

Exploring the future of data-driven product design

Katerina Gorkovenko
University of Edinburgh
Edinburgh, United Kingdom
K.Gorkovenko@ed.ac.uk

Dan Burnett
Lancaster University
Lancaster, United Kingdom
D.Burnett@lancaster.ac.uk

James Thorp
Lancaster University
Lancaster, United Kingdom
J.Thorp1@lancaster.ac.uk

Daniel Richards
Lancaster University
Lancaster, United Kingdom
D.Richards@lancaster.ac.uk

Dave Murray-Rust
University of Edinburgh
Edinburgh, United Kingdom
D.Murray-Rust@ed.ac.uk

ABSTRACT

Connected devices present new opportunities to advance design through data collection in the wild, similar to the way digital services evolve through analytics. However, it is still unclear how live data transmitted by connected devices informs the design of these products, going beyond performance optimisation to support creative practices. Design can be enriched by data captured by connected devices, from usage logs to environmental sensors, and data about the devices and people around them. Through a series of workshops, this paper contributes industry and academia perspectives on the future of data-driven product design. We highlight HCI challenges, issues and implications, including sensemaking and the generation of design insight. We further challenge current notions of data-driven design and envision ways in which future HCI research can develop ways to work with data in the design process in a connected, rich, human manner.

Author Keywords

Data-driven design; Design research; IoT; Smart devices; In the wild; Human-centred design

CCS Concepts

•**Human-centered computing** → **Human computer interaction (HCI)**; **User studies**;

INTRODUCTION

Data-driven approaches are standard within the digital industry, where usage analytics and A/B testing are routinely employed to inform new features or layouts intended to improve the original design. The growth of connected products, such as *smart* consumer devices, is creating opportunities to similarly rethink the way that physical products are developed. The

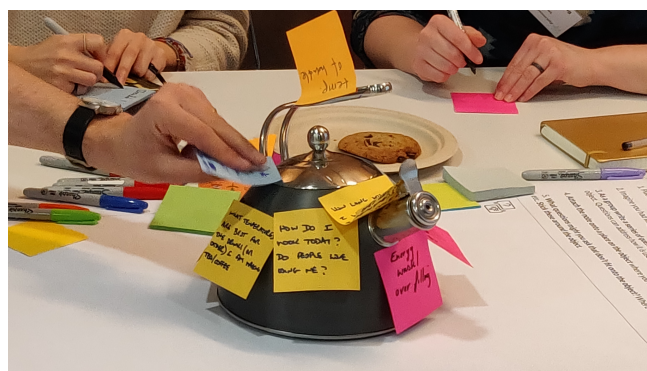


Figure 1. A kettle annotated with design questions that could be explored through data collection, generated by workshop participants at our Lancaster event.

design of physical products for commercial applications requires rigorous prototyping and testing before manufacture to establish that the product will be functional, robust and safe. However, once released onto the market, products can be used in contexts and for purposes that the designers may not have anticipated or observed within controlled testing. Insights into these contexts could be valuable for product design and development [29]. Within academia a variety of research methods, such as technology probes and design ethnography, have illustrated how behaviours observed in a lab setting differ from those in the wild [1]. Challenges remain around scaling those methods up for industrial adoption, due to their grounding in the collection of thick qualitative data, which can be costly to collect and analyse [1].

The market is becoming increasingly saturated by consumer products made "smart" through computation and sensorization, such as thermostats, vacuum cleaners, cars, and home assistants. Their core functionalities often rely on the processing of live data, such as audio commands in the case of a home assistant. Auxilliary sensors generate data about the environment around the product and how it is being used, such as accelerometer data from a smartwatch. There is an opportunity to investigate the potential for this data to support in the wild research, which can in turn generate human-centred design

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI '20, April 25–30, 2020, Honolulu, HI, USA.

Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6708-0/20/04 ...\$15.00.

<http://dx.doi.org/10.1145/3313831.3376560>

insights that directly relate to user behaviour – designing *by, with and from* data [47].

There is a significant amount of literature associated with the use of simulated and real-time data streams to inform late-stage design processes associated with optimisation [51, 52], such as Building Information Modelling (BIM) for large-scale architectural projects [27] and product life-cycle management for high-value assets [48]. These systems are utilised late in the design cycle as the accuracy of the output is dependent on the information used in the models. For example, an energy model of a building needs detailed construction information which is generated after the feasibility and concept design stages. Many open questions remain about how such data can add value to design processes and at the scale of consumer products. As highlighted by [34], one of the major challenges is the difficulty that design practitioners experience in incorporating data-driven approaches to their work, which force them to use optimisation too early in the design process.

This work aims to map the space of choices available to design teams. To try and shed light on how real-time data streams might be better utilised to accelerate and enhance design cycles, we conducted a series of workshops with experts working within product design and IoT. Through a series of collaborative activities, they helped us investigate the following question:

What opportunities and challenges do IoT and product design experts identify for data-driven product design?

We use the phrase *data-driven product design* to broadly define the approach of utilising data from devices in the wild for design purposes. This is facilitated by the use of sensors and functional data, alongside various qualitative sources, such as feedback from users.

This paper sets out to scope HCI challenges in adopting real-time data within design. It contributes an expert perspective on the future of data-driven product design. The participants draw on their professional experiences, identifying the potential impact and best practices that may emerge around data-gathering. They further speculate on how design knowledge may be derived from data, and identify key questions that need to be addressed. Within the discussion we outline how these insights can be used for the development of new research practices and tools that support product design. While our primary concern is in understanding how industry might utilise data, we also discuss directions for further academic exploration.

BACKGROUND

Connected devices in the wild can generate an array of *use* data, which may be valuable for design. This can be functional data, such as a Nest thermostat gathering data about user preferences alongside room temperature readings to decide when to turn the heating on. It can be background information gathered to support the product's life-cycle, such as logs generated for predictive maintenance. Devices can also be used to collect research-driven design data, such as differences in use in an A/B test. Finally, they can collect data for other purposes, such as building a large corpus of voice recordings to support the development of better speech recognition software. Below we

outline industry and academia practices that utilise such use data and identify gaps in the way it is used to support design research.

Approaches to Data Driven Design

Digital products benefit from being continuously developed and optimised. Common research practices include A/B testing, usage analytics and crash reports [23, 32, 35]. Findings can be directly fed into software updates, design, and hyper-personalised content [23, 32, 46]. Research indicates that software engineers are especially interested in questions around user behavior, which can be explored through data [7]. Within the gaming industry, analytics utilise sophisticated visualisations and analysis tools to represent user behaviour, such as heatmaps and advanced spatial clustering techniques [6, 22]. Insights from these visualisations are used to improve the design of game maps and make believable adversarial bots [6].

In contrast, physical products are subject to an intensive design process pre-manufacture, including prototyping stages and user testing [55]. There are many places within the design lifecycle where data can affect the process, from early conceptual and requirements gathering stages through to optimisation and maintenance. Following the sections of the British Design Council's Double Diamond [19], in this paper we use the phrase *early-stage design* to encompass the discover, define, and develop stages where the majority of conceptual work takes place, and *late-stage design* to frame the delivery process, where we observe optimisation and refinements. This thinking compliments a holistic process which encompasses the full lifespan of the product, as discussed by work on research in the wild [1, 15, 18] that can affect thinking about design as well as optimisation.

There is a drive within industry to make use of big data collected in the wild, and data gathering has been seen to enrich and support product design, for example, the Hövding bicycle helmet was developed in response to 3000 staged accidents and 2000 hours of data recorded by cyclists [2]. This data can come from various sources, including sensors on the physical device, and can be used to facilitate predictive maintenance and service products [21, 44, 51]. Predictive maintenance practices are common within the aerospace industry, where sophisticated algorithms can help engineers make accurate predictions about when to replace and service components before issues occur [21]. As an example, Rolls-Royce lease rather than sell their jet-engines to airlines. In use, the engines generate telemetry data, which is used to diagnose faults at the earliest opportunity and service them responsively [5]. Although widely accepted for maintenance this research technique has limited recorded impact on design.

