Computers in Industry

Challenges for Data-Driven Design in Early Physical Product Design: A Scientific and Industrial perspectives

--Manuscript Draft--

Manuscript Number:	COMIND-D-21-01488
Article Type:	VSI: DataScienceForEngDesign
Keywords:	Data-Driven Design, Data Science, Design Research, Internet of Things
Corresponding Author:	Tristan BRIARD FRANCE
First Author:	Tristan BRIARD
Order of Authors:	Tristan BRIARD
	Camille Jean
	Améziane Aoussat
	Philippe Véron
Abstract:	With the rapid development of digital technologies, complex products are becoming more connected. Alongside, the usage of data in the product development process keeps on increasing. Data is an essential means of monitoring the behaviour of the products and their users for optimisation purposes. It is in the digital sector that data is already commonly used to identify new opportunities, to support decision-making and to reduce development time. To replicate this process with physical products, novel design approaches based on data are emerging. Designers need to anticipate from the early stages of product design the right data to capture and analyse. However, research is still in its infancy and faces numerous challenges. Thus, the question addressed in this article is: “what are the challenges of data-driven design research in the early phases of the physical product development process?” A workshop involving 10 researchers was held to answer this question. In addition, a campaign of 12 interviews with connected products manufacturers completed this research. Through a literature review, the workshop and the campaign of interviews, this article synthesizes both a scientific and an industrial outlook on the challenges of data driven design. It offers a first glimpse of future research leads.

Highlights

- A review of data driven design practices in product design is presented
- Data driven design challenges in product design are also presented
- A workshop with scholars exploring further challenges is conducted
- Interviews with industrials are held to complete the workshop challenges
- A synthesis of all the key data driven design challenges is presented

Title Page (with Author Details)

Title: Challenges for Data-Driven Design in Early Physical Product Design: A Scientific and Industrial perspectives

Authors: *Tristan Briard*[a,b], Camille Jean[a], Améziane Aoussat[a], Philippe Véron[b]

Affiliations:
[a] Laboratoire de Conception Produits et Innovation, Arts et Métiers – Campus de Paris, 151 Boulevard de l'Hôpital, 75013 Paris, France

[b] Laboratoire d'Ingénierie des Systèmes Physiques et Numériques, Arts et Métiers - Campus d'Aix-en-Provence, 2 cours des Arts et Métiers, 13617 Aix-en-Provence, France

Corresponding author: Tristan Briard tristan.briard@ensam.eu

1. INTRODUCTION

Nowadays, with the rapid development of digital technologies, complex solutions are becoming more accessible. This trend has enabled the evolution of products by making them both smarter and more connected. Alongside, the collection and usage of data naturally followed this development. The Design Society report by Isaksson and Eckert (2020) forecasts the future of products and development processes in 2040. The products of the future will involve even greater connectivity between the user, the product and the manufacturer. Data is an essential mean of monitoring the behaviour of existing products, processes and users for optimization purpose. Its use is also predicted to keep on increasing in the future product development process. In this context, data-driven design approaches are emerging. Data-driven design is often associated with the integration of data during the design process. However, there is a lack of unified definition (Zheng et al., 2020). Despite being acknowledged as an important approach of design for the upcoming years (Kim et al., 2017), data-driven design is limited and far from being mature yet in product design. A review of the existing literature by Bertoni (2020) helped grasp the current data integration in the product development process. Most of the research in the early stages concern the phase “identifying customers’ needs” and rely on data mining. Companies don’t have the tools and resources to effectively integrate data in their product development phases. They tend to emphasise on low effort and cost solutions, for example text mining on online reviews. Research in design science should showcase the potential of a structured and dedicated data-driven design approach. However, to this day, only a handful of research tries to integrate data early in the product development stages. In the face of this observation, the research question addressed in this paper is: “what are the challenges of data-driven design research in the early phases of the physical product development process?” The objective of this paper is to highlight research leads for future works. To answer this question, a dual approach from both scientific and industrial perspectives was conducted. In section 2, the frame of our research work is presented. Section 3 presents a literature review of data integration in the early product development stages and the currently related challenges. Section 4 present the scientific aspect of our approach. A workshop among scholars was indeed organised to answer the research question. The industrial aspect of our research is presented in the section 5. The industrial perspectives on data-driven design challenges were explored through an interview campaign among connected products manufacturers. Finally, section 6 synthesises and discusses the identified challenges of data integration in the early stages of physical product design.

2. APPROACH

2.1 Research Context

This article focuses on the integration of data during the early design process of physical products. For the sake of simplicity, we chose the design thinking approach (Rowe, 1987; Brown, 2008) as a reference for this article. This is a simple and common way of representing the product design process often illustrated by a double diamond (Design Council UK, 2005). As the scope of the research is "the early stages of physical product design", the product is considered up to before production. Thus, only the first four phases "Research", "Definition", "Ideation" and "Prototyping & Test" are considered (Faste et al., 1993). Other design approaches such as the systematic approach to product design (Pahl and Beitz, 1996) could have been chosen as well. In this case, the phases corresponding to the upstream design would be the phases "Task Clarification", "Conceptual Design" and "Embodiment Design".

