both are valid routes forward, however when their epistemological roads intersect, common understanding will not necessary be the *immediate* destination. Whilst these definitional anxieties are not unique to IoT and PbD, they must be acknowledged for the sake of intentional specificity. Although the argument in this paper intersects with much wider debates, for the most part the *specific* issues we are interested in are (1) *Privacy by Design* as it was described by by the Ontario Information and Privacy Commissioner [8] and (2) *Data protection by design and by default* as it appears in article 25 of the GDPR [39]. In the following we unpack aspects of both of these areas in more detail.

2.1 Privacy by Design

The term can be traced back to a 1995 report[1] on Privacy-enhancing technologies which was bootstrapped for the modern era in a 2012 report *Privacy by Design in the Age of Big Data* [8]. In the foreword to the report Ann Cavoukian quotes a 13th century Persia poet's words to articulate that to reinvent the world, one must speak a new language. The subtext is that technological progress is itself a new language, but one that brings with it new challenges with respect to privacy. Within the foreword Cavoukian then gives an example of what PbD is: a way of processing data that has been cryptographically hashed. This means that although patterns can be found in the data, apparently there is no way to reverse engineer it and find out what the original data was. While the example seems compelling, it is arguably also a little naïve. Although such technical approaches can, in particular circumstances, protect the privacy of individuals who are represented in the data, to make the assumption that a system built with this technique can be operated free from privacy concerns would be remiss. In this particular example, assuming an actor had access to the source data, and also knew the details of the hash algorithm, then the privacy protection would be circumvented with ease.

The 2012 report is largely based on technical work done by Jeff Jonas and focuses on seven principles that for creating systems that are private by design. The principles include:

- Full attribution for each data record;
- Data tethering (that any changes to data are recorded immediately);
- Analytics on anonymised data;
- Tamper-resistant audit logs;
- Engineer systems to tend toward false negative rather than false positive in borderline cases;
- Self-correcting conclusions (check prior conclusions based on new data);
- Information transfer accounting (wherever the data goes, it should be trackable and traceable—whether that is to a hard copy, monitor, or another system)

These principles are specifically concerned with 'sensemaking systems' (i.e systems that synthesise data from multiple systems such as payroll, customer relationship management, financial accounting) and make sense out of it. While the principles do make some sense within *that specific* context, they are regrettably *too* specific to make much sense in the heterogeneous user groups and devices of the IoT (even within domestic IoT there is a huge variety) and hence these guidelines aren't a particularly useful start point if one wishes to understand the acceptability (and associated adoptability) of domestic IoT devices.

When exploring the challenges of PbD, Sarah Spiekermann notes "Data is like water: it flows and ripples in ways that are difficult to predict" [29]. Her notes going forward suggest PbD is rather idealistic and, in practice, can be resolved to a utilisation of Privacy-Enhancing Technologies combined with additional security, aiming to create a "fault-proof" landscape. As she puts it, "the reality is much more challenging". Spiekermann highlights the idealism by referring to the advertising revenues garnered by, for example, Google and Facebook, pointing out that "without personal data, such services are unthinkable" and that proponents of PbD "hardly embrace these economic facts in their reasoning". Put differently, it may not be *possible* to create systems which are profitable and feature-rich without contravening some of the current ideals of PbD.

That is not to say that PbD isn't successful, and indeed a powerful way to approach system design. Responding to Spiekermann, Cavoukian notes her broad agreement with the analysis, but also insists "the challenges of PbD are not as great as Spiekermann suggested; the engineers I have met have embraced the PbD principles, finding implementation not difficult" [7], citing experience of discussions with large tech companies and project experience on the Ontario Smart Grid. It seems to be the case that part of Spiekermann's critique is to point out that to do PbD effectively, it must become part of organisational culture, cutting through management, marketing, design and engineering. Cavoukian agrees saying this *is* possible. That specific projects can be carried out with some semblance of PbD spanning from the top down, and the bottom up, is true. *However*, the more interesting part of Spiekermann's commentary touches on potentially systemic shortcomings are the core of PbD's rhetoric: a 'fault-proof' landscape is unrealistic and the 'economic facts' of many business models seem to be conveniently ignored.

These two factors, are perhaps what causes PbD to stagnate, and to struggle to move from principles to practicalities. Another way to look at it echoes Shapiro's suggestion that neither engineers nor customers are able to properly articulate, understand or analyse the impact of 'non-functional' requirements (those which do not impact on a system's behaviour, but do effect its acceptability) like privacy [28]. These hard-to-grasp requirements operate at a completely different level of abstraction to what engineers and customers are comfortable dealing with.

So, to recap. The new language of technology is making a new world, but, we're quite fluent in the language or comfortable in the world yet. While purely technical responses to privacy sometimes *appear* to have faultless solutions (e.g. processing irreversibly hashed data), however, rarely will that solution be generally applicable, and often it may be defeated by similar means to which it was implemented. Fundamental principles appear to be useful mechanisms; however, their fundamentality often seems compromised when seen through the lenses of specific contexts, or types of context. Although PbD has demonstrably helped to inform the delivery of privacy-aware projects with buy-in from developers, customers, and management alike, examples of this appear to be isolated and do not necessarily

[1] http://www.ontla.on.ca/library/repository/mon/10000/184530.pdf

cut through the aforementioned difficulties. Rhetoric around PbD hints at the feasibility of creating a 'fault-proof' approach to privacy and often fails to appreciate the economic realities of what makes data-centric businesses viable.

2.2 GDPR Art. 25: Data protection by design and default

When the GDPR becomes active across the European Union on 25th May 2018 the data protection legislation across a swathe of Europe will immediately change. Given that the regulations protect citizens regardless of where data pertaining to them is held, it will also impact upon a huge number of organisations who hold data *about* European citizens, who will have to comply or risk sanctions. There is little certainty about how GDPR will play out in practice though, as it is a complex legal assemblage, and in parts has scope for sovereign nations to interpret it as they see fit while other aspects will be enforced by EU bodies. Notwithstanding differences in local interpretation, guidelines about how to interpret the regulations will take time to emerge. Among the gamut of these challenges, various facets of minutiae may cause further complications; for example, particular phrases may have quite contrasting meanings in the different languages that GDPR is written in [15].

Within the GDPR, article 25 deals specifically with *Data protection by default and design*, and hence whenever those concepts are drawn upon by other articles or recitals, they, in turn, cite article 25 [37]. The opening words of the article say that data controllers must take "the state of the art" into account however no scrutiny is given to what state of the art might really mean. Given that this assertion is made in under the heading 'data protection by design and default' we might infer that there's a relationship between the two, but the *terms* of that relationship are left unwritten. Article 25 also makes reference to the 'by default' trope, saying that appropriate measures be taken to ensure that, *by default*, "only personal data which are necessary for each specific purpose of the processing are processed". So, on face value it seems that GDPR's interpretation of data-protection by design, and relatedly default, is at best ambiguous. It certainly doesn't progress the PbD debate about how to move beyond principles and towards specific guidelines. This lack of specificity that regulators have with respect to PbD (and its relatives) is not confined to the text of the GDPR. The UK Information Commissioners Office (ICO) the body in the UK responsible for interpreting and enforcing GDPR provides guidance on PbD, however within their guidance they do not proffer any clear definition of what it is or how to implement it[2], only referring to projects run by Cavoukian's department in Ontario. While it may seem unreasonable to expect European regulators define a working model and clear definition of PbD, given that scholars across various disciplines have struggled to do so, including the terminology in the GDPR *without* attending to PbD's lack of clear definition, could become problematic.