Tao et al. [51] envision that big data generated throughout a product's life-cycle can be visualised and utilised for design. They propose a cyclical data-driven process, where digital twin visualisations - real-time representations of physical objects with their associated spatially bound data - are utilised for manufacturing, design, and servicing of products. While the authors discuss the potential impact of this type of visualisation on industries that manufacture high value goods (e.g.

aerospace), it is less clear what applications a similar data-driven approach might have within mass-produced connected products.

In The Wild HCI Research

Human Computer Interaction (HCI) research of product use in the wild has much to offer in terms of methods, including the use of research products, technology probes, experience sampling (ESM), design ethnography, and field studies [1, 13, 33, 37, 41], where researchers collect contextually grounded insight. These methods often rely on the collection of dense and detail rich qualitative data, but here we would focus on a few of those that also utilise remote quantitative data capture.

Contingent experience sampling, uses triggers from data to send response requests to participants [24, 45, 53, 58]. An example of this is an ESM tool called Paco, which allows researchers to design surveys, triggers and log device sensor data [24]. Although contingent ESM has the potential to be disruptive to participants' activities, the method alleviates issues around retrospective recall present within self-reporting approaches, such as diary studies [53]. Furthermore, unlike conventional ESM, the response request questions can be relevant to the activity the participant is performing [53].

Contemporary ethnographic methods, such as thing ethnography and ethnomining, collect quantitative data from various sources to support research. Ethnomining collects data from a smart device, which is then visualised, and used in contextual inquiry and collaborative sense-making interviews with participants post-hoc [4, 8]. Meanwhile, thing ethnography, leverages a thing perspective by attaching cameras and sensors to objects, which are used as part of ethnographic research [16, 28, 29]. An example of an investigation through sensorizing devices can be seen in the work of Chang et al. [16], who investigated the relationships between vehicles and drivers in an urban setting using a series of scooters.

More relevant to our work is the emergence of data-enabled design, where technology probes collect quantitative data in the wild alongside qualitative responses to support design in the prototyping stage. Bogers et al. [9] used an iterative process to design a baby bottle, involving contextual inquiry and the deployment of sensorized technology probes, where the design process was supported by remote-sensor data capture alongside interviews. Similarly, Bourgeois [11] derived design insights about home energy systems from a longitudinal study in the wild, supported by data capture. Finally, Burnett et al. [14] merge academic research practices with industry led data visualisation, as discussed by [51], where digital twin representations of smart products are used as research tools to support designers and ethnographers. Insight generated through these methods in the wild can facilitate Human Centered Design (HCD), which focuses on the needs of users [30].

Emergent Issues With In The Wild Data Collection

Advances within in the wild HCI research have the potential to impact industry design practices but thorny issues remain, including ethics and research at scale. While research within academic contexts is subject to regionally enforced ethical codes of conduct, industry has to abide by varying data-privacy

regulations like GDPR in Europe [42], which requires the explicit need for consent [10]. It is yet unclear if time and resource heavy academic research practices that merge big and thick data from connected products, such as [9, 11], can be scaled up to industrial contexts beyond the prototyping stages of design.

Furthermore, data-collection more generally has been criticized for perpetuating industry practices that are seen to disempower the consumer like *surveillance capitalism* and *prosumption* [39, 60, 61]. Prosumption is characterised by companies harnessing consumer labour to increase profit [39]. Examples include customers using self-service checkout machines in a supermarket rather than a cashier, and the shifting of responsibility for recycling from the producer onto the consumer [57]. While prosumption can be seen as exploitative, the value associated with digital products is more varied in nature and distribution. Contributing online content can have benefits for the individual as well as the corporation - a YouTube vlogger gaining fame and associated perks for example [39]. Meanwhile, surveillance capitalism is an industry practice where data about online behaviour is collected and sold by companies for profit [60, 61]. Facebook and Google using click-through models to target users with highly personalised ads is an example of this. Finally, the way data is analysed and visualised can have far-reaching ethical implications [17], highlighting the importance of ethical considerations throughout the whole lifespan of data, from the moment of capture through to its utilisation.

Zuboff argues that surveillance capitalism translates to the use of smart products [60, 61], but when data collection and behavioral modelling is an integral part of a device's function, assessing its ethical ramifications can be ambiguous. This research is based in the belief that an open dialogue between industry and academia has the potential to identify the opportunities for data-driven product design research that can be scaled appropriately, be beneficial for both industry and society, and counter concerns of surveillance capitalism and prosumption. Despite advances in in-the-wild research methods around smart products, key questions remain around how to use this knowledge for design.

RESEARCH OVERVIEW

Workshop Structure

We conducted three workshops with a total of 32 participants from industry and academia who reported having expertise within IoT or product design. Each workshop was three hours long and took place across the United Kingdom, in Lancaster (P1-P9), Edinburgh (P10-P20) and London (P21-P32) in order to recruit a varied cohort of experts.

The workshops aimed to explore how the participants viewed the future of data driven product design, where connected products are used as research tools for their own re-design and development. Questions around how the process would work in practice were purposefully left open ended to investigate the different ways data can be gathered, turned into design insight, and fed into the design process. To introduce the participants to the idea of visualising live data from a connected

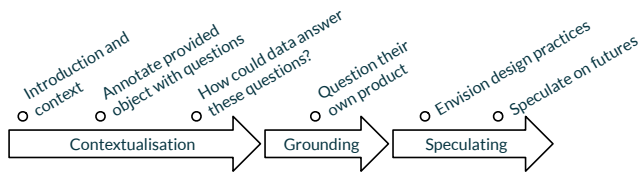


Figure 2. Workshop structure showing flow from contextualisation, through grounding in professional practice into speculation

product for design purposes, we presented a demo consisting of a bluetooth speaker fitted with sensors. Live data was transmitted by WiFi to a digital on-screen representation of the speaker, including values for acceleration, magnetometer, gyroscope, temperature, audio metadata, button presses, and volume change data. After presenting the demo, participants worked in groups of 3-5 people on five collaborative activities, which were centred around large A1 sheets of paper with instructions (Figure 2):

1. Generating design questions around an object that we provided, including kettles (Lancaster), electric scooters (Edinburgh), and audio-recorders (London) (See Figure 1).
2. Selecting a question from the first activity and exploring how it could be answered through data gathering.
3. Each group selected one of the products that the experts had worked on, and generated questions related to its design that could be answered through data. Fictional products were used where necessary, and groups repeated the exercise for multiple products if time allowed.
4. After an overview of the kinds of data that could be collected, the experts were asked to envision a day in the life of a designer who utilises a data-driven design approach, in particular new design practices driven by data gathering.
5. Finally, the participants were asked to identify their hopes, fears, and utopian futures of a data-driven design process as it relates to their own practice. With this activity we hoped to draw out preferable futures as mapped onto Voros's Futures Cone [3], which depicts projected paths in terms of possible, plausible, probable or preferable. This allowed the participants to envision positive possibilities for data-driven product design, free of practical implementation challenges.

Participants: Recruitment and Expertise

The experts were recruited with the help of several organisations, snowball sampling, and through emailing lists. We recruited from the industry connections of the larger Chatty Factories project this research is part of, an innovation centre in Glasgow advertised the events on their social media, a digital adoption agency in London hosted and advertised the London event on their website, and we were allowed to recruit during an industry meetup event in Edinburgh. We sent out invitations and promotional material to industry and academia experts, which were local to each workshop, in order to recruit from a wider spread of institutions and industries and accommodate for people's busy schedules and availability.

The experts reported having a mixture of expertise within product design (n=25), IoT (n=30), computer science (n=4), and management (n=4). They worked in industry (n=9), academia (n=11), and a combination of the two (n=12). The participant's experience ranged from six months to 25 years. Their average age was 37 years with SD of 9 years. To preserve their identity we refer to them as P1 to P32.