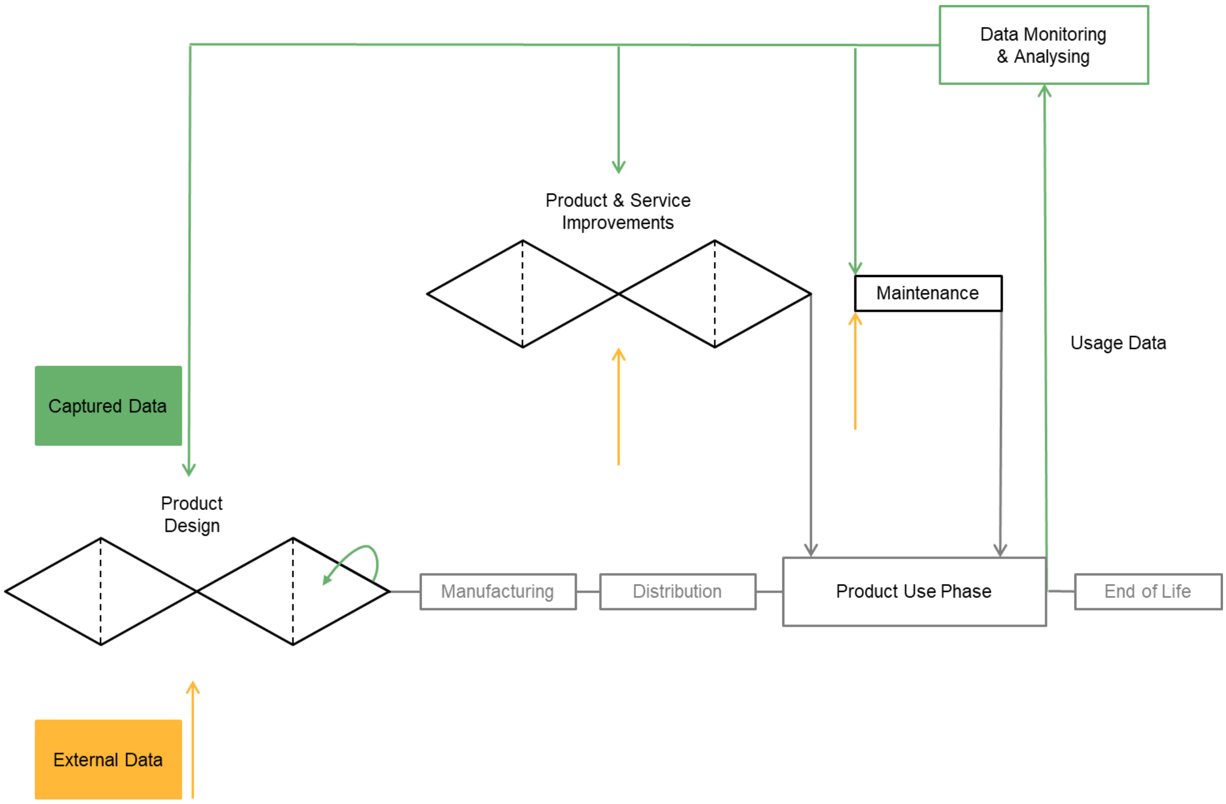

Figure 1. Research Context

Figure 1 illustrate our context of research. We have distinguished two types of data: captured data and external data. By captured data, we mean data monitored by embedded sensors and transmitted through the connectivity of physical products or prototypes for further use. For example, this could be data captured by accelerometers, gyroscopes, etc. or captured by the product interface. In contrast, external data is considered as data that does not originate from the product itself but rather from external sources. Examples include data from social media, online reviews, patents, historical data or even data monitored through external sensors. Both captured and external types of data can inform different phases of the product lifecycle management (Li et al., 2015). During the maintenance phase of physical products, the data is often exploited to provide information on the functional status in real time through sensors. The data can also be analysed and used to anticipate failures, this is predictive maintenance (Dalzochio et al., 2020). The use of data in maintenance is a developed and well-documented area of research. Although the emerging opportunity to access product data in a real-world context offers the potential to improve products during the use phase (Chowdhery et al., 2020). However, research into the use of data for product/service improvements is less common. Indeed, today only few physical products have the necessary infrastructure and scalability to ensure sufficient modifications. For example, in the automotive industry, a few connected cars seem to be the pioneers of this type of product. They are more intelligent and more easily modulated than conventional vehicles. Via remote software updates, the car's performances are improved, services are enhanced and new features can even be implemented (Lyyra and Koskinen, 2016). The integration of data during the product use phase or during the maintenance phase is not addressed in this article. Indeed, in the product lifecycle, our research scope is solely the early design phase where data can be just as valuable. Designers can use data as a decision-making tool to better meet the user's needs (Porter and Heppelmann, 2014). However, the use of captured data especially via embedded sensors to access real-world usage is not widespread (Deng et al., 2019). Indeed, it appears that the use of external data is more favoured as they are easier and cheaper to access (Bertoni, 2020).

2.2 Research Approach

In this context, to answer the research question “what are the challenges of data-driven design research in the early phases of the physical product development process?”, we conducted the following research approach (Figure 2).