2.3 Unsinkable by Design

Hubris can be dangerous. The infamous ocean liner *Titanic* developed reputation for being unsinkable prior to its maiden voyage across the Atlantic Ocean. While the ship *was* technologically advanced and had state of the art safety measures at the core of its design, given its untimely demise, it seems laughable to suggest that it could have been unsinkable. Reportedly, the ship received ample warning to avoid the ice berg that eventually tore open one side of her hull, however given the level of confidence about her unsinkability, the message was never delivered to the Captain. Furthermore, there were not enough lifeboats for those aboard and because of lack of preparation for such an emergency panic and disorganisation took hold as the ship went down. The tale of the titanic shows us that simply embracing the 'state of the art' can actually be counter-productive, and, under some circumstances can encourage systematically unhelpful behaviour (not having enough lifeboats, or telling the captain about an iceberg on collision course, for example).

Allegorically similar, bicycle locks come in a myriad of shapes, sizes and at various price points; chain-locks, D-locks, combination locks, cable locks, smart locks and so on. Some locks are more secure than others, however, given the right tools, expertise, and amount of time a thief can break *any* lock. Hence, considering the most prestigious lock totally 'safe' or the tackiest cheap lock 'not safe' represent a not-particularly-helpful hubristic position. One alternative is to look at locks in terms of an average or estimated protection *time* (i.e. to give you an idea of how long it might take a dedicated thief to break it). Arguably we need similarly interpretive schemes to understand where a given system intersects with its privacy requirements.

To consider the traps such perceptions represent more generally we utilise Heffalumps which are type of fictional elephant, which appear in A.A. Milne's Whinnie the Pooh stories. Even within the realms of the stories Heffalumps are imaginary, only ever appearing in Pooh's mind. Nonetheless, one day Pooh and his sidekick Piglet attempt to catch one by building a trap. Unfortunately, after setting the trap, they get lost and—somehow—stumble into a trap *themselves* (it is inferred it is the same trap they laid). The term *Heffalump Trap*[3] has become a term in political journalism used to describe occasions when a trap is set up to catch an opponent, but ends up catching the person set the trap (see figure 1).

Each of these tales can be considered an allegory for PbD. Even though it was not the designers or engineers of *Titanic* which began and proliferated stories about her unsinkability, the myth took hold and even the *crew* believed it. Similarly, to believe that one's top of the range whizz-bang bicycle lock sold by a prestigious brand, is completely unbreakable would be remiss, and could easily lead to an unanticipated theft. Promoting binary positions—whether they regard sinking ships, bicycles, *or privacy*—promotes irrational beliefs. From

[2] https://ico.org.uk/for-organisations/guide-to-data-protection/privacy-by-design/

[3] https://en.wikipedia.org/wiki/Heffalump#Cultural_impact

these irrational beliefs a toxicity grows, and while the thin end of the wedge may result in inconvenient bicycle thefts, at the other end of the spectrum it can result in a systemic miscalculation of risk, which in the case of *Titanic* meant a needless loss of life. An obvious mitigation strategy is to provide interpretable information about *relative* merits of a particular solution, whatever the domain is.

Relating this back to our core discussion, we suggest that if not handled carefully, PbD, could act as a Heffalump trap. If the level to which a system protects privacy is assumed to be high (or 'high enough') simply because it followed PbD principles, then the person holding this belief—regardless of whether they are user, consumer, or designer—puts themselves at risk of being caught out on the occasion that the system isn't as private as they thought. This can have an irrecoverable effect on future adoption and acceptability.

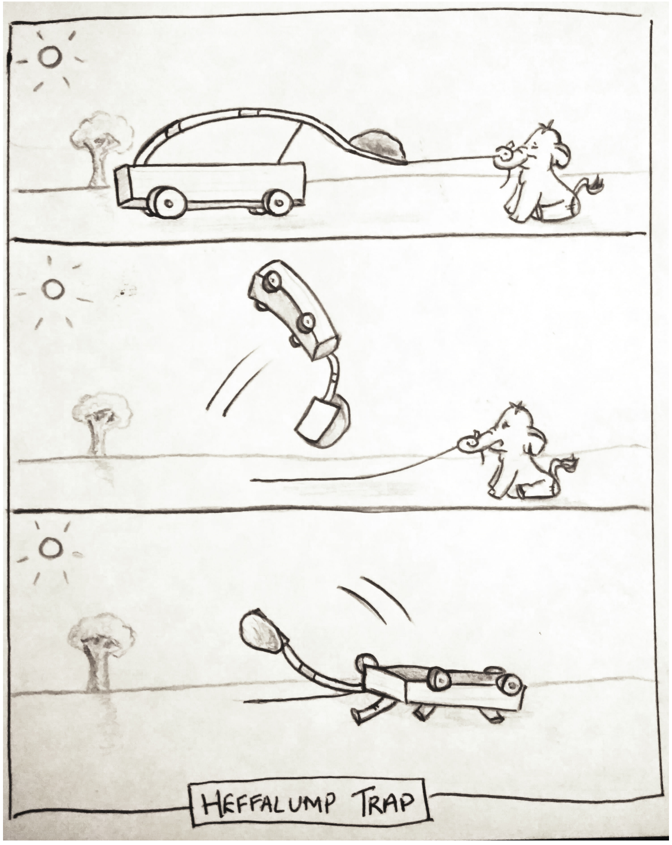


Figure 1. Depiction of a Heffalump Trap.

In terms of the how GDPR, there seems to be a danger that, given 'data protection by design and default' is barely specified, that data controllers will interpret the approaches mentioned in article 25 (e.g. data minimisation, pseudonymisation, et cetera) and treat systems which utilise those as compliant without critically examining tangible impacts and risks associated with a system's data usage. By extension the other articles and recitals which depend upon data protection being delivered by design and default will inherit the same shortcomings. Plausibly the engineers, lawyers, a customers of particular data controllers may, given that their products 'do data protection by design and default' *assume* that doing this means that their systems operate in accordance with GDPR's aims. In reality, they would simply be adhering to the minimum legal suggestion of what GDPR lists under the header 'data protection by design and default', and which doesn't bear much resemblance to more involved attempts to understand how to actually enact PbD.