The experts drew from their professional experiences, which ranged across a variety of contexts and projects, including working with aircraft engines (P1), smart projectors (P3), food 3D printers (P32), bespoke smart products (P21), work-site management technology (P29), robots used as an educational resource (P14), solar power plants (P25), etc. Their professional roles included designers, computer scientists, lecturers, researchers, consultants, managers, business owners, and PhD students.

Data Collection and Analysis

The workshops were audio and video recorded, which were anonymously transcribed into 87,241 words. The workshops also generated written material captured on large A1 sheets of paper containing the collaborative activities.

Transcripts and written materials generated from the workshops were inductively thematically analysed [12] using NVivo. The lead researcher coded the data independently generating 233 codes. A segment of the codes was verified, discussed, and clustered with a group of four researchers. The clusters and renaming of codes was integrated into the Nvivo file and the rest of the codes were clustered in agreement with the discussions and changes generated during the analysis with the other researchers. Through an iterative process of independent clustering of codes by the lead researcher and verification and discussion with another member of the research team eight final themes were identified and agreed upon. The themes broadly relate to data sense-making, finding value for design, implications of the process, and ethical and implementation issues.

RESULTS

The experts contributed their visions, concerns, and perceived needs, as they relate to the future of data-driven product design. Below we present the themes, beginning with a short summary:

Live Data Gathering

The capturing of sensor and functional data to observe user behaviours, wear and tear, functional failures, and contexts of use. Participants identified that data collection from products in the wild can be achieved both through passive data gathering and actively engaging with users through triggered feedback requests. Passive data gathering included the collection of sensor data, and logs of device functions. To support sense-making around the relationship between design and data, experts explored correlating functional data with other information, such as the times of product use, function quality, fail detection, and the experiences of the user.

Different objects were annotated in each workshop, but the questions we explored with experts were the same. User behaviours were always central in discussions, in particular what factors contribute to increased or decreased use, leading to discussion of wear and tear and functional failures of products. The contexts of product use were also of interest, explored through sensor data, such as GPS tracking, temperature, chemical composition of substances around the product (e.g. air, water), pressure, movement, vibration, proximity, and the forces acting upon the product. A key consideration within this process raised by P5 and P27 was the "accuracy of the data", which can be influenced by sensor limitations or underlying data analysis issues.

Understanding these contexts and environments of use highlighted factors that may lead to faster deterioration of components, e.g. P14 worked with a robot to teach STEM skills in education, and revealed that the children repeatedly break parts of the robot. They used a "current sensor" within the joints of the robot to measure the pressure applied, and were "able to identify that the main reason this one was getting broken is that too much pressure was applied to the knee", which could be addressed in subsequent design iterations.

Participants thought that a data-driven design process has great potential to unravel some of the complexities of human behaviour, gauging the emotions, habits, and engagement of users through the use of audio, bio-information, use of functionality, and user identification. Understanding behaviours and motivations can in turn inform the design of better products that address underlying needs and expectations. For example when annotating an electric scooter in the Edinburgh workshop, the experts identified that by measuring the excitement of a person at certain speeds of an electric scooter the designer can explore if "speed [is] important of the joy of using" (P12) the product. Furthermore, by comparing behavioural data across a large number of individuals, the designer can identify different user groups based on "patterns in use" (P22).

"What you would want to do is to identify people within that population who you think are otherwise similar so that you can understand when the profile of one changes, is that because of a factor that is particular to them, or is it because it is affecting everybody?" -P16

Generate Design Insight

Explore research questions through A/B testing, usability tests and ethnographic explorations, supported by data dashboards, ML, and user feedback. A key potential for a data-driven design process is to "challenge [the] assumptions" (P29) of a designer about product use in the wild through the collection of factual supporting information. Participants suggested this could be achieved through research and testing, including "user testing" (P24), "A/B testing" (P18). Ethnographic research investigating the "perceptions of the user" (P25) could be supported by live contextual inquiry with "live feedback" (P14), while A/B testing would deploy different versions of a product for comparison: "if it was a small thing we could maybe rapidly prototype and design small changes like you might see with websites" (P32). Product 'speciation'

came up, where different user needs can be met by developing products with varying functionality:

"we can also then develop a persona around an individual, so we can then develop a product for certain geographies for example. Say if we are doing it in the far East vs America we could have quite different results in different products." -P14

Participants felt that for data gathering to be ethical and financially viable within the process, pre-defined research questions should inform the data gathering practices. For example, when annotating a kettle with questions P5 discussed: "I don't know if the barometric pressure of the environment affects the efficiency of the kettle so maybe you put a sensor on to answer that question and when you have you can take it off again". They favoured the initiation of in-the wild research early in the product development, even throughout the prototyping stage, placing sensors on a subset of existing products to answer a question to e.g. balance the development of two design attributes: "I would like data to support my decision" (P32). This approach has been explored in 'transition products', where existing domestic appliances are fitted with sensors to inform future designs, particularly for usability and environmental sustainability [25].

"I can make the case to my boss to make 50 or 500 of these devices because we really need to know things that we will get back into the design cycle." -P5

Dashboards for exploratory data analytics, the utilisation of machine learning, and the ability to gain feedback from users were seen as vital to the process. "Talking to the user" (P28) is a direct and reliable way to answer design questions. Data visualisation can empower designers to identify and question emerging correlations between user behaviour and product functions, and participants felt that data visualisation tools utilised within data-driven design should support collaborative data exploration (P10-P12). They suggested this could be achieved by producing views of amalgamated as well as raw data, allowing exploration across different timescales and parameters: "you have aggregate to make life easier because there is too much of it but the answer is in the disaggregated data and the question is how do you get to this happy medium" (P29). Combinations of machine learning algorithms could help to spot patterns, identify behaviours that deviate from the norm, and make predictions about the meaning of live data. Discussions often focused on the potential of such a system to make re-design recommendations or automatically assess the impact of a design change on its function and user engagement. P32 envisioned a design process where:

"You have taken a recommendation, you have looked at the compromise, you matched that against your brief, you virtually design it, you test it, you prototype it, you submit a change request, you get feedback." -P32

Product Development and Re-design

Ever evolving algorithms extend the product's capability and continuously optimise its design, while designers focus on developing products for the fringes of the user

group. We were particularly interested in how the experts envisioned translating data into a design change of the physical product. They had a sense that the process would be ideal for parameter optimisation, identifying failures, and understanding use, but found it more difficult to envision how it could support the designer's creativity. Some felt that the ultimate goal of the process would be that *"the product can redesign itself"* (P15), where ML algorithms continuously update and optimise the design based on user behaviour and manufacturing requirements. Others felt that to support creativity the process should be more open ended, allowing for designers to explore and question the data but ultimately draw their own conclusions and inspiration:

"Data would probably tend to lead you towards optimisation rather than innovation. Because it comes back to what is an average user, if you use machine learning on something it homogenizes it doesn't innovate." -P2

Data-driven design can make devices more intelligent, responsive, give users accurate predictions, and support a tailored user experience. Sensorized products already out in the wild could be made "smarter" by adapting to the ways they are being used. For example, P32's food 3D printer: *"because it is a chemical process, temperature influences a function, so if hot liquids are more runny, the print process might be changed in response to the temperature of the environment"*. The algorithms dictating the function of the product would need to understand the context of use, anticipate the needs of users, and adapt their function accordingly, *"designing themselves on the fly"* (P24). In this scenario the development of products is strongly dependent on the design of sophisticated AI.

The process of making products "smarter" can also impact the product's physical design, such as assessing the least intrusive sensors that allow them to sustain their functionality and identifying suitable changes to the physical interface of the product. Smart products could also extend their functionality beyond their original design, accessing external online data sources, communicating with other devices around them, or crowd-sourcing knowledge – the Edinburgh participants suggested that a smart scooter could tell the user which streets to avoid based on current traffic, or send information to the council about issues with road safety.