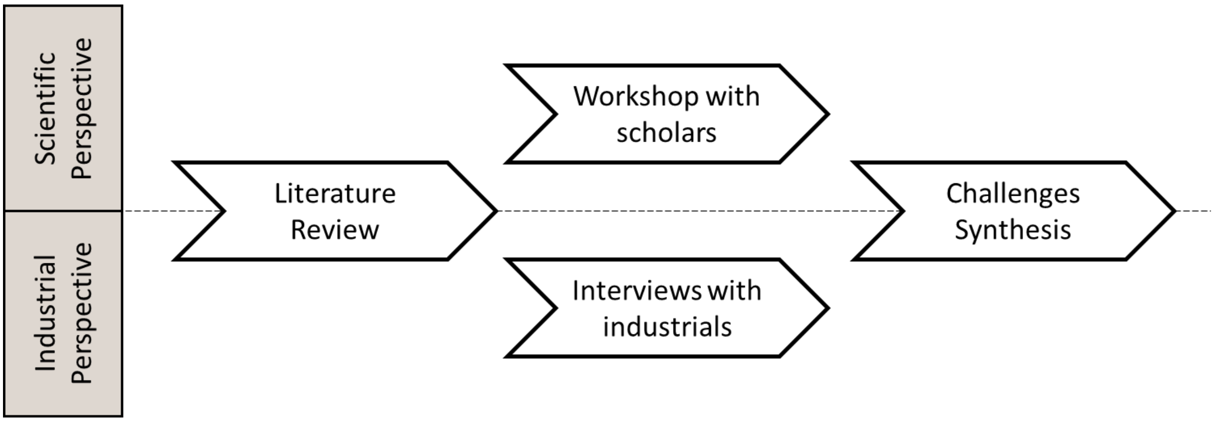

Figure 2. Research Approach

Firstly, we established a literature review covering the integration of captured and external data in physical product design. This allowed us to grasp the current practices with both types of data in each of the four stages of the design thinking process. In addition, a partial answer to the research question was established through a review of the challenges found in the scientific literature. Due to the novelty of the subject, it seemed more appropriate to carry out an exploratory study with a qualitative approach to the challenges (Gil et al., 2008). Thus, a workshop with scholars and an industrial interview campaign were conducted independently in order to access both the scientific and industrial perspectives of the challenges. The literature review was used as a basis to build the workshop framework and the interview guide. On the one hand, the workshop was organised with scholars at an international scientific conference. The participants were guided to share their views and reflect on the challenges of data in design. On the other hand, the interview campaign was organised with manufacturers of connected products who answered a semi-structured questionnaire. They were invited to share their experiences and insights on the challenges of data-driven design. Both the workshop and the interviews were designed to obtain extensive answers to the research question. Given the novelty of the field and the complexity of the subject, it is necessary to identify the priority challenges to be addressed. Thus, the last phase of our research process is a synthesis of the issues identified in the previous three phases. A reflection on the challenges, based on the literature review, the workshop and the interviews, is carried out to prioritise the challenges and to suggest leads to overcome them.

3. LITERATURE REVIEW

The following subsections present the literature review. We first address the use of data in early product design. Key references are presented for each of the first four stages of the design thinking process for both external and captured data. Next, the challenges of data integration in the design process are explored, from a design science perspective and from a data science perspective.

3.1 Data integration practices in early product design

3.1.1 Research Phase

The “Research” phase is an empathic study stage to understand the users, their different uses and their expectations of the product.

a) External Data

During the research phase, data is mainly used to identify user needs by means of external data mining (Bertoni 2020). Empathic user research is conducted through data exploration and analysis. For example, Bae and Kim (2011) used data mining on questionnaires to identify the functional characteristics that influence the purchase of digital cameras. Likewise, Lin et al. (2016) and Chien et al. (2016) have also proposed data mining on questionnaires to link user experience with product aesthetics on wearables and notebooks respectively. Yang et al. (2019) proposed a similar but more advanced approach on a smartphone use case. User experience is extracted through data mining from a large volume of online customer reviews. For a given product, the proposed tool automatically extracts three main elements: the product functionality used, the context of use and the associated user experience. From the analysis of this data, designers can emphasis on key features and re-think product issues accordingly to the users’ needs. To assess users’ sentiments about products, Chiarello et al. (2020) filtered Twitter information using their new lexicon proposal. The information obtained were found to be more relevant than traditional sentiment analysis and could be of great value to the research phase.

b) Captured Data

Bogers et al. (2016) made a connected version of a baby bottle in order to access its uses. They distributed it to a panel of testers in real-life conditions. By correlating the data collected through sensors with the testers' feedback, they were able to access a deep understanding of the users and their usages of the product. For their research, Van Eck et al. (2016) also provided connected products to volunteer users. Through patterns recognition, the captured data are processed to identify the different uses of the product. Once translated, the sensor data revealed very interesting information about the product and its usage. Designers could then access the frequencies of each different uses. They could also identify diverted uses and even discover unexpected correlations.

3.1.2 Definition Phase

To frame the problem, the "definition" phase combines the previous investigations in order to translate them into actual user needs.

a) External Data

Chiu and Lin (2018) proposed a data mining tool that fits into both the research and definition phases. Similar to the work presented for identifying customer needs, the tool scans a large database of online customer reviews. For a given product, in this case a bicycle, it identifies users' preferred components and then automatically incorporates them into the specifications. Romelfanger and Kolich (2019) applied data mining to feedback on car seats in order to maximise their perceived comfort. Based on the analysed feedback data, they defined new ergonomic prescriptions and integrated them into the product specifications. From a more technical point of view, Chiarello et al. (2017) proposed a method to automatically extract from patent texts the advantages and disadvantages of technologies. The extracted knowledge is then organised to assist designers in the further design process. Kwon et al. (2018) proposed a tool based on Wikipedia data. For a given product, the tool automatically identifies the key features and their associated design possibilities. Thus, it automatically generates a morphological matrix for the designers. The tool is presented through the design of a drone.

b) Captured Data

The captured data can reflect the real uses of the products in the field. Thus, they prove to be valuable information in the definition phase. For example, Shin et al. (2015) and Ma et al. (2017) have proposed similar work based on the analysis of sensors data on a locomotive brake system and on a crawler crane respectively. Defective parts related to abnormal field data are identified. They can then be redesigned to improve the product. Captured data can also provide information on the actual specifications needs of the users (Klein et al. 2019). Products are then defined to be the most suitable for their purposes. Lützenberger et al. (2016) proposed a case study on the bearings of a washing machine. Real world usage data are collected by sensors and used in mathematical equations to define the most suitable bearings. Another case study by Vegte et al. (2019) focuses on the design of connected refrigerators. Real-world usage data are also collected by sensors. They are utilised in virtual simulations to improve the design of the next generation of products.

3.1.3 Ideation Phase

In the ideation phase, the designers generate creative ideas answering the identified needs.

a) External Data

During the product ideation phase, data can be integrated into tools that stimulate creativity by generating many design alternatives (Kusiak 2009). To inspire designers, Chen et al. (2019) developed two data-driven ideation tools: a semantic tool and a visual tool. The semantic tool proposes relevant word associations from data mining to suggest new and innovative ideas. As for the visual tool, it generates random visuals by synthesising two selected sets of images to inspire the designers. Ranscombe et al. (2017) proposed a digital comparison tool highlighting the differences in shapes between similar products. This tool can inspire designers with novel shapes to help them differentiate their products. Song and Luo (2017) proposed a method for exploring the patent database through data mining. Based on the search terms, it offers designers a set of relevant patents for inspiration. The authors illustrated their method with the use case of spherical rolling robots.

b) Captured Data

At the time of this article writing, the authors have not found scientific research on captured data integrated to the ideation phase of the design process.