Practitioners using PbD are at risk in a similar way, whether or not they are governed by the GDPR. Despite the temptation to believe that following privacy can be 'achieved' by following the various principles of PbD, it simply is not the case. Notions of such 'fault-proof' systems are akin to unsinkable ships or impenetrable locks; they are *illusions* and do not exist. Even accepting that such systems don't exist, we must also contend with the fact that to move to a digital economy that can function without the gathering and collecting of personal data is akin to a physical economy that can work without oil; some major disruption needs to take place before this can be a reality. The lack of coherence in PbD methods, relatively immature models and frames for understanding what privacy is, and hence difficulty in actually *implementing* PbD confound all these issues too.

Ship captains, bicycle owners, and Heffalump trappers, and system developers; *beware.* Don't claim the ship is unsinkable, don't claim lock is unbreakable, don't try to catch the imaginary animal—most importantly, don't assume it's possible to build a system that's flawlessly *'private by design'*. Instead, strive for a system which is transparent and informative, *as well as* embracing concepts like data minimisation and pseudonymisation.

3 Informed by Design

Although we wish to highlight the danger of attempting to catch Heffalumps and build unsinkable ships—i.e. misplaced beliefs that PbD can, with some assurance, deliver privacy-compliant systems—we also wish to commend the endeavour. In the the following we introduce a subtly different framing of privacy as *Informed by Design* (IbD). By describing a notion at a higher level of abstraction this form of words is immediately advantageous and avoids the somewhat messy and easily conflated privacy by design, security by design and/or data protection by design (and/or default). Relatedly, by suggesting that the thing being enabled or delivered *by design* is *information,* as opposed to fiendishly-hard-to-define concepts like privacy or security, the scope of expectations with IbD is arguably easier to comprehend *and* deliver than it is with PbD.

Our approach to IbD has developed as a Research through Design [13,14,21] derived outcome, of a series of *Design Fiction* projects [3,10,20]. Widely attributed to the writer Bruce Sterling [31,32], but significantly further developed by designer and researcher Julian Bleecker and his collaborators at the Near Future Laboratory [3,4], Design Fiction is a method of speculative design which focuses on building fictional worlds. Design fictions contain users, environments, and technologies. The technologies they contain may often be 'real' insofar as they exist in some early or prototypical form in the present, however they are 'fictional' in the way they only exist *diegetically* in an artificially fabricated future world [10]. As well as using Design Fiction as a methodology, as-is, we specifically used it to experiment with Object Oriented Ontology (OOO) as a theoretical means to respond to the new types of network agency that the IoT has enabled. In the following we discuss OOO and Design Fiction in a little more detail, before describing the projects from which we derived these insights.

3.1 Object Oriented Ontology

As we are not philosophers we willingly defer the task of arguing OOO's validity and/or critiquing its merits to those more qualified than ourselves. However, what follows aims to articulate an accessible summary of our interpretation of OOO and to contextualise the subsequent account of the interplay between our design practice and engagement with theory.

In his seminal work *Being and Time*, Heidegger presents his view of ontology. By laying the foundations for OOO this highly influential 20th century philosophical text has taken on a new life in the 21st century [18]. The traditional Heideggerian view argues that things—objects—are all but impossible to understand in their *own* phenomenological terms, and therefore, we should make sense of them in relation to human use. Heidegger coined neologisms to communicate his argument, and famously uses a hammer as an example. When a hammer (or other object) is in its normal context of use it is 'ready-to-hand' and if that context is disturbed (for example if the head of the hammer fell off) then it is described as 'present-at-hand'. The metaphysics of this distinction are complex and must be negotiated outside of this paper, but the important point to note is that the hammer only comes into being via a human use (or perhaps non-use, in the case of the broken hammer). Central to the Heideggerian position is the notion that existence is a "correlate between [the human] mind and world" [5]. That these two constructs are inseparably linked is what Meillassoux refers to as 'correlationism' [16]. OOO *rejects* this correlationism and instead entertains the idea that objects have their *own* realities which are distinct from human use and from the mind/world correlate. From this post-correlationist position, anything—literally *any thing,* from a fibre optic cable, to a blade of grass, to a quantum computer, to a gooseberry fool—may be cast in the limelight of its own ontological resolve. If we consider the amalgamated glow emanating from the bazillions of tiny *lights-of-non-correlationism* then the resulting plane of luminescence is what illuminates OOO's "flat ontology" [6]. Having departed from familiar and intuitive human-centric ontologies, the vantage point one assumes when considering the nature of OOO's flat ontology is a strange and conflicted place:

"In short, all things equally exist, yet they do not exist equally […] This maxim may seem like a tautology—or just a gag. It's certainly not the sort of qualified, reasoned, hand-wrung ontological position that's customary in philosophy. But such an extreme take is required for the curious garden of things to flow. Consider it a thought experiment, as all speculation must be: what if we shed all criteria whatsoever and simply hold that everything exits, even things that don't? […] none's existence fundamentally different from another, none more primary nor more original." [5:11]

This open-endedness is necessary because in OOO 'objects' are not just *material* objects, but extend to anything. Such a categorisation requires special appreciation and a theory which allows for multiple types of 'Being' to meaningfully coexist. Exemplifying this Bogost uses the famously ill-fated video game *E.T. the Extra-Terrestrial* as an example. He muses that the game is simultaneously many different things:

- 8 kilobytes of opcodes;
- a compilation of source into assembly code;
- a flow of radio frequency into a television;
- a plastic cartridge;
- memory etched on wafer;
- a consumer good;
- a set of rules and game mechanics;
- intellectual property;
- 'the worst game ever made';
- a constituent of 728,000 Atari games *buried* in New Mexico[4];
- all of the above.

There is no elementary thing which comprises the video game, it is never a *single* one of the objects above, nor is it their conglomerate. Bogost tells us Latour refers to this as 'irreduction'—or the idea that no single thing can be truncated to another. Irreduction's consequence is that in most cases inter-object relations are devoid of intimacy or mutual-knowing, 'Being' for different objects is completely distinct and thus "objects only unlock each other's realities to a certain extent" [18]. Although incandescently challenging for us human objects to comprehend, this view of ontology is evocative, powerful, and represents an enticing philosophical renaissance "the epistemological tide ebbed, revealing the iridescent shells of realism they had so long occluded" [5].

Beyond a shared rejection of correlationism there is much disagreement between OOO scholars. Our interpretation aligns with that Ian Bogost presents in *Alien Phenomenology* [5]. Of particular influence is the notion of *carpentry*; the practice of creating "machines" that attempt to reveal clues about the phenomenology of objects. While it's accepted that objects' experiences can never be fully understood, the machines of carpentry act as proxies for the unknowable. They proffer a "rendering satisfactory enough to allow the artifact's operator to gain some insights into an alien thing's perspective" [5:100]. He cites a range of examples some of which he created as deliberate acts of carpentry whilst others

[4] cf. https://en.wikipedia.org/wiki/E.T._the_Extra-Terrestrial_(video_game)

simply demonstrate his argument serendipitously. One of Bogost's examples is software to visualise how a 90s games console stores and constructs sprites and palettes in the limited memory available, the result is a unique view on the connection between the 'raw' image storage, and the game as we see it on the screen (ibid). Another example, the *Latour Litanizer*, is a carpentered machine which queries Wikipedia, calls upon the random article feature, extracts the article title, and presents a list of several random titles. While its instrumental purpose is to quickly and easily generate Latour-like litanies, it also provides a portal of sorts into the interior reality of Wikipedia's content: "Not only does the diversity and detachment of being intensify with each fresh litany, but those very qualities also invite further discussion of the object in question at Wikipedia" [5:96].