It was less clear how insights around product function and user behaviour would translate into changing aspects of the physical design. Ultimately creativity was seen to stem from ignoring the behavioral norms observed and allowing the designer to investigate and create for users at the fringes of the data set.

"Is your job to look at the data that the machine learning is generating, or is your job to just look at the outliers? So the job of the designer becomes either to ignore the data, or to look at the weird things in it." -P2

User-centred Design

A combination of live data and communication with users can enable user-centred design and facilitate co-design or co-creation. One of the primary values of utilising a data-driven product design process is in understanding users. P29

reflected that *"products are never used in the way people think"*. By gaining insight into people's habits, needs, behaviours, and engagement, we can create better user-experiences and more intuitive user interactions. In turn this process of making more user-centred products could have broader societal benefits, such as improving people's well-being.

Beyond optimisation and making a product "smarter", experts felt that feedback and communication is central to generating design insight. Questions remained about whether focusing narrowly on data would lead to design changes that are detrimental to the product and its desirability. P20 identified that physical sliders within a home media system may be wanted by users *"aspirationally"* because it *"makes [users] feel happy that they could switch it if they really wanted to"*, while the true usefulness and functionality of the product lies elsewhere.

The participants felt that live user feedback would contextualise the quantitative data coming in about their experience. An example of this is P32's consideration for how data gathering and user feedback can contextualise the designer's understanding of the performance of a 3D food printer:

"How it knows if it was printed correctly, so how that could be measured, whether you can ask if it was what someone was expecting [...] even if it was not necessarily what the designer intended the machine to do in that case, has it achieved the user satisfaction" -P32

Experts identified that active user engagement in research on product use may lead to creative and nuanced design suggestions. User-designer communication could facilitate co-design and co-creation, and workshop participants envisioned a future where users actively hack, customise and personalise products, transforming the role of the designer to include capturing these activities, analysing data and approving changes for manufacturing:

"Why don't we get the user to design the product that they need and provide, the mechanisms with which we can do that, and them monitor and provide a system that helps them engage with the products rather than impose our own product design." -P11

Support and Service Products

Data-driven processes will enable predictive maintenance, live user support and software updates. A key challenge for industry is to define the objectives of data gathering. The participants felt that data-driven design would be particularly useful for companies with a service business model where they are remotely supporting customers, servicing their products, and continuously providing software and hardware updates. If industry has a vested interest in prolonging the life of their products they would be more likely to invest in sensorizing them. Alternatively, functional data generated by a product may also identify gaps in the service. For example, when discussing a fictional home media system P17 identified that the content users search for but do not find may be an indication that the media service should commission more of that content.

By continuously collecting data, manufacturers can provide live support and tailored software updates. When discussing a high tech welding mask P24 identified that there is an opportunity to correlate the movement of a welder with the quality of their work using different imaging systems, algorithms, and remote support: "You can have an analyst and the data scientist working in parallel to the person doing the job". Information could then also be used to "develop an AI based model that would correct the welders as they weld". Similarly, a data-driven approach is particularly suited for predictive maintenance, where live monitoring can inform when an intervention needs to take place on the part of the service provider. A frequently referenced example was Rolls-Royce, who lease their jet engines to airlines with continuous servicing. Live remote monitoring enables service providers to actively support their customers, by providing live guidance, addressing questions, and relaying up-to-date knowledge about best practices.

"I hope it would give us an insight into what is actually happening in real time. If one of our machines has a problem I hope to be able to see the data and see it immediately and be able to react to it. I hope to be able to use the live data streams in order to debug. Is it overheating, has someone smacked it? These things would help me resolve issues quicker and offer better customer service as a result." -P3

If manufacturers are also invested in servicing and supporting their products once they are on the market, they may in turn benefit from designing their products in a modular way, so that parts can be replaced when needed. Participants suggested this way of designing products has the potential to be sustainable and appeal to customers who value quality and service. Conversely, "physical manufacturing", [which accounts] "for changeability" (P5), was seen as challenging to implement in practice.

Ethical Issues

Data-driven design can disempower users by reducing privacy, manipulate behaviours, and limit informed consent. Additionally, designers may be disempowered by handing over creative practice to data scientists and algorithms. Data-driven design has to deal with a series of thorny ethical issues in order to bring value to all the stakeholders involved, including customers, industry and designers. A clear issue with data gathering from smart products was the identification of individuals, which can both increase the adaptation potential of functionality and jeopardise the privacy of users. For example, a fictional home media system discussed by P17-P20, could be made more adaptive and responsive by recognising the voice of the person using it, but "obviously that is a privacy problem" (P20). Finally, even when data is anonymous, the inferences that could be made about individual's identity raise serious concerns around data security.

Although we focused the workshops around generating value for design, participants highlighted that industry would have financial incentives to utilise a data-driven approach. An industry that seeks to make better and more desirable products through gathering vast amounts of data about their customers through their products may be inherently exploitative. The

experts repeatedly stressed the importance of not selling the data onto third parties, specifically insurance companies. Generating revenue through the sale of data, may in turn shift the purpose of the data-gathering process from design to surveillance, "if it is to sell the data then you're no longer making [products] you're making sensors" (P2).

A key issue around gathering data is "informed choice" (P12) on the part of users. Consumers need to understand what data they are sharing, what inferences can be made about them, and what they can gain and lose from taking part in a data gathering process. The *acceptability* of the process would be crucial to adoption, and the experts felt this could be increased by giving users access to and ownership of their own data, "everyone has visibility and control of the data collected about them" (P16). The value (financial, relating to design, etc.) generated with their help should be made clear throughout the process, and users should benefit explicitly from allowing companies to gather data. While it is clear that there is a need for transparency, the composition of terms and conditions needs to become clearer and more user-centred:

"The big problem with transparency though ... is the huge overhead of time for people. Why should I spend five minutes to try and understand something about my kettle?" -P17

In addition to data gathering having potential disempowering effects to consumers, it may also negatively affect designers. Experts feared that the advancements of data-driven design can lead to the transformation of design roles from creative practitioners to data scientists, who maintain data and algorithms so complex that they lose understanding of their meaning and implications. "The question was if the machine learning is doing all the work I guess what you're doing is just maintaining the black box" (P2).

Behaviour manipulation of users through data-driven models was another key concern within discussions. The ability to understand the relationship between design and behaviour could empower industry to guide and manipulate consumers. For example P11 expressed that he fears "live data will be used to monitor and control user's behaviour and influence the use of the product", catering to the requirements of the producer rather than the interest of the consumer. Questions were also raised about the negative behaviours that would be observed through data gathering and how might industry deal with them. A fictional example discussed by P2 focuses on a car tire manufacturer that uses data-driven design for innovation:

"If you find out that actually people need better tyres, because there are driving over the speed limit, do you then produce tyres that encourage people to go faster? And if you don't, do you then become responsible for advising them to buy tyres that don't meet their needs?" -P2

Implementation Issues

Issues: cost, including sensorizing products, analysing and storing data; the need for flexibility with the manufacturing infrastructure; data quality and bias. The experts

felt that there is a tension between the need to sell finished and well-designed products and continuously collect data to redesign them. Questions were raised about the price of sensorizing products and analysing the data, a process which requires an *"incredibly compelling argument about a return on that investment"* (P5). The experts identified that the products need to be within a certain cost range to make the process worthwhile:

"If we say that we're software company and we're going to get this data and the next 5 years, they're going to say that you are investing in the next Uber, the next Google, but if you say I'm going to make kettles and they have got data in them, they're going to ask how much are your production costs and how much are you going to sell the kettles for?" -P2

The reality of implementing a design change within the production environment adds an extra layer of complexity, as a large amount of infrastructure would need to be in place to enable flexible manufacturing. A rapid redesign process may build *"obsolescence to the previous product"* (P11). Furthermore, the sensors themselves might add risks – P3, a business owner, recalled: *"We have had instances where sensors have actually taken the machine down. It's actually part of a ecosystem and that has its own risks"*.