3.1.4 Prototyping & Tests Phase

Different solutions are prototyped, tested and evaluated during this last phase in order to keep the most suitable solution.

a) External Data

During prototyping, external data collection tools can be temporarily set up. They allow designers to obtain data that is difficult or impossible to access during normal use of the product. This is for instance the case for physiological data collection tools such as heart monitors, electroencephalograms or eye movement monitoring tools. They can then be analysed and translated into empathic data. For example, Maia and Furtado (2019) proposed correlations between the measurement of feelings and specific physiological data. Thus, access to the evolution of feelings during the test phase can inform the designers to best meet the needs of product users. Although their uses are widespread in the field of human-computer interfaces, these data appear to be poorly used to assess the user experiences of physical products (Maia and Furtado, 2016). Peruzzini et al. (2019) proposed a mixed prototyping method involving empathic data. They translated the physiological data recorded during the tests of the prototypes into mental workload. This enabled them to compare different design variants of tractor cabin interiors and select the one that minimised the perceived workload.

b) Captured Data

It is also common to use embedded sensors during the prototyping phase. Indeed, sensors are often added to products solely during simulation of field testing in order to improve them iteratively (Camburn et al., 2017). Some work goes beyond this traditional use of data focused on product robustness. Ghosh et al. (2017) prototyped a shoe with embedded sensors to monitor walking-related data such as pressure, temperature, etc. They have successfully translated this digital data into perceived comfort for the user. This empathic feedback appeared to be more accurate in the representation of comfort than surveys' answers. Such feedbacks drawn from data can be used to improve the product during the prototyping phase. Van Eck et al. (2019) captured data from connected medical imaging systems in the field to determine their different uses and the associated frequencies. Thus, based on this knowledge, they created suitable scenarios for ergonomic testing. Voet et al. (2018) used a connected prototype of a handheld grinder to establish correlations between the data recorded by the embedded sensors and the different possible uses of the product. They then validated these learned correlations with the help of different testers.

Figure 3 provides a summary of this review of practices. The research works on data integration are aligned with their corresponding design thinking phase.

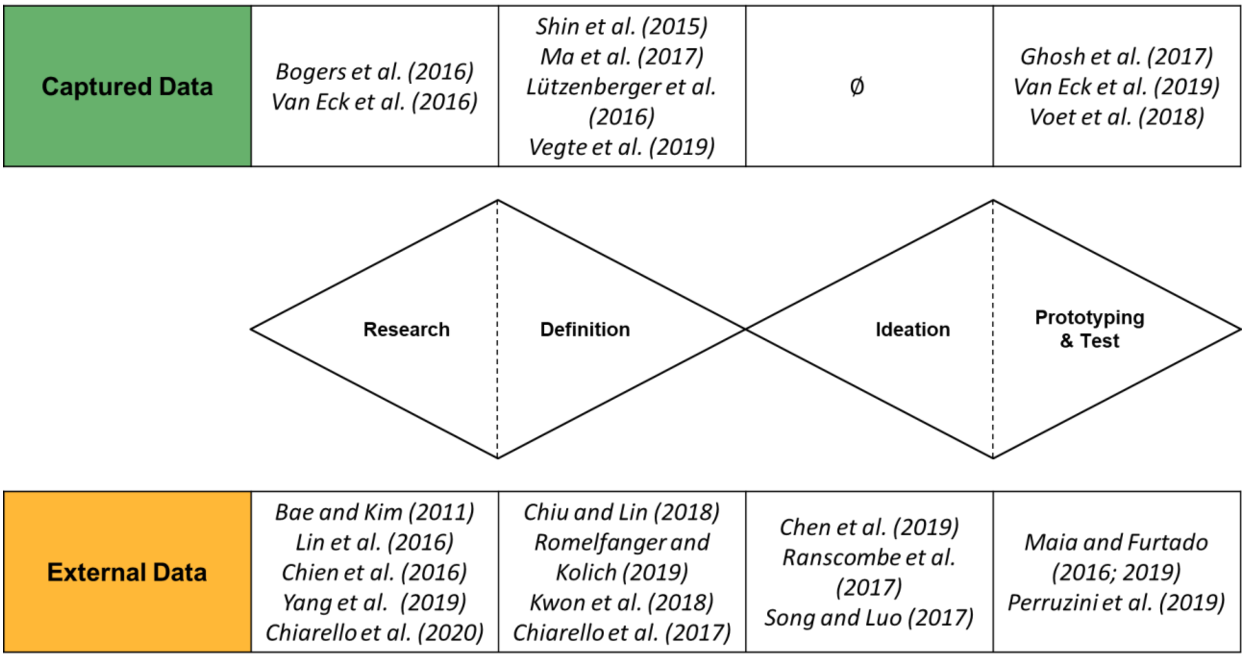

Figure 3. Synthesis of data integration practices in early physical product design

3.2 Challenges of Data Integration during Early Product Design Development

The introduction of data into product design offers new opportunities to improve its time and cost. In addition, the resulting products should also be better designed, more robust and more durable (Porter and Heppelmann, 2014). Nevertheless, with the introduction of data into product design also come challenges. They are presented in this sub-section in two separate parts. The first part presents the challenges that data integration brings to design science. The second part presents the challenges that are more specific to data science but impact design science as well i.e. the challenges at the intersection of these two science domains.