Whether achieved by leveraging computer code or some other craft, it is "through the making of things we do philosophy" [38]—*that* is the essence of carpentry. Wakkary et al. do their carpentry through material speculations (ibid), and while Bogost sees himself as a philosopher-*programmer*, he notes that philosopher-chefs, philosopher-astronomers, and philosopher-mechanics are all uniquely equipped as carpenters in their own right. In our case, we are exploring the possibilities for philosopher-*designers*. Ensuring that OOO finds an outlet in some kind of applied practice is, in fact, crucial. It lends OOO a concrete legitimacy that metaphysics usually evades:

> *"If a physician is someone who practices medicine, perhaps a metaphysician ought be someone who practices ontology. Just as one would likely not trust a doctor who had only read and written journal articles about medicine to explain the particular curiosities of one's body, so one ought not trust a metaphysician who had only read and written books about the nature of the universe."* [5:91]

Hence, *material* engagements with OOO are what make the theory compelling, and carpentry is the process by which that engagement happens. Having realised that computers have, by virtue of programming languages, relatively accessible inner worlds, Bogost realised computers are a particularly compelling place to do carpentry. There is some shared ground between Bogost's computer-centric approach to OOO and the approach that we adopted with this work. We might say that computer programmers, emboldened by the ultimate control code has over the computer, allows them to 'play God' (within the realm of the computer or system they happen to be programming). This demiurgic gift affords the philosopher-programmer a great deal of freedom to explore the objects of the computer realm. As we discuss below, a similar quality is afforded when designers unshackle themselves from the preconceptions of contemporaneous truths, and, with new freedom, make speculative forays into the near future.

3.2 Design Fiction

Design usually seeks to answer questions, and thus to create futures. Speculative design, in contrast, uses design to *asks* questions about possible futures[5]. Hence the family of approaches which we collectively refer to as speculative design do not aim to create a products for sale, or that necessarily solve a problem, rather they go through the motions of design in order to elicit thought and provoke deeper understandings [2,11,12]. There are many nuanced views on the speculative design landscape which are beyond the scope of what we can address in this paper[6], however the specific method of speculation we are employ is *design fiction*.

There are a number of concurrent yet incongruent perspectives on what design fiction is, a disagreement that merges with discussions about the most productive ways to create and use them. The school of thought referred to as 'Design Fiction as World Building' [10] most exactly describes the approach we adopt with this work. The world building approach argues that design fiction is the creation of multiple artefacts that, when viewed together, describe the coordinates of, or 'entry points' into, a fictional world (ibid). Entry points also tend to depict parts of that world at different scales (figure 2). So, a given constituent artefact of a design fiction may either represent a large area of the world (providing a 'zoomed out' summary view), or a smaller area (providing a 'zoomed in' detail view).

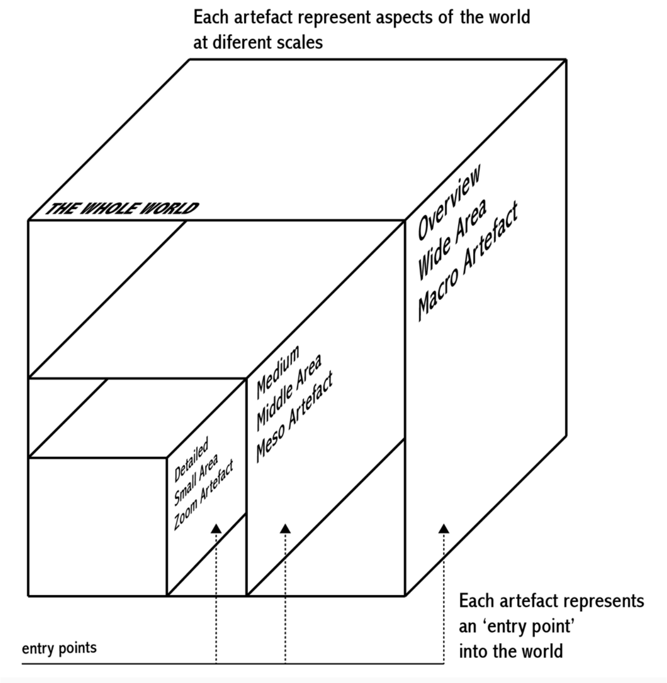
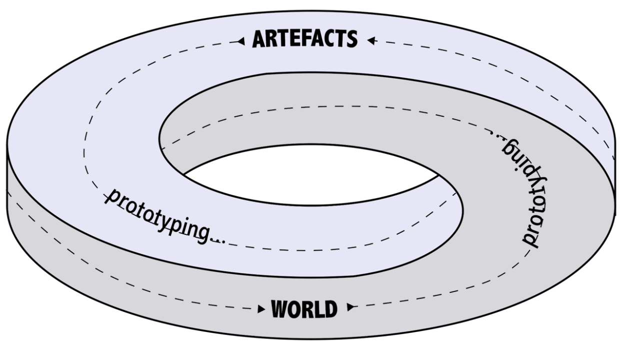

Each artefact represent aspects of the world at diferent scales
THE WHOLE WORLD
Overview
Wide Area
Macro Artefact
Medium
Middle Area
Meso Artefact
Detailed
Small Area
Zoom Artefact
Each artefact represents an 'entry point' into the world
entry points

Figure 2. Design Fiction as World Building

[5] Although more couched in the related practice of critical design, this asking/answering contrast is summed up nicely in "A/B" [26]

[6] For an overview and disambiguation of critical and speculative design we recommend Dunne & Raby [12] and Tonkinwise's review of the same text [35].

By creating multi-scaled worlds like this, design fictions produce a reciprocal prototyping relationship (figure 3); the artefacts define the contours of the fictional world and simultaneously prototype the nature *of* that world. *Meanwhile*, the world that emerges from the artefacts reciprocates and prototypes the nature of those artefacts (ibid). As we discuss below, we adopted the world building strategy for this research. We also suggest that both the individual artefacts, and the whole design fiction, may be seen in terms of Bogostian carpentry. Returning briefly to the notion of a carpenter playing God by manipulating computer code, the same logic plays out with design fiction but rather than the subroutines, APIs and procedure calls of the computer domain, a design fiction-philosopher has the texture and contours of the artificial world—and the artefacts that define those attributes—at their disposal.