Some of the experts identified that evolving legislation worldwide, for example GDPR, may complicate and restrict data gathering from smart devices. Experts were concerned by the potentially large volumes of data required to generate design insight, with the associated big data storage, processing capacity and energy demand making the process unsustainable. P22 reflected that currently industry tends to collect everything, while they may only *"use maybe two or three parameters, so you just wasted large amounts of time and money on storage that you don't need"*. This may in turn can lead to *"drowning in data [and] decision paralysis"* (P10).

Data was seen as being potentially uncontextualized, misleading, and susceptible to hacking (*"how long would it take before they pay money to hire people to play the model in order to get an advantage"* -P15). It was felt that the quality of the raw sensor data can bias machine learning models, and that data driven models in general may see the world in terms of existing behaviours and information, *"it is what you [were] like in the past, and your future is being dictated by your own past"* (P17). Finally, even technically well versed designers and ethnographers may misinterpret the meaning of the data and model outputs they are seeing.

Broader questions were also raised about the unintended consequences of data-driven product design. Would we inevitably create black boxes of data and algorithms that are unmanageable and unsustainable? Would this process have a negative environmental impact?

Utopian Vision

The process has positive potential: businesses benefit financially; creatives gain a valuable design and research resource that helps them solve real problems; consumers'

well-being is increased; and products are more sustainable and adaptive. When prompted to explore a Utopian future for data-driven design, the preferable futures as mapped onto Voros's Futures cone [3], the experts could imagine a process that would create new jobs around design, data, analysis and ethnography. It would allow designers to *"solve real problems"* (P13), address the needs of consumers, and engage them in a dialogue facilitating co-design while collecting data in ethical ways that hand over ownership of the process to the users. Businesses would benefit financially by creating desirable long lasting products, which are repairable, sustainable and contribute to the well-being of users. Everyone would have access to an appropriate segment of the data, and be able to understand and explore it in a way that brings about insight.

Although there were many reservations about the implementation and ethics of the process as discussed in the previous sections, a data-driven process can support sustainable consumer and manufacturing practices. P10 envisioned: *"You go out and you choose where it is built, you choose how it behaves and it is customisable to you"*. They envision an industry where data helps solve real world problems and address the needs of individuals, facilitating a sustainable, self-improving, human-centred, socially beneficial vision.

"This information will allow us, as designers to improve our process and make it a lot quicker to get a product to the market and make things with the consumer in mind."
-P14

An example of data serving a socially beneficial role was given by P5, who recalled how Nike were able to spot a flu epidemic through their wearables by identifying decreased levels of user activity across geographies.

"I think we are going to have a world of sand and gemstones in a way, where every little optimisation that we get from this way of thinking is like a grain of sand and overall we get a beach that we all like but occasionally we will get a few gems like the Nike example where you have these wonderful things that really leap us forward in a big way." - P3

DISCUSSION

The work presented here primarily contributes a current industry and academia perspective on the future of data driven product design, exploring what happens when live data collected from smart products in the wild is visualised and analysed to support sense-making and research. The experts felt that usability studies, A/B testing, and ethnographic studies can be implemented remotely around these products with the aid of feedback and engagement of users. In turn this data-driven approach can enable the re-design of hardware and software, enable user-centred thinking, and allow manufacturers to service products. Meanwhile, issues around ethics, privacy, security, and costs remain open. We propose several implications for data-driven product design. We aim to remain critical of the value and implications of the process by highlighting practices that decentralize power dynamics and address the ethical and implementation issues discussed within the workshops.

Expanding Notions of Data-Driven Design

The term 'data-driven design' tends to initiate discussions around the use of data to optimise, and refine products in relation to easily quantified performance metrics. However, as we moved through the workshop activities we found that participants were often more interested in using data to understand less easily quantified phenomena. Capturing aspects of the way products behave and fit into the world around them is seen as a means to consider new capacities to innovate and to support users. Whilst there are already good examples of data-driven design optimisation systems [36], HCI research that actively challenges the term 'data-driven design' and explores alternative human-data interactions to support early-stage design may prove valuable in this space. Against a tendency to rely on more data and more computation to solve problems, we would question how much data is needed to drive the process, rather than being seen as a supporting material. To follow Shea et al. from 2005, and equally relevant today: "the real challenge is to make systems that designers want to use" [43] – and in this case it means understanding where data is beneficial and supportive to creativity, rather than an external *driver* dictating solutions.

The experts saw value in data collection for the purposes of design research, which can help answer questions, challenge designers' assumptions about how their products are used, and enable a user-centred process. While data gathering from products in the wild was seen to have a very strong potential to be used for optimisation and predictive maintenance – reflected by current industry practices [5, 21] – the experts were still able to see some opportunities for data-gathering to be part of an innovative and creative design practice. Here discussions gravitated towards ethnographic research, focusing on the subtleties and characteristics of behaviours observed within the incoming data.

We propose that a desirable data-driven design tool would enable designers to actively investigate research questions which relate to the experiences of product use in the wild to inform design. This can be done in two ways. Firstly, by sensorizing prototypes, such as in the work of [9], where insights from thick and big data enable the development of user-centred products. Secondly, by utilising data from connected products post-manufacture, to inform the early-stage design of new species or versions of that product.

A future oriented vision for idealised data-driven design practice may lie with Sterling's [50] Spimes. Spimes are manufactured objects: "*material instantiations of an immaterial system*", sustainable, upgradeable, and made from materials that are continuously recycled back into their own production. They would be designed so that users have access to both their physical and digital instantiations. Stead et al. [49] argue that spimes are inherently more "*cyclical, ongoing and sustainable*" than IoT products, which are currently designed with a limited lifespan, made obsolete by the latest hardware or software. Those visions match well to the expert's Utopian discussions around sustainability and empowerment. One can imagine different user interfaces for spimes that enable

groups to examine the data history of the object, linking to the provenance of its data-driven design process.

Developing Data-Driven Practices

It is clear that data driven practices need to be carefully cultivated to avoid several pitfalls. The experts felt that a data-driven product design model might be too costly in terms of data storage, analysis, and sensorization beyond that necessary for the device's function. This process was plausible to them in cases where data gathering is already a vital part of the function of the device, or within a service-driven model, which requires robust and long-lasting design. There was a suggestion that data gathering outwith the device's desired functionality should only happen if there are specific questions that the design team is interested in investigating. Rather than sensorizing all manufactured devices, a subset could be instrumented as probes or transition products[25], as well as early iterations of product testing and prototyping [9].

Implementing data-driven design in practice requires techniques that bridge the "*gap between the designers, technologists and the users*" (P28), in a way that does not unduly inconvenience any of them. Users are asked to contribute their effort and give informed consent to share details about their lives and usage; designers are asked to develop literacy around the affordances of sensors, the limitations and intricacies of data capture as well as developing robust strategies for interpretation. All of this requires a temporal and cognitive investment that must be well balanced.

A growing body of work into smart product data visualisations, highlights that data has potential to be used as a resource for design research [14]. For example, digital twins have the potential to visualise data associated with a design in a spatially bound manner [14, 51], which is particularly suited for the visually oriented practices of many designers. Meanwhile sense-making around visualised data can be supported by annotation [40, 59] and collaborative data exploration [31].

How designers should make sense of data was a recurring theme in the workshops. Data was not seen as an end in itself, and designers were expected to question the accuracy of information they are seeing, being able to account for the uncertainty of sensor readings, the bias of collection and analysis algorithms and the limitations of visualisation and interpretation. Correll [17] highlights that by visualising the end product of data analysis alone "*but not the process by which it was created, we risk propagating false, misleading, or unreplicable findings*". Conversely, data visualisation with too much extraneous detail increases their complexity and weakens their rhetorical impact. Finding methods of visualisation and analysis of data to support sense-making and work within design workflows is essential. As suggested by the experts, data-driven design tools should show designers raw data alongside amalgamations and inferences, support collaborative data exploration, and give designers the opportunity to gain user feedback and reflections.