3.2.1 Data Integration in Early Physical Product Design Challenges

In their research, Labrinidis and Jagadish (2012) presented the challenges of data in the design process. Faced with an overwhelming amount of data available, the first step for designers is to define the useful data. Sorting through the data is not an obvious task, they need to anticipate in early product design which data will be relevant to capture. Moreover, in order to reduce the costs of storing large volumes of data and the costs of integrating additional sensors on products, designers of connected products are often faced with a challenging exercise. They need to minimise the integration of sensors and data to be collected but maximise the capture of meaningful information (Hou and Jiao, 2020). Furthermore, once collected, interpretations of the data must also be unbiased. Indeed, Bertoni (2018) has identified challenges regarding data interpretation issues. Product design is a multidisciplinary process involving different stakeholders (engineers, designers, ergonomists, etc.). Thus, their professions and personal experiences may influence their interpretations of the data. Similarly, the choice of data to be collected and their perceived importance may be biased. From a more holistic point of view, a lack of support for the successful integration of data into physical product design was identified. The need to develop novel methodologies emerged as a challenge. Bertoni (2020), in a literature review of data-driven design research, and Gorkovenko et al. (2020), in data-driven design themed workshops, both highlighted this need for methodologies. These methodologies should enable the best use of data in the design process and foresee the most relevant data to be captured by the product. Cantamessa et al. (2020) also looked at design from a holistic point of view and identified the challenges brought about by the rise of digital technology. To overcome the lack of support for data integration during the design process, one of the challenges identified would be to rethink it while integrating the new data professions. Bstieler et al. (2018) organised a workshop with researchers to discuss promising questions, including data in innovation. Participants also emphasised the challenge of integrating data experts and embedded electronics experts in the product design process. Integrating their tools and expertise should help to go further in the design of connected products and to maximise their potential. Finally, as data-driven design is, to date, an emerging field. There is also a need to define ethical design practices that respect the privacy of users (Gorkovenko et al., 2020). Table 1a compiles the challenges identified with the associated references.

Challenges	References
Identifying Useful Data	*Labradinis and Jagadish (2012), Hou and Jiao (2019), Cantamessa et al. (2020), Gorkovenko et al. (2020)*
Interpretation of data	*Labradinis and Jagadish (2012), Holler et al. (2016), Bertoni (2018), Hou and Jiao (2019)*
Lack of methodologies	*Bertoni (2020), Cantamessa et al. (2020), Gorkovenko et al. (2020), Chiarello et al. (2021)*
Integration of Data Experts	*Bstieler et al. (2018), Bertoni (2020), Cantamessa et al. (2020)*
Ethical Issues	*Gorkovenko et al. (2020)*

Table 1a. Synthesis of the challenges of data integration in early physical product design process

3.2.2 Data Science in Early Physical Product Design Challenges

Data is increasingly integrated into the design process as a decision support tool. As a result, the fields of design science and data science are becoming intertwined. Chiarello et al. (2021) have conducted a comprehensive literature review of the intersection of engineering design and data science. They highlighted many challenges relevant to our research. On the one hand, they identified the need to develop data-driven tools, e.g. based on artificial intelligence or machine learning, to assist engineering design. By drawing on the strengths of data science, these tools could facilitate the problem definition phase, the conceptual design, the inventive and optimisation process. On the other hand, they underlined the need to adapt current data science methods to the engineering design process. They need to be redesigned so that they work best in this specific context. For Labrinidis and Jagadish (2012) who presented few challenges related to the use of big data, extracting useful information from data is not a straightforward process. The collected raw data must be transmitted and put into a structured form for analysis. Possible errors must then be avoided during their capture, transmission and structuring. Furthermore, conclusions drawn from data are not fool-proof. The data must be clean and trustworthy but also thoroughly analysed. The review of the scientific literature also highlighted another limiting factor in the interpretation of the data: the complex context in which physical products are used. Hou and Jiao (2020) and Holler et al. (2016) have reported this challenge. Extracting valuable information from usage data can be difficult for designers as it is often associated with complex usage contexts. Holler et al. (2016) also added that the highly individual nature of products makes the insights from the data not always transferable to design other similar products. In the research of Cantamessa et al. (2020), Hou and Jiao (2020) as well as Labridinis and Jagadish (2012), the same challenge for designers to deal with large volume of data, multiple sources and in various formats is identified. It is not easy for designers to aggregate, analyse and correlate all these data to translate them into useful product knowledge. In this sense, it is expected that in the future of product design, data skills will become essential (Isaksson and Eckert, 2020). Indeed, the increasing accessibility of data will push designers to develop new data-related skills while opening up new opportunities. For example, companies today tend to integrate data into their products for descriptive purposes. Bstieler et al. (2018) highlighted the challenge of going beyond this classical use and integrate the predictive and prescriptive usage of data in addition. They may unlock further potentials for products and drive innovation. Finally, in Holler et al. (2016) research, the costs and investments involved in data integration appeared to be a major challenge. They mentioned the difficulty for companies to assess the benefits of investing in data, especially as there are also costs associated with integrating sensors into products, data storage infrastructure and jobs dedicated to data analysis. Table 1b compiles the challenges identified with the associated references

Challenges	References
Large volume of unstructured data	*Labradinis and Jagadish (2012), Hou and Jiao (2019), Cantamessa et al. (2020)*
Error in data process	*Labradinis and Jagadish (2012), Cantamessa et al. (2020),*
Costs of data and sensors	*Holler et al. (2016), Hou and Jiao (2019)*
Complex context of use	*Holler et al. (2016), Hou and Jiao (2019)*
Developing data driven tools	*Chiarello et al. (2021)*
Further Data usage	*Bstieler et al. (2018)*

Table 1b. Synthesis of Data Science in Early Physical Product Design Challenges

In this section reviewing the literature, a look at data-driven design practices and its challenges was proposed. In the next section, the challenges are explored from a scientific perspective through a workshop between scholars.

4 SCIENTIFIC WORKSHOP

4.1 Set-up

In order to complete the literature review and access a scientific perspective, a workshop was organised at the international scientific conference DESIGN 2020 (Briard et al., 2021). The objective of this workshop was to explore the views of scholars on the challenges that data brings to the early stages of physical product design. This workshop took place online and was divided into three distinct parts (Figure 4).