In design fiction as world building multiple artefacts come together (left) to define multiple entry points into an artificially created world. Each entry point describes that world at a different scale. The effect is a reciprocal prototyping relationship, where the world is prototyping the artefacts and the artefacts are prototyping the world (right).

Figure 3. Design Fiction as World Building's reciprocal prototyping relationship.

3.2 Carpentry for the 21st Century; Prototyping Informed by Design

Our discussion and critique of PbD defines the area with which our research is concerned. The notes on OOO and Design Fiction introduce a theoretical position and method which we used to explore that area. Here we recount the details of two projects which probed these ideas, and in doing so laid the foundations for our IbD proposal. These Design Fictions were built around two fictional technologies. The first involved a fictional smart kettle, Polly [22] and the second describes a privacy-enhancing technology named Orbit [23]. In the following we discuss how both projects led us towards the IbD thesis.

3.3 Polly, The World's First *Truly* Smart Kettle

Our decision to build a Design Fiction world around a kettle (as opposed to some other object) was driven by two factors. First we were motivated to build upon the trope of mundane domestic devices as exemplars of IoT use cases, e.g. the much talked about smart refrigerator [1]. Second there are several existing smart kettles available in the consumer market which exhibit common concerns relating to IoT privacy, security and trust.

In the development of this Design Fiction we produced several artefacts to help us conceive of the world where Polly makes sense. These include a press release describing the product and its history, product packaging, and user interfaces. Through the creation of these artefacts the texture and detail of Polly's world emerge, in which the kettle becomes situated [34]. The press release we created describes many of the kettle's features, these include intelligent notifications, synchronization with social media feeds, automatic updates, voice activation, usage tracking, location-based boiling, *JustRight* smart fill level. Beyond revealing features, the press release also tells us more about the world that Polly exists within, for example the product was originally crowdfunded before subsequently being bought out by *Amazon's IoT* division. Finally, the press release also reveals that Polly is accredited by a government IoT regulator named *OfIoT* and as part of its accreditation utilises an alternative to current standards for transferring data over the internet named *Minimum Necessary Datagram Protocol*.

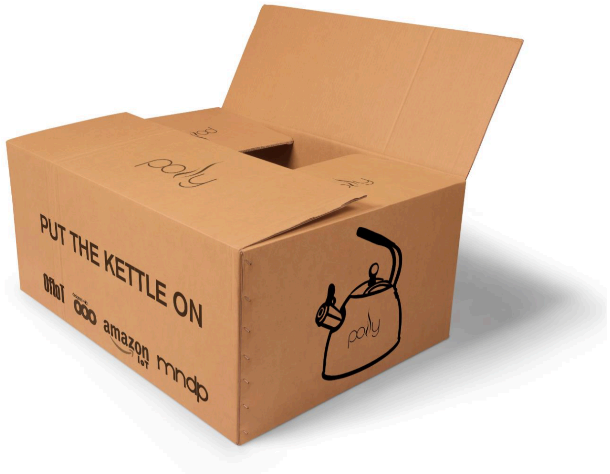


Figure 4. Polly's packaging, featuring logos for OfIoT, Amazon IoT, and MNDP.

Our assumption when building Polly's fictional world was that in the future the pervasiveness and ubiquity of data collecting devices will grow hand in hand with IoT adoption [33]. Presumably IoT kettles will (continue to) collect data too. The visibility of the data shared by these devices today is at best opaque and at worst absent. In this way shielding the user from the underlying data transactions. While PbD principles *may* protect the user from wanted or nefarious processing of their personal data, on occasions where that sort of processing is *necessary* to facilitate the device's functionality, on those occasions IbD would suggest making the underlying processing visible. We may liken this to an autonomous car that would chose an optimized route to its destination. Although optimized travel is desirable, *if* the car was unable to reveal precisely what that route was, it would not be surprising to feel somewhat mistrustful of it.

Responding to this need we constructed two IbD informed features.

Figure 5 shows a Polly timeline with several events taking place over a period of several hours on 15th April 2017. Each event is coded with one or more of 4 icons. A 'home' icon is used if the data transaction is between the kettle and another device on the local network. Two cloud icons, one with an up arrow and one with a down arrow, indicate if data is going to (up) or coming from (down) the Internet. Finally, a 'gear' icon is used to denote whether this specific data transaction is having a direct effect on the operations or configuration of the kettle or other networked devices. From the timeline, we can tell that Polly was dormant for over 4 hours since the 'daily cloud pingback', which uploads usage data to the cloud and downloads configuration, security, and update data from the cloud. Because this activity involves upload, download and configuration it includes those three icons.

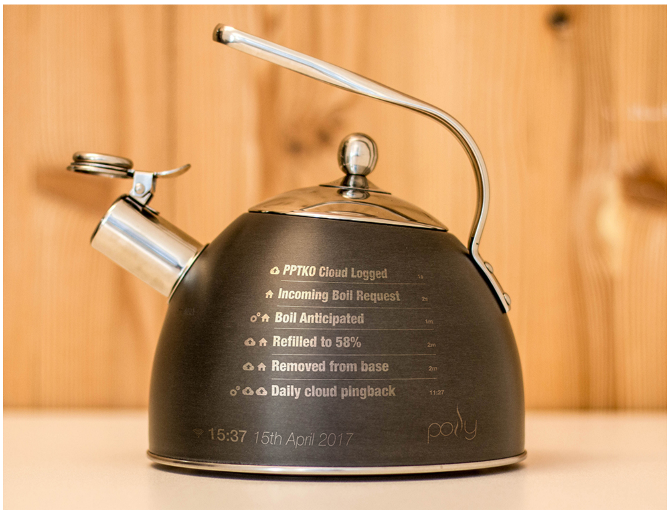


Figure 5. Polly's IbD timeline.

Next we can tell that the kettle is picked up from its base, refilled to 58% full, at this point it the software inside Polly anticipates (based on being refilled and previous user behavior patterns), that it will be boiled soon. We can see that removing the kettle from the base and refilling it result in immediate sharing of data to the cloud. The anticipation event however does not share data to the *cloud*, but *does* share data with the home's smart meter and other appliances to inform them of an impending power-consumption spike. Next, we see an incoming boil request, initiated from within the home, hence no upload or download, this is swiftly followed by a 'PPTKO' event ('Polly Put The Kettle On'), which *is* logged to the cloud. By interacting with the timeline details could be revealed showing precisely what data was sent or received, where it was sent or received from, and what purpose it plays in the constellation.

Figure 6 depicts a volumetric representation of the data uploaded from Polly, downloaded *to* Polly, and moving to or from Polly around the local network. This display differs from the timeline in that we cannot tell from it *why* data is moving around. However, what we *can* tell is the relative amount of data this smart kettle consumes and gives off, as well as the relative volume of those external transactions in relation to any 'chatter' with other devices on the local network. Both displays are intended to be used in conjunction with each other such that Polly is quite transparent about to what it communicates, for what purpose, and what – in terms of volume – the significance of that communication is. We might infer from these two displays that Polly 'gets' much less from the cloud than it 'gives'. Our Design Fiction does not explicitly communicate *why* this is the case. It could be that Polly's voice recognition software relies on the cloud, hence large audio files are uploaded frequently. It could be that because Polly appears to log almost every event it detects with its cloud provider (figure 6) and thus over time the volume of data builds up.