Supporting Ethical Design Practices

Privacy, disempowerment of users and designers, behaviour manipulation, and issues with informed consent, were key

ethical issues raised by the experts within the workshops. Wor- ryingly, research into pervasive surveillance more generally indicates that people become accustomed to data collection even when initially opposed to it [38]. Coulton and Lindley [18] argue industry needs to be explicit in the ways that they create value from data gathering, rather than justifying ethically am- biguous practices through the promise of user-centred design. In line with discussions raised by Shoshana Zuboff [61], a key challenge for data-driven design would be to give users control of their data. To do this, the way value is generated for all stakeholders - including manufacturers, designers, and users - needs to be made explicit. A databox approach, where data is stored and analysed on a local physical device with the user, may address concerns around privacy [20]. Meanwhile limiting data gathering to a subset of consenting users, can counter the narrative of mass surveillance.

The active and continuous engagement of users through simple and minimal feedback requests at key times, may have two beneficial implications on the ethics of the process. Firstly, that users would be continuously reminded and informed that their data is being collected and inferences about their be- haviour being made. Secondly, that users would be able to influence the researcher's understanding of the data being col- lected, enabling proactive collaborative sense-making. Within home sensing, data logs have been shown to be a valuable tool for collaborative sense-making and creating actionable knowledge between specialists and homeowners [26]. We feel that a similar collaborative process of contextualising data can support generating design insight. The user's engagement has the potential to enable co-design when the process is transpar- ent [56]. We see potential for smart product use data to inform when these interactions take place [54], but questions remain about how to appropriately enable co-design.

Limitations and Future Work

Due to the use of a series of workshops, the work presents speculations unconstrained by the complexities of how man- ufacturing and design operate. Further explorations of data- driven design practices would need to investigate best prac- tices, ethics, and the feasibility of such a process. Informed consent and active engagement with users are important but difficult challenges for data-driven design. Future work could explore how to explain data gathering appropriately, establish clear ethical and consent practices, and to engage users with the design process. Having clear and meaningful purposes for data gathering may assist with this, such as improving the accessibility or environmental sustainability of a design. Even so, there may be a limit to the proportion of the population that would agree to such data-gathering. A strict opt-in policy may inadvertently exclude individuals on the fringes of user groups who may benefit from improved design and be of in- terest to designers. This questions the universal application of data-driven design if data is gathered from a selective sec- tion of the population who are less worried or better informed about data privacy and protection. Finally, the work explores an expert perspective on data-driven product design, as such it does not account for the views of consumers. The experts raised concerns about the ethics of data gathering through

smart devices. There is an opportunity to further explore how the public might experience this process.

Implications for Design

Taken together, these results suggest implications for the future of using live product data in the design process:

- Usage data is currently under-explored as a means to gener- ate design insight. While many of the design experts were positive about the possibilities, few were actively working towards them, and there was an assumption that data was a means towards optimisation and reduction of human agency.
- Practices should not just focus on quantitative sensor data, where uncertainties and analysis methods may bias a de- signer's understanding of product use, but include qualita- tive data, such as user feedback.
- Data driven practices need to be developed that support human creativity, that contextualise data and its interpreta- tion while remaining legible and fit design workflows. This means challenging the notion of 'data-driven' and finding more nuanced ways to bring data into the design process.
- Designers were all sensitive to the ethical challenges of increased emphasis on data within the design process, both with regards to their own practices and their relations with end users. For usage data collection to be acceptable at scale, it needs to be clearly differentiated from surveillance, with negotiated consent, and with reasonable burden to users.
- Not every product should be collecting data; for reasons of cost, privacy and effort, a judiciously chosen subset of production such as early prototypes or a small percentage of release products may be more appropriate than the scat- tergun application of sensors.
- In line with this, models of value that support the new rela- tionships between designers, manufacturers and end-users need to be articulated, to ensure appropriate deployments where increased production costs, incisiveness and cogni- tive effort for users and designers are balanced by insights.

CONCLUSION

Through three workshops with industry and academia experts, we explored the opportunities and challenges for the future of data-driven product design. The outcomes from the workshops provide an overview of the HCI and HDI challenges associated with bringing live data into the product design process, and exploring areas beyond optimisation. In particular, we found opportunities for collaborative sensemaking, for broadening views on what data could be used for, and the possibilities for personalisation and customisation of physical objects. The experts tempered these utopian visions with a critical sense of the pitfalls and dangers of datafication and extractive data processes. This supports the exploration of new roles for designers and end-users supported by configurations of sensors and algorithms, generating insight in a humane and sustainable manner.

ACKNOWLEDGMENTS

This research was funded by the EPSRC under the New In- dustrial Systems theme through the Chatty Factories project (EP/R021031/1).

REFERENCES

- [1] Andy Crabtree, Alan Chamberlain, Rebecca E. Grinter, Matt Jones, Tom Rodden, and Yvonne Rogers (Eds.). 2013. Introduction to the Special Issue of ‘The Turn to The Wild’. *ACM Trans. Comput.-Hum. Interact.* 20, 3, Article 13 (July 2013), 4 pages. DOI : <http://dx.doi.org/10.1145/2491500.2491501>
- [2] Luke Abrahams. 2019. The ‘world’s safest’ bike helmet has its own airbag. Metro. (14 September 2019). Retrieved September 16, 2019 from <https://metro.co.uk/2019/09/14/the-worlds-safest-bike-helmet-has-its-own-airbag-10742254/>.
- [3] Roy Amara. 1974. The futures field: Functions, forms, and critical issues.. In *Futures*. 289–301. [https://doi.org/10.1016/0016-3287\(74\)90072-X](https://doi.org/10.1016/0016-3287(74)90072-X)
- [4] Ken Anderson, Dawn Nafus, and Ryan Rattenbury. 2009. Numbers have qualities too: experiences with ethno-mining. In *In Ethnographic Praxis in Industry Conference Proceedings (EPIC 2009)*. 123–140. <https://doi.org/10.1111/j.1559-8918.2009.tb00133.x>
- [5] Jim Austin, Tom Jackson, Martyn Fletcher, Mark Jessop, Peter Cowley, and Peter Lobner. 2004. Chapter 5 - Predictive Maintenance: Distributed Aircraft engine diagnostics. In *The Grid 2 (Second Edition)* (second edition ed.), Ian Foster and Carl Kesselman (Eds.). Morgan Kaufmann, Burlington, 69 – 79. DOI : <http://dx.doi.org/https://doi.org/10.1016/B978-155860933-4/50009-2>
- [6] Christian Bauckhage, Rafet Sifa, Anders Drachen, Christian Thurau, and Fabian Hadiji. 2014. Beyond heatmaps: Spatio-temporal clustering using behavior-based partitioning of game levels. In *2014 IEEE Conference on Computational Intelligence and Games*. 1–8. DOI : <http://dx.doi.org/10.1109/CIG.2014.6932865>
- [7] Andrew Begel and Thomas Zimmermann. 2014. Analyze This! 145 Questions for Data Scientists in Software Engineering. In *Proceedings of the 36th International Conference on Software Engineering (ICSE 2014)*. ACM, New York, NY, USA, 12–23. DOI : <http://dx.doi.org/10.1145/2568225.2568233>
- [8] Yash Bhavnani, Kerry Rodden, Laura Cuzzo Guarnotta, Margaret T. Lynn, Sara Chizari, and Laura Granka. 2017. Understanding Mobile Phone Activities via Retrospective Review of Visualizations of Usage Data. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17)*. ACM, New York, NY, USA, Article 58, 10 pages. DOI : <http://dx.doi.org/10.1145/3098279.3119841>
- [9] Sander Bogers, Joep Frens, Janne van Kollenburg, Eva Deckers, and Caroline Hummels. 2016. Connected Baby Bottle: A Design Case Study Towards a Framework for Data-Enabled Design. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. ACM, New York, NY, USA, 301–311. DOI : <http://dx.doi.org/10.1145/2901790.2901855>
- [10] Jacky Bourgeois, Gerd Kortuem, and Fahim Kawsar. 2018. Trusted and GDPR-compliant Research with the Internet of Things. In *Proceedings of the 8th International Conference on the Internet of Things (IOT '18)*. ACM, New York, NY, USA, Article 13, 8 pages. DOI : <http://dx.doi.org/10.1145/3277593.3277604>
- [11] Jacky Bourgeois, Janet van der Linden, Gerd Kortuem, Blaine A. Price, and Christopher Rimmer. 2014. Conversations with My Washing Machine: An In-the-wild Study of Demand Shifting with Self-generated Energy. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14)*. ACM, New York, NY, USA, 459–470. DOI : <http://dx.doi.org/10.1145/2632048.2632106>
- [12] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. DOI : <http://dx.doi.org/10.1191/1478088706qp0630a>
- [13] Barry A. T. Brown, Abigail J. Sellen, and Kenton P. O’Hara. 2000. A Diary Study of Information Capture in Working Life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '00)*. ACM, New York, NY, USA, 438–445. DOI : <http://dx.doi.org/10.1145/332040.332472>
- [14] Dan Burnett, James Thorp, Daniel Richards, Katerina Gorkovenko, and Dave Murray-Rust. 2019. Digital Twins As a Resource for Design Research. In *Proceedings of the 8th ACM International Symposium on Pervasive Displays (PerDis '19)*. ACM, New York, NY, USA, Article 37, 2 pages. DOI : <http://dx.doi.org/10.1145/3321335.3329685>
- [15] Alan Chamberlain, Andy Crabtree, Tom Rodden, Matt Jones, and Yvonne Rogers. 2012. Research in the Wild: Understanding ‘in the Wild’ Approaches to Design and Development. In *Proceedings of the Designing Interactive Systems Conference (DIS '12)*. ACM, New York, NY, USA, 795–796. DOI : <http://dx.doi.org/10.1145/2317956.2318078>
- [16] Wen-Wei Chang, Elisa Giaccardi, Lin-Lin Chen, and Rung-Huei Liang. 2017. “Interview with Things”: A First-thing Perspective to Understand the Scooter’s Everyday Socio-material Network in Taiwan. In *Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17)*. ACM, New York, NY, USA, 1001–1012. DOI : <http://dx.doi.org/10.1145/3064663.3064717>
- [17] Michael Correll. 2019. Ethical Dimensions of Visualization Research. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 188, 13 pages. DOI : <http://dx.doi.org/10.1145/3290605.3300418>