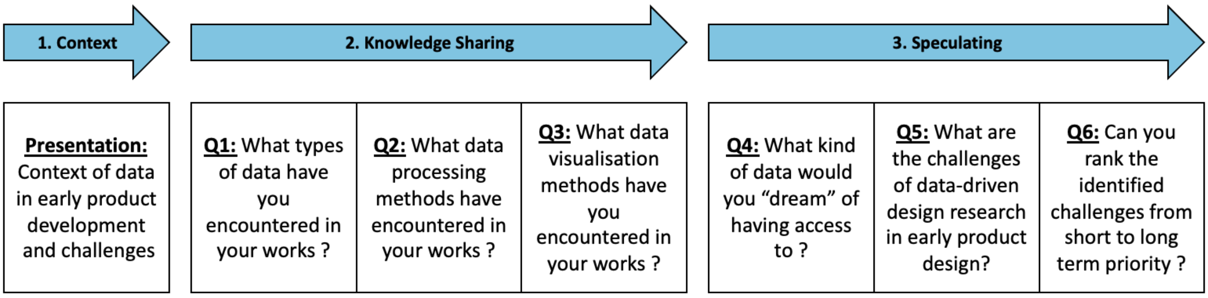

Figure 4. Workshop flow and structure

4.1.1 Structure

The objective of the first part “Context” of the workshop was to share the research context with the participants. A presentation of the literature review provided both a common background knowledge and a framework for the research. This helped participants to have a clear understanding of the research question. The presentation also emphasised on the challenges related to data in the early stages of physical product development, in order to provide insights before addressing the following parts. Then the "Knowledge Sharing" part encouraged participants to exchange their personal experiences on data in product design. This restitution part also allowed to establish a dialogue between the workshop participants before addressing the third and more creative part, the "Speculating" part. This next part aimed to generate answers to the research question and then to rank these challenges.

4.1.2 Participants and Expertise

For the workshop, participants were asked to join an online voice channel to communicate with each other. In addition, they were invited to join an online interactive whiteboard. 10 participants actively took part in the proposed workshop, including 4 PhD students, 3 university professors, 2 research associates and 1 data analyst. Each participant had an expertise in at least one of the following domains: product design, product service systems or data science. Ideally, the individual perspective of each participant (e.g. background, profession, expertise, etc.) was a source of diversity in the proposed answers.

4.1.3 Analyses

After a quick introduction of themselves, the participants addressed the series of questions detailed in Figure 4. The questions were discussed one by one from Q1 to Q5 in the same way. First, the participants individually proposed several answers to the question posed by means of virtual post-it on the interactive whiteboard. Then a discussion phase took place in which participants were invited to share, elaborate and discuss their answers with one another. Question Q6 was an open debate in which participants agreed on a ranking of data integration in early design challenges (Table 2). In the end, the workshop generated an idea board and a ranking of data challenges in the early design process. In addition, the discussion phase of each question was transcribed anonymously. The authors have analysed and discussed this generated material in order to convey as accurately as possible the relevant contributions of the participants in the next section.

4.2 Results

4.2.1 The Knowledge Sharing Phase

During this part of the workshop, participants were first asked to share about the type of data they have used in their work (Q1). A wide range of data types was mentioned according to the specialities of each participant such as “historical data”, “use phase data”, “simulation data” or “prototype data”. There is one important observation that could be already drawn. The fact that most of the data mentioned by the participants have not been thought of in advance to inform the design. They are not dedicated to a data-driven design approach. They would have been produced anyway even with other product development approaches. Indeed, designers often deal with already existing datasets or easily implemented data collections (Bertoni 2020). However, to maximise its potential, Cantamessa et al. (2020) argue that data collection and use should be considered as an essential part of the design process, rather than a secondary tool. The questions of data processing (Q2) and data visualisation (Q3) were then addressed. As with the types of data, the methods of data processing vary according to the specialities of the different participants. However, most participants used the same visualisation tools in their work like “graphs”, “heat maps” or “top lists”. Another significant observation from this knowledge sharing phase is that these tools are conventional and not specific to data science. It could be interesting to measure the benefits of using specialised tools in a data-driven design approach (Chiarello et al., 2021).

4.2.2 The Speculating Phase

A speculation phase was then conducted: participants were first asked to think about their dream data (Q4), i.e. the data they would find most interesting to have access to in a data-driven design approach. The participants' propositions like "interaction data" and "environmental data" respectively expressed a common need for a product capable of monitoring both user-product and product-environment interactions. While such a product might have seemed utopian a few years ago, the increasing accessibility and connectivity of sensors is allowing designers to get closer to it (Isaksson and Eckert 2020). However, monitoring every parameter is also counterproductive and unrealistic. Therefore, designers need to think carefully in advance about which interactions to measure and how. These concerns were underlying many of the challenges proposed during the workshop. In order to answer the research question, participants were finally asked to identify data-driven design challenges and rank them by importance (Q5 & Q6). The participants' suggestions highlighted 5 main ideas of challenges for data-driven design in early product development. Challenges ranked from high to low priorities can be found on Table 2.

The first research challenge that participants unanimously chose for high priority is the need to develop robust and dedicated data methodologies. This is a direct consequence of the lack of guidance when integrating data into the physical product design process. Indeed, due to the recent democratisation of product connectivity, research on data-driven design is still in its infancy (Yu and Zhu 2016). During the discussions, participants formulated many questions that future methods should answer, including the following: “Depending on the product type, which data should be collected?”, “What is the minimum granularity of the data to collect?” or “How can the value of data be judged?” In addition, participants added the difficulty of identifying useful use cases on which to base their own work. In addition, participants added the difficulty of identifying useful use cases as a high priority challenge. The participants hoped to gain in quality and time in their research thanks to these already experienced solutions. However, as data integration in design is still a recent field, there seem to be few case studies. Moreover, the very specific contexts of each use case make the knowledge transfer difficult. In this respect, one participant mentioned the IoT use case catalogue of Wilberg et al. (2018) as a potential source of inspiration.

Next, two different challenge ideas were ranked as medium priorities by the workshop participants. The first one concerns the means of capturing usage data. Once the data to be monitored are identified, the appropriate sensors to be integrated must be defined. This translation step is not straightforward and raised questions among the participants. During the discussions, one participant suggested the idea of a "sensor library". A solution that would offer, depending on the data needs of the designers, the associated sensors, their advantages, disadvantages and illustrative case studies. The other challenge idea raised in middle term priority is not exclusively a problem of design science. It addresses the contemporary need for protection of the users’ data, its privacy and its security. Participants appeared genuinely concerns about the capture of usage data and raised concerns about ethics. Participants were worried that already ambiguous practices on the internet would be replicated with physical products through their increasing connectivity. Discussions among participants focused on the need for a "code of conduct" to be followed when designing with data in order to "make the most of the data without infringing on the privacy of users".