Potentially Polly could have been compromised, and the large upload volume is because Polly is part of a botnet. The 'truth' of why Polly uploads so much data is not, in fact, important to the paper and is a piece of information that will remain in the interior of the Design Fiction world. The world we have built for Polly to exist in, within which we have prototyped two features inspired by concerns to do with privacy and opportunities presented by OOO, serves as a first step to explore how thinking in terms of a flat ontology can be beneficial for the design of the IoT.

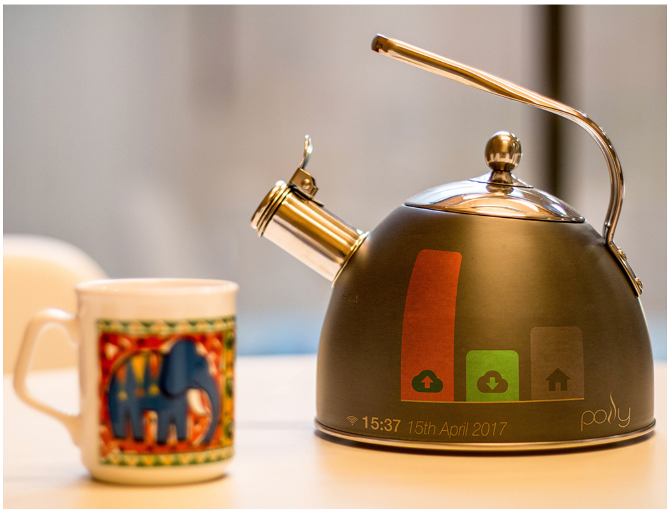


Figure 6. Polly's data usage graph.

3.3 Orbit Privacy

This project specifically focused on the interaction between a user and a technology whereby the user consents for their data to be collected, profiled, or otherwise used. We did this with the GDPR's specific protections in mind. Although legal interpretations are so far untested in courts the articles of the GDPR theoretically protect the right:

- To be aware what personal data is held about an individual;

- To access any personal data that is held;
- To rectify inaccurate personal data that is held;
- Of data portability (i.e. to extract data in a readable form to be taken elsewhere);
- To refuse permission for processing or profiling of personal data;
- That any consent obtained relating to personal data must be verifiable, specific, unambiguous and freely given.

The apparatus of consent (i.e. how information is presented to users, and how that consent is recorded) is the problem area that became of particular interest to us. Although some progress has been made recently, for example pre-ticked checkboxes non-consensual cookie usage were both outlawed in Europe in 2011[7], tick boxes for users to indicate they have, understood, and agree to conditions of use is still the norm. There are fundamental problems with this approach, the most obvious being that while pre-GDPR laws assume a tick in a box as legal consent, in practice it is very rare that users *actually* have read the terms, and even less so that they have understood them. Crudely but vividly demonstrating how such agreements are not an effective way to gain meaningful consent, a 2016 study found that only 25% of participants looked at the agreement at all, and only 2% could demonstrate reasonable comprehension [25]. One-size-fits-all approaches, whereby user agreements are written in such a way as to obtain *all* the permission the device could *ever* need, structurally remove the ability for users to be selective about which features of a system they actually want to use, and thus denies them the GDPR protection for 'specific unambiguous' consent. These systems also fail to account for temporality meaningfully; once consent has been given it is often difficult, and sometimes impossible, to revoke all or part of it at a later date.

We elected to use an IoT lock as the fulcrum, around which other aspects of the Design Fiction world would coalesce. Inspired by IoT locks that already exist on the market[8] the fictional lock has the following features:

- Keyless opening using NFC;
- Geofencing (automatically lock/unlock depending on user's location);
- Providing guests temporary access via smartphone;
- Voice activation (via a voice agent such as Amazon Echo);
- Interfacing with other services (via integrative platforms such as IFTTT).

In terms of the design problem, each feature requires a subtly different relationship with collected data, where data is stored, and how it is processed. Keyless opening using NFC operation only requires that data be stored within the user's own network; geofencing requires that data be processed by the lock company; and services such as IFTTT could lead to data being shared with any number of 3rd parties. Given that the design fiction's primary concern was GDPR, we opted to give technical implementation only cursory consideration and working around the assumption the lock is activated, via a hub, by an IoT radio standard such as ZigBee and that suitable APIs facilitate integration with external services such as IFTTT.

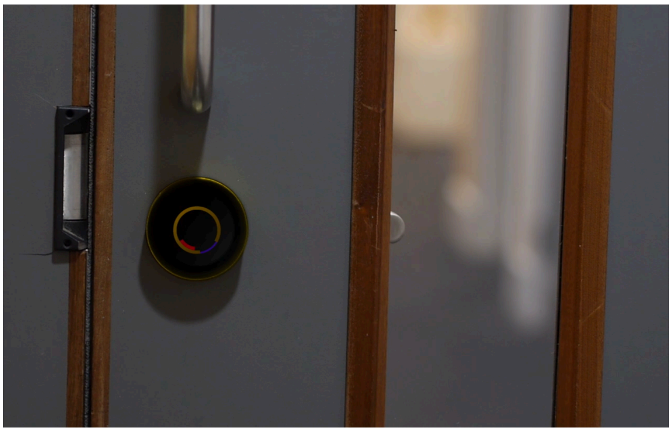


Figure 7. Still from supporting video showing the simple lock design. The electronics are housed in this plate which would replace one side of the standard lock plate with the remainder of the lock mechanism remaining the same.

Initially we wished to design a kind of map that could be used during the consent process to illustrate to a user what data would go where, so that they would be, insofar as the consent process, be *informed by design*. This task, however, was fraught with difficulties arising from how complex potential data-relationships are, even in relatively straightforward IoT systems such as a smart lock. Figure 8 illustrates a data scenario around an IoT lock which has been configured to trigger smart lighting to turn on when the user unlocks the door. While the cause and effect are straightforward and visible to the user (opening the door results in the lights coming on), there are in fact at least three cloud services behind the scenes making the hardware work, and as shown in this example there may be unknown 3rd parties using the data too. To translate this into a static map that details absolutely and concretely where data goes and when, is not viable. Confounding the already difficult task, our human appreciation of context makes the challenge *even harder*. To cartographically represent or respond to each human object's understanding of context-specific 'acceptability' (i.e. when it is okay to share data and when it is *not* okay) is something that needs to be done on a case-by-case basis [27].