- [18] Paul Coulton and Joseph Galen Lindley. 2019. More-Than Human Centred Design: Considering Other Things. *The Design Journal* 22, 4 (2019), 463–481. DOI : <http://dx.doi.org/10.1080/14606925.2019.1614320>
- [19] British Design Council. 2019. Design methods for developing services. (2019). Retrieved August 30, 2019 from <http://tiny.cc/5h42cz>.
- [20] Andy Crabtree, Tom Lodge, James Colley, Chris Greenhalgh, Kevin Glover, Hamed Haddadi, Yousef Amar, Richard Mortier, Qi Li, John Moore, Liang Wang, Poonam Yadav, Jianxin Zhao, Anthony Brown, Lachlan Urquhart, and Derek McAuley. 2018. Building accountability into the Internet of Things: the IoT Databox model. *Journal of Reliable Intelligent Environments* 4, 1 (01 Apr 2018), 39–55. DOI : <http://dx.doi.org/10.1007/s40860-018-0054-5>
- [21] Jim Daily and Jeff Peterson. 2017. *Predictive Maintenance: How Big Data Analysis Can Improve Maintenance*. Springer International Publishing, Cham, 267–278. DOI : http://dx.doi.org/10.1007/978-3-319-46155-7_18
- [22] Anders Drachen and Matthias Schubert. 2013. Spatial game analytics and visualization. In *2013 IEEE Conference on Computational Intelligence in Games (CIG)*. 1–8. DOI : <http://dx.doi.org/10.1109/CIG.2013.6633629>
- [23] Anders Drachen, Magy Seif El-Nasr, and Alessandro Canossa. 2013. *Game Analytics – The Basics*. Springer London, London, 13–40. DOI : http://dx.doi.org/10.1007/978-1-4471-4769-5_2
- [24] Bob Evans. 2016. Paco—Applying Computational Methods to Scale Qualitative Methods. In *In Ethnographic Praxis in Industry Conference Proceedings (EPIC 2016)*. 348–368.
- [25] Eleonora Fiore, Paolo Tamborrini, and Silvia Barbero. 2017. Design for Next Connected Appliances. *The Design Journal* 20, sup1 (2017), S2634–S2644. DOI : <http://dx.doi.org/10.1080/14606925.2017.1352775>
- [26] Joel E. Fischer, Andy Crabtree, James A. Colley, Tom Rodden, and Enrico Costanza. 2017. Data Work: How Energy Advisors and Clients Make IoT Data Accountable. *Computer Supported Cooperative Work (CSCW)* 26, 4 (01 Dec 2017), 597–626. DOI : <http://dx.doi.org/10.1007/s10606-017-9293-x>
- [27] Tristan Gerrish, Kirti Ruikar, Malcolm Cook, Mark Johnson, Mark Phillip, and Christine Lowry. 2017. BIM application to building energy performance visualisation and management: Challenges and potential. *Energy and Buildings* 144 (2017), 218 – 228. DOI : <http://dx.doi.org/https://doi.org/10.1016/j.enbuild.2017.03.032>
- [28] Elisa Giaccardi, Nazli Cila, Chris Speed, and Melissa Caldwell. 2016a. Thing Ethnography: Doing Design Research with Non-Humans. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. ACM, New York, NY, USA, 377–387. DOI : <http://dx.doi.org/10.1145/2901790.2901905>
- [29] Elisa Giaccardi, Chris Speed, Nazli Cila, and Melissa L. Caldwell. 2016b. *Things As Co-ethnographers: Implications of a Thing Perspective for Design and Anthropology*.
- [30] Joseph Giacomini. 2014. What Is Human Centred Design? *The Design Journal* 17 (12 2014). DOI : <http://dx.doi.org/10.2752/175630614X14056185480186>
- [31] Katerina Gorkovenko, Lars Lischke, and Paweł W. Woźniak. 2018. Large Displays and Tablets: Data Exploration and Its Effects on Data Collection. In *Proceedings of the 10th Nordic Conference on Human-Computer Interaction (NordiCHI '18)*. ACM, New York, NY, USA, 664–675. DOI : <http://dx.doi.org/10.1145/3240167.3240192>
- [32] Galen Hunt, Kirk Glerum, Kinshuman Kinshumann, Steve Greenberg, Gabriel Aul, Vince Orgovan, Greg Nichols, David Grant, and Gretchen Loihle. 2009. Debugging in the (Very) Large: Ten Years of Implementation and Experience. In *Proceedings of the 22nd ACM Symposium on Operating Systems Principles (SOSP '09)* (proceedings of the 22nd acm symposium on operating systems principles (sosp '09) ed.). Association for Computing Machinery, Inc. <https://www.microsoft.com/en-us/research/publication/debugging-in-the-very-large-ten-years-of-implementation-and-experience/>
- [33] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B. Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, Nicolas Roussel, and Björn Eiderbäck. 2003. Technology Probes: Inspiring Design for and with Families. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 17–24. DOI : <http://dx.doi.org/10.1145/642611.642616>
- [34] David Edward Jones, Chris Snider, Lee Kent, and Ben Hicks. 2019. Early Stage Digital Twins for Early Stage Engineering Design. *Proceedings of the Design Society: International Conference on Engineering Design* 1, 1 (2019), 2557–2566. DOI : <http://dx.doi.org/10.1017/dsi.2019.262>
- [35] Mario Linares-Vasquez, Kevin Moran, and Denys Poshyvanyk. 2017. Continuous, Evolutionary and Large-Scale: A New Perspective for Automated Mobile App Testing. In *2017 IEEE International Conference on Software Maintenance and Evolution (ICSME)*. 399–410. DOI : <http://dx.doi.org/10.1109/ICSME.2017.27>
- [36] Justin Matejka, Michael Glueck, Erin Bradner, Ali Hashemi, Tovi Grossman, and George Fitzmaurice. 2018. Dream Lens: Exploration and Visualization of Large-Scale Generative Design Datasets. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 369, 12 pages. DOI : <http://dx.doi.org/10.1145/3173574.3173943>