The idea of the last research challenge is clearly expressed in the note "Synergies/Limits of Data Integration in Design". This challenge implies a more mature knowledge of data-driven design and is therefore ranked as a low priority. For the workshop participants it seemed interesting to explore the strengths and weaknesses of this novel design approach. It seemed interesting to explore synergies with other design methodologies.

Challenges	Priority
Lack of Methodologies	*High*
Identifying Useful Use Cases	*High*
How to collect data during product use	*Medium*
Ethical aspects (data privacy and security)	*Medium*
Synergies/Limits of Data Integration in Design	*Low*

Table 2. Ranking of the workshop challenges (Q5 & Q6)

In this section, a scientific perspective of the challenges related to data integration in product design has been highlighted. The next step of our research is to study the industrial perspective of the challenges.

5 INDUSTRIAL INTERVIEWS

5.1 Set-up

To complement the literature review and the scientific perspectives, an interview campaign with connected product companies was conducted. The intent was to explore the challenges of data-driven design from an industrial perspective.

5.1.1 Structure

Because of the novelty of the data-driven design approach for physical products, we did not find common and shared methods yet in the industry. A wide range of practices, different for each company and context, is to be expected. For these reasons, a semi-structured approach was chosen for this campaign interview. An interview guide was established with the aim of letting the interviewees express themselves freely on their practices. In addition, since this research focuses on the early part of the design process, it was important to frame the interviews on this part. The objective was to explore the various practices of designing physical products with data as well as the difficulties encountered. The interview guide was constructed with the literature review and recommendations of Schultze and Avital (2011). The guide was divided into four parts and is illustrated on figure 5. First, an introductory part in which the participants present themselves, their position, their expertise and their data related projects. Then, for the second part, each of the 4 steps of the design thinking process is addressed independently (research, definition, ideation and prototyping & test). For each step, the company's data practices are investigated through the participants' responses. In the third phase, the interview guide considers the use phase, the focus is on usage data. Thus, the participants are questioned about their utilization of the data generated by the products use. In particular, how do they anticipate and integrate sensor data into the design process? Finally, in the last phase of the interview guide, the participants are asked to answer the research question.

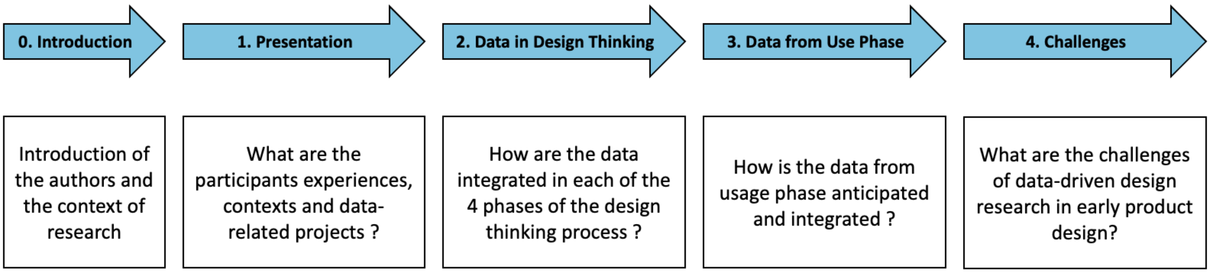

Figure 5. Interview flow and structure

5.1.2 Compagnies and Participants

To ensure the relevance of this research, it was necessary to interview companies and stakeholders with experiences in data for physical products. We therefore selected companies that offered connected consumer products. In that regard, companies manufacturing connected household appliances, sports, medical, leisure and mobility products were reached. Different profiles were approached: compagnies of different sizes (start-ups, medium-sized and multinationals) and of different maturity (from those producing connected products for several years to those just starting out). Several profiles of participants were also interviewed to ensure a variety and diversity of answers. They all had a strong expertise in data and/or sensors in design, the majority of the participants had at least 5 years of experience, but not the same professions (IoT R&D engineers, project managers, designers, etc.) All of these choices in selecting companies and participants should provide a wide and varied spectrum of responses to the research question, while ensuring the relevance and expertise of the answers given. In the end, 12 interviews were conducted with 15 different data-driven design stakeholders in 9 different companies (Table 3).

Company	**Industry**	**Approx. Revenue**	**Approx. Employee**	**Interviewee Profiles**	
				Interview 1	Interview **2**
1	Automotive	20B €	100K	IoT R&D Engineer	
2	Mobility	10M €	25	CEO	Data Engineer
3	Mobility	33M €	100	Head of R&D	
4	Leisure	200M €	1000	Test Engineer & Technical Director	
5	Leisure	30M €	50	IoT R&D Engineer & Test Engineer	
6	Sports	3B €	15K	Head of Data Project & Head of IoT Project	
7	Sports	100M €	200	IoT R&D Engineer	
8	Household	7B €	30K	Designer	Head of R&D
9	Medical	50M €	200	Head of Design	IoT R&D Engineer

Table 3. Compagnies and Interviewed Profiles

5.1.3 Analyses

All the interviews were conducted online via a voice and video sharing software. They were completed from April 2021 to June 2021 with a minimal interview length of 36 minutes, a maximal interview length of 73 minutes and an average of 51 minutes. Interviews were anonymized and transcribed in over 16 500 words. To answer the research question, a thematic analysis was applied to the interview transcripts. The data were analysed through a deductive coding approach (Boyatzis, 1998). Codes are labels used to describe content of text extract. Our initial set of codes is derived from the literature review and consists of the four following codes: “design science challenge”, “data science challenge”, “sustainability and ethics challenge” and “financial challenges”. First, an individual analysis of the interviews transcripts was carried out to assign excerpts to codes. Then, for each code, a comparative analysis of the associated data was conducted to highlight similar challenges among the different transcripts. The following section presents the shared challenges that occurred in the transcripts of at least two different companies.

5.2 Results

Even if the companies and participants did not share the same expertise and maturity. Many common challenges emerged from the interviews. This sub-section presents the challenges derived from the deductive coding approach. The Tables 4a, 4b, 4c and 4d present for each one of the four codes, the associated challenges with the number of occurrences among distinct companies.