To resolve our mapping-misgivings we needed to make two compromises, although this changes the nature of the design space it does not hamper our enquiry's overall aim to explore practice and philosophy. First, we reduced the scope of our interest from 'GDPR compliance' to 'personal identifiability'. Second, we had to reject the wholly deterministic view that our exploration of data packets brought, and instead build a

[7] http://www.bbc.co.uk/news/world-europe-15260748

[8] cf. http://uk.pcmag.com/surveillance-cameras/77460/guide/the-best-smart-locks-of-2017

map with the ability to communicate aspects of context, risk, and probability. Hence, it turned out that our dalliance with OOO took a route that we had not initially expected. We anticipated that exploring the tiny ontologies of the IoT lock itself, the data it produces, and its users, would lead us toward carpentry applicable to one of *those* objects, what *actually* happened is that we arrived at an artefact of carpentry around an entirely new object—one that communicates the *likelihood of identifiability*—and whose own tiny ontology, offers a new way to view any specific assemblage of devices, services, data, and people. By communicating the chance that a person could be identifiable based on the data associated with device use, and presenting that in terms of whether the data is held on devices owned by a user, servers owned by companies the user knows, or servers owned by 'anyone else', we aimed to have defined a construct that could represent both sides of the human/computer dichotomy that OOO helped us comprehend.

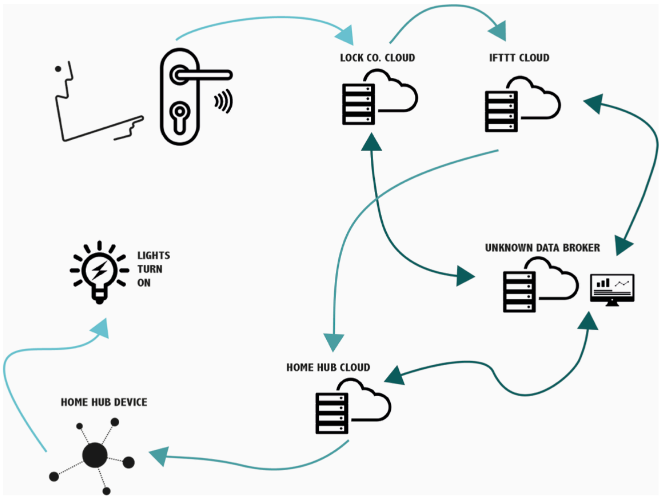
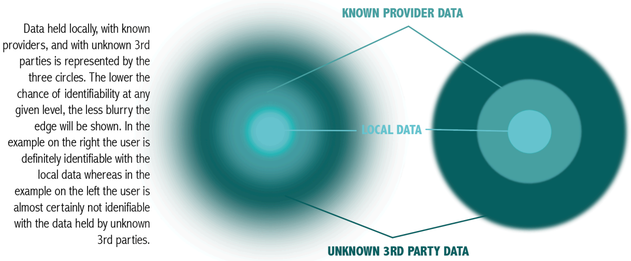


Figure 8. Triggered by the user unlocking a door data flows across and is processed on different networks and does not necessarily have a specific end point.

The most basic forms of our aware identifiability maps that appreciate notions of risk and probability, are shown in figure 8. Due to some metaphorical and visual similarity to the Bohr model of the hydrogen atom[9] we have referred to these as 'Orbits', or Identifiability Orbits. These maps represent data that is generated, stored and processed as part of an IoT system, and specifically *where* that data is held. A circular band represents each 'level' of data and our key privacy construct of identifiability is communicated by how sharp or blurry the edge of that level is. Hence, if the circle is the middle is razor-sharp, it is almost definite that the user could be identified by the data at that level. The blurrier the edge of any level is, the less likely it is that a user could be identified.



Data held locally, with known providers, and with unknown 3rd parties is represented by the three circles. The lower the chance of identifiability at any given level, the less blurry the edge will be shown. In the example on the right the user is definitely identifiable with the local data whereas in the example on the left the user is almost certainly not idenifiable with the data held by unknown 3rd parties.

Figure 9. Early prototype design for the identifiability Orbits.

Exploring how the design might be implanted, and how a user might interact with it, we implanted the Identifiability Orbits into our design fiction world by creating a film that depicts a user adding a lock to their smart home. The interaction in our film is triggered by instructing a voice agent to detect new devices; once the lock is detected the home's, the voice agent instructs the user to use the supporting 'Orbit Privacy App' on the user's phone so they can configure their privacy settings. By using a slider, the various functions of the lock can be enabled/disabled and the data implications of those choices visualised using an Identifiability orbit.

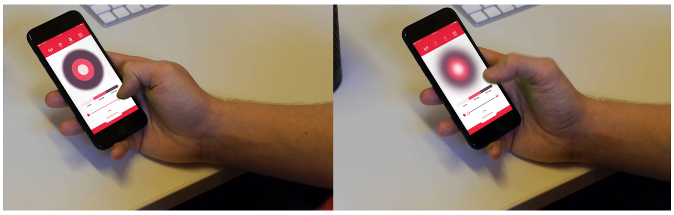


Figure 10. Stills from the design fiction film. A user uses the Orbit-based privacy app to configure which functions their IoT lock will have permission to use. On the left they have enabled maximum functionality, Orbit updates to show that the user will most likely be identifiable at all levels, although that is not certain with third parties. The inverse is shown on the right.

With the basic interaction demonstrated the film proceeds to demonstrate how a user may use such an app to dynamically modify their choices (figure 10). In our scenario, the user originally configured their lock for maximum privacy. The scenario extends to show that, if notified by a delivery company who require access to the house, the Orbit app communicates to the user identifiability implications of the data flow associated with provisioning temporary access to the delivery company, before revoking it again once delivery is completed. Although this work was completed before it was announced, this is a data flow very reminiscent of that supporting Amazon Key[10].

4 Discussion

Our OOO-informed Design Fictions work within this sentiment: "Security by design and privacy by design can be

[9] https://en.wikipedia.org/wiki/Bohr_model

[10] https://www.theverge.com/2017/10/25/16538834/amazon-key-in-home-delivery-unlock-door-prime-cloud-cam-smart-lock

achieved only by design. We need a firmer grasp of the obvious" [28]. While privacy is an extremely challenging notion to meaningfully grasp, and involves a huge amount of ultimately particularity, our argument suggests that concepts like PbD, or the GDPR's article 25, are misleadingly reassuring. By referring to technologies both specifically (pseudonymisation, for example) and ambiguously ('state of the art', for example) PbD masquerades as a 'solution' to privacy, yet in practice it arguably represents a Heffalump trap. *Informed by Design* approaches may offer a counterpoint. Where PbD tries to be methodical, IbD is intentionally contingent; where PbD strives for answers, IbD intentionally highlights questions; PbD may be criticised for its ambiguity, IbD is *intended* to be interpreted. A simple way to measure something that is implemented with IbD at its core is to ask, "are the users informed about any privacy implications of this system?" Although summarising the proposal to such a simple question makes it seem straightforward, and perhaps simple, alas it is not.