- [37] William Odom, Ron Wakkary, Youn-kyung Lim, Audrey Desjardins, Bart Hengeveld, and Richard Banks. 2016. From Research Prototype to Research Product. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 2549–2561. DOI : <http://dx.doi.org/10.1145/2858036.2858447>
- [38] Antti Oulasvirta, Aurora Pihlajamaa, Jukka Perkiö, Debarshi Ray, Taneli Vähäkangas, Tero Hasu, Niklas Vainio, and Petri Myllymäki. 2012. Long-term Effects of Ubiquitous Surveillance in the Home. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. ACM, New York, NY, USA, 41–50. DOI : <http://dx.doi.org/10.1145/2370216.2370224>
- [39] George Ritzer and Nathan Jurgenson. 2010. Production, Consumption, Prosumption: The nature of capitalism in the age of the digital 'prosumer'. *Journal of Consumer Culture* 10, 1 (2010), 13–36. DOI : <http://dx.doi.org/10.1177/1469540509354673>
- [40] Hugo Romat, Nathalie Henry Riche, Ken Hinckley, Bongshin Lee, Caroline Appert, Emmanuel Pietriga, and Christopher Collins. 2019. ActiveInk: (Th)Inking with Data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 42, 13 pages. DOI : <http://dx.doi.org/10.1145/3290605.3300272>
- [41] Tony Salvador, Genevieve Bell, and Ken Anderson. 1999. Design Ethnography. *Design Management Journal (Former Series)* 10 (10 1999), 35 – 41. DOI : <http://dx.doi.org/10.1111/j.1948-7169.1999.tb00274.x>
- [42] Iskander Sanchez-Rola, Matteo Dell' Amico, Platon Kotzias, Davide Balzarotti, Leyla Bilge, Pierre-Antoine Vervier, and Igor Santos. 2019. Can I Opt Out Yet?: GDPR and the Global Illusion of Cookie Control. In *Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security (Asia CCS '19)*. ACM, New York, NY, USA, 340–351. DOI : <http://dx.doi.org/10.1145/3321705.3329806>
- [43] Kristina Shea, Robert Aish, and Marina Gourtovaia. 2005. Towards integrated performance-driven generative design tools. *Automation in Construction* 14 (03 2005), 253–264. DOI : <http://dx.doi.org/10.1016/j.autcon.2004.07.002>
- [44] Ruben Sipos, Dmitriy Fradkin, Fabian Moerchen, and Zhuang Wang. 2014. Log-based Predictive Maintenance. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '14)*. ACM, New York, NY, USA, 1867–1876. DOI : <http://dx.doi.org/10.1145/2623330.2623340>
- [45] Annelien Smets and Bram Lievens. 2018. Human Sensemaking in the Smart City: A Research Approach Merging Big and Thick Data. In *In Ethnographic Praxis in Industry Conference Proceedings (EPIC 2018)*.
- [46] B. Smith and G. Linden. 2017. Two Decades of Recommender Systems at Amazon.com. *IEEE Internet Computing* 21, 03 (may 2017), 12–18. DOI : <http://dx.doi.org/10.1109/MIC.2017.72>
- [47] Chris Speed and Jon Oberlander. 2016. Designing from, with and by Data: Introducing the ablative framework. In *Proceedings of the International Design Research Society Conference*.
- [48] John Stark. 2015. *Product lifecycle management*. Product lifecycle management, Vol. 1. Springer Cham, New York, NY. 1–29 pages.
- [49] Michael Stead, Paul Coulton, and Joseph Lindley. 2019. Spimes Not Things: A Design Manifesto For A Sustainable Internet Of Things. *The Design Journal* 22 (06 2019). DOI : <http://dx.doi.org/10.1080/14606925.2019.1594936>
- [50] Bruce Sterling. 2005. *Shaping Things*. MIT Press.
- [51] Fei Tao, Jiangfeng Cheng, Qinglin Qi, Meng Zhang, He Zhang, and Fangyuan Sui. 2018. Digital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced Manufacturing Technology* 94, 9 (01 Feb 2018), 3563–3576. DOI : <http://dx.doi.org/10.1007/s00170-017-0233-1>
- [52] Fei Tao, Fangyuan Sui, Ang Liu, Qinglin Qi, Meng Zhang, Boyang Song, Zirong Guo, Stephen C.-Y. Lu, and A. Y. C. Nee. 2019. Digital twin-driven product design framework. *International Journal of Production Research* 57, 12 (2019), 3935–3953. DOI : <http://dx.doi.org/10.1080/00207543.2018.1443229>
- [53] Niels van Berkel, Denzil Ferreira, and Vassilis Kostakos. 2017. The Experience Sampling Method on Mobile Devices. *ACM Comput. Surv.* 50, 6, Article 93 (Dec. 2017), 40 pages. DOI : <http://dx.doi.org/10.1145/3123988>
- [54] Niels van Berkel, Jorge Goncalves, Peter Koval, Simo Hosio, Tilman Dingler, Denzil Ferreira, and Vassilis Kostakos. 2019. Context-Informed Scheduling and Analysis: Improving Accuracy of Mobile Self-Reports. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 51, 12 pages. DOI : <http://dx.doi.org/10.1145/3290605.3300281>
- [55] Ellen van Kleef, Hans C.M. van Trijp, and Pieter Luning. 2005. Consumer research in the early stages of new product development: a critical review of methods and techniques. *Food Quality and Preference* 16, 3 (2005), 181 – 201. DOI : <http://dx.doi.org/https://doi.org/10.1016/j.foodqual.2004.05.012>
- [56] John Vines, Rachel Clarke, Peter Wright, John McCarthy, and Patrick Olivier. 2013. Configuring Participation: On How We Involve People in Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 429–438. DOI : <http://dx.doi.org/10.1145/2470654.2470716>

- [57] Kathryn Wheeler and Miriam Glucksman. 2015. It's Kind of Saving them a Job isn't it? The Consumption Work of Household Recycling. *The Sociological Review* 63, 3 (2015), 551–569. DOI: <http://dx.doi.org/10.1111/1467-954X.12199>
- [58] Shaozeng Zhang, Bo Zhao, and Jennifer Ventrella. 2018. Towards an Archaeological-Ethnographic Approach to Big Data: Rethinking Data Veracity. *Ethnographic Praxis in Industry Conference Proceedings* 2018 (10 2018), 62–85. DOI: <http://dx.doi.org/10.1111/1559-8918.2018.01197>
- [59] Qiyu Zhi, Suwen Lin, Poorna Talkad Sukumar, and Ronald Metoyer. 2019. GameViews: Understanding and Supporting Data-driven Sports Storytelling. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 269, 13 pages. DOI: <http://dx.doi.org/10.1145/3290605.3300499>
- [60] Shoshana Zuboff. 2015. Big Other: Surveillance Capitalism and the Prospects of an Information Civilization. *Journal of Information Technology* 30 (2015), 75–89. DOI: <http://dx.doi.org/10.1057/jit.2015.5>
- [61] Shoshana Zuboff. 2018. *The Age of Surveillance Capitalism : the Fight for the Future at the New Frontier of Power*. Profile Books.