First, as demonstrated particularly in the Orbit example, refining the ideals of OOO into an actionable process is not always straightforward—translating philosophy into design practice, although fruitful, is challenging [23,30]. Second, to translate context-rich and human abstractions such as 'privacy' into specific design constraints falls under the complications of any given design's ultimate particularity. However, the argument presented here aims to articulate that the barebones of PbD are not, alone, enough to meaningfully develop interactive systems that go beyond the letter of regulations such as GDPR to *also* respect their ideals. IbD, however, goes some way to resolve this issue.

Above all else this paper advocates for a more considered engagement with the complexities of designing for the IoT such that the non-functional requirements engineers have traditionally struggled to deal with, may be approached heuristically. In order to drive the adoption of acceptable IoT devices designers, engineers and policy-makers need to set aside beliefs that a system can, *by design*, circumvent any concerns about a system's privacy. Instead, the IoT needs to be designed with a considered approach to privacy that accepts IoT devices *do* pose problems for individuals' privacy, but that those problems can be tempered by transparently *informing* users about how the system works, and consciously *not* attempting to trap Heffalumps.

Acknowledgements

This work was made possible by the RCUK PETRAS Cyber Security for the Internet of Things Research Hub (petrashub.org) under EPSRC grant EP/N02334X/1.

References

1. Charles Arthur. 2014. Internet fridges: the zombie idea that will never, ever happen. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2014/jan/07/internet-fridge-lg-ces-2014.
2. James Auger. 2013. Speculative design: crafting the speculation. *Digital Creativity* 24, 1: 11–35.
3. Julian Bleecker. 2009. Design Fiction: A short essay on design, science, fact and fiction. *Near Future Laboratory*.
4. Julian Bleecker. 2013. A Design Fiction Evening with the Near Future Laboratory. Retrieved from http://vimeo.com/84826827.
5. Ian Bogost. 2012. *Alien phenomenology, or, what it's like to be a thing*. U of Minnesota Press.
6. Levi R. Bryant. 2011. *Democracy of Objects*. Open Humanities Press.
7. Ann Cavoukian. 2012. Operationalizing privacy by design. *Communications of the ACM* 55, 9: 7.
8. Ann Cavoukian and Jeff Jonas. 2012. *Privacy by Design in the Age of Big Data*. .
9. William E. Connolly. 2013. The "New Materialism" and the Fragility of Things. *Millennium: Journal of International Studies* 41, 3: 399–412.
10. Paul Coulton, Joseph Lindley, Miriam Sturdee, and Michael Stead. 2017. Design Fiction as World Building. *Proceedings of the 3nd Biennial Research Through Design Conference*.
11. Anthony Dunne. 2006. *Hertzian Tales: Electronic Products, Aesthetic Experience, and Critical Design*. The MIT Press.
12. Anthony Dunne and Fiona Raby. 2013. *Speculative Everything*. The MIT Press, London.
13. Christopher Frayling. 1993. Research in Art and Design. *Royal College of Art Research Papers* 1, 1: 1–9.
14. William Gaver. 2012. What should we expect from research through design? *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*: 937.
15. Maximilian Von Grafenstein and Christina Douka. 2017. The "state of the art" of privacy- and security-by-design (measures). *Proceedings of MyData*.
16. Peter Gratton and Paul J Ennis. 2014. *The Meillassoux Dictionary*. Edinburgh University Press.
17. Graham Harman. 2002. *Tool-Being: Heidegger and the Metaphysics of Objects*. Open Court Publishing, Peru, IL.
18. Graham Harman. 2002. *Tool-being: Heidegger and the metaphysics of objects*. Open Court Publishing.
19. Graham Harman. 2015. Object-Oriented Ontology. In *The Palgrave Handbook of Posthumanism in Film and Television*. Palgrave Macmillan UK, London, 401–409.
20. J. Lindley and P. Coulton. 2015. Back to the future: 10 years of design fiction. *ACM International Conference Proceeding Series*.
21. Joseph Lindley. 2015. A pragmatics framework for design fiction. *Proceedings of the European Academy of Design Conference*.
22. Joseph Lindley, Paul Coulton, and Rachel Cooper. 2017. Why the Internet of Things needs Object Orientated Ontology. *The Design Journal* 20, sup1:

S2846–S2857.

23. Joseph Lindley, Paul Coulton, and Rachel Cooper. 2018. Turning Philosophy with a Speculative Lathe: Object Oriented Ontology, Carpentry, and Design Fiction. *In Press*.

24. Joseph Lindley, Anya Skatova, Rob Proctor, Coulton Paul, and Rachel Cooper. *PETRAS Landscape Review and Gap Analysis for Adoption and Acceptability*. .

25. Jonathan A. Obar and A. Oeldorf-Hirsch. 2016. The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services. *The 44th Research Conference on Communication, Information and Internet Policy*.

26. Fiona Raby and Anthony Dunne. 2009. A/B. Retrieved October 27, 2014 from http://www.dunneandraby.co.uk/content/projects/476/0.

27. m.c. schraefel, Richard Gomer, Alper Alan, Enrico Gerding, and Carsten Maple. 2017. The Internet of Things: Interaction Challenges to Meaningful Consent at Scale. *interactions* 24, 6: 26–33.

28. Stuart S. Shapiro. 2010. Privacy by design. *Communications of the ACM* 53, 6: 27.

29. Sarah Spiekermann. 2012. The challenges of privacy by design. *Communications of the ACM* 55, 7: 38.

30. Liesbeth Stam and Wouter Eggink. 2014. Why Designers and Philosophers should meet in school. *Proceedings of the E&PDE 2014 16th International conference on Engineering and Product Design, University of Twente, The Netherlands*.

31. Bruce Sterling. 2005. *Shaping Things*. The MIT Press.

32. Bruce Sterling. 2012. Bruce Sterling Explains the Intriguing New Concept of Design Fiction (Interview by Torie Bosch). *Slate.com*. Retrieved February 9, 2014 from http://www.slate.com/blogs/future_tense/2012/03/02/bruce_sterling_on_design_fictions_.html.

33. Bruce Sterling. 2014. *The Epic Struggle of the Internet of Things*. Strelka Press.

34. Lucy Suchman. 1987. *Plans and situated actions: the problem of human-machine communication*. Cambridge University Press.

35. Cameron Tonkinwise. 2014. How We Intend to Future Review of Anthony Dunne. *Design Philosophy Papers* 12, 2: 169–187.

36. I Van der Tuin and R Dolphijn. 2012. *New Materialism: Interviews & Cartographies*. .

37. Nicholas Vollmer. 2017. Article 25 EU General Data Protection Regulation (EU-GDPR). Retrieved January 15, 2018 from http://www.privacy-regulation.eu/en/article-25-data-protection-by-design-and-by-default-GDPR.htm.

38. Ron Wakkary, Doenja Oogjes, Sabrina Hauser, et al. 2017. Morse Things: A Design Inquiry into the Gap Between Things and Us. *Proceedings of the 2017 Conference on Designing Interactive Systems*: 503–514.

39. Summaries of Articles contained in the GDPR. Retrieved September 15, 2017 from http://www.eugdpr.org/article-summaries.html.