5.2.1 Design Science Challenges

Challenges	Occurrence
Lack of illustrating use case	*5*
Lack of guidelines	*4*
Lack of methods	*4*
Lack of organisation in design teams	*2*

Table 4a. Synthesis of the Design Science challenges identified

The first challenge highlighted with the highest number of occurrences among the companies is the lack of illustrating use cases. Several companies mentioned their need for use cases during the interviews, but not with the same intentions. Most companies wanted to utilise the use cases as communication tools. The principles and potentials of data integration in physical product design would be presented to the decision makers through relevant use cases. The idea is to convince them to unlock investments in data-driven design within the company. Use cases are also interesting when presented to product design stakeholders. They can inspire them to think about the integration of data applied in their own specialties. "There is a need to "evangelise" the potential of data within companies," noted one participant. Participants considered use cases to be essential to raise awareness of data opportunities. In their view, once data-driven design is seen as valuable, a change in design practices toward data integration should follow naturally. For some companies experienced with data integration, use cases are mentioned more as sources of information to speed up the design by helping to identify relevant data to collect or data correlations. After the use cases, it was the lack of guidelines and methods that emerged as important in the interviews. All the companies interviewed agreed that a structured approach is needed to successfully integrate data and maximise its potential. Some companies admitted that they lacked experience in integrating data during design. Indeed, even though data is an inherent part of their connected products, it appeared as little collected and even less exploited. These companies do not anticipate usage in the design and they do not have a long-term vision with data, only short-term objectives. Faced with the trend towards the digitalisation of physical products, they have become aware of the value of data and have expressed a need for guidance. Not only the less experienced companies expressed a need for methodologies, indeed companies more experienced with data integration did too. Even if with their longer experiences, they have a better understanding of data integration, they still expressed a lack of methods. However, given the novelty of the subject, they wanted reliable support to guide their integration of data and speed up the design. For example, methods to find correlations, as they experiment in the laboratory by trial and error. In this regard, one participant said "to identify the relevant data to collect and the associated sensors, we have to experiment a lot. Sometimes we have good picks and sometimes not". Another challenge raised concern the design from an organisational point of view. As with the methods, the organisation of data-driven design, its actors and their roles have not yet been clearly defined in the scientific literature. Companies organise themselves internally according to their means. This lack of robust design organisation makes it difficult for them, as "between the different design departments, everyone passes the buck, leaving the others responsible for tasks that are not entirely specific to their domains". Thus, time costs and design quality are impacted by an unstructured internal organisation.

5.2.2 Data Science Challenges

Challenges	Occurrence
Accessing context of use with data	*5*
Accessing user information	*3*
Physical limitation of the product	*3*
Transferability of correlations	*2*

Table 4b. Synthesis of the Data Science challenges identified

For the field of data science, the challenges identified mainly concern the trio of the product, its user and its environment. As they interact together, they create/define the context of use. The challenge most often mentioned by companies was to access this context of use of products through data. Indeed, interpreting data into useful information for design is not straightforward. It requires sufficient reliable data to correlate and extract useful knowledge. Designers found it difficult to translate the usage data collected from products into design resources. The complexity of the physical product in a real environment greatly exceeds that of digital products and therefore limits conclusions based on data. Participants said, "Very often the context is missing to draw interesting conclusions from data" and "The interest of the data collected by the products seems a bit limited. The product embedded sensors lacked an additional link to the environment." Some companies, with more experience in the field of data, have also expressed the specific challenge of accessing information about the user experience. Again with the aim of obtaining valuable information to inform the design, it is about accessing the emotional states of the users while using the product. The correlation of data to translate a user experience is still a very complex research topic. "Human beings are complex and evolve in a complex world" said a participant. Another participant added "The [outside] environment is not controlled at all. The same surroundings with the same users can give very different results because the conditions and the context are always different. The data cannot be controlled from one day to the next, from one user to the next." Participants also expressed limitations in integrating sensors into physical products. This challenge is mainly related to the current technological capabilities of sensors and means of data transmission. Some designers felt that these were insufficient and therefore hindered their projects and ambitions. To be able to capture and transfer data, a product needs to have on-board electronics which require dedicated space and energy resources. Plus, there are also cost issues: "we simply implement the minimum sensor that goes well for cost reasons. For example, if an engine overheats, we put a temperature sensor on it for safety reasons to shut it down in an emergency, but we don't add memory to it to follow the temperature changes and thus try to understand the faulty conditions." Finally, even if the knowledge from the data can be properly captured, transferred and interpreted to inform product design. The uniqueness of products, their context of use and their own environment make it difficult to transfer the knowledge gained to other similar products.

5.2.3 Sustainability and Ethics challenge

Challenges	Occurrence
Justifying the add of connectivity	*4*
Privacy and data security	*3*
Ethical limits to data capture and interpretation	*2*
Environmental costs	*2*

Table 4c. Synthesis of Sustainable and Ethical challenges identified

"IoT [in physical product] represents a huge environmental, time and financial cost," said one participant. In the current context of limited resources, data and sensors must be considered in advance to avoid unnecessary energy consumption. In this respect, the first point that emerged through the interview campaign is the justification of the intention of connectivity. Designers feel that it should add something to the product, benefit the users and be a useful and sustainable solution. "Mass producing these products is not realistic. On a washing machine why not, but here [about some products] it doesn't add anything." Secondly, it appeared important to also consider the need for privacy and security. The increasingly digital part of physical products introduces the same issues as for the internet. Some interviewees were aware of these issues and wanted to anticipate them. In addition, some also expressed the need for an ethical framework to be respected when designing connected products. They felt like limits must be set on collected data and the conclusions that can be drawn in order to guarantee users' anonymity.

5.2.4 Financial Challenges

Challenges	Occurrence
Price of data and sensors	*8*
Return on investment	*4*
Lack of IoT costs modelisation	*3*

Table 4d. Synthesis of the financial challenges identified

For many designers interviewed, the price of sensors and data was identified as the main barrier to data integration in physical product design. Almost all companies mentioned this challenge during the interview campaign. Investment in sensors and data is considered too expensive. "Putting sensors in products, collecting data and storing it in a data lake is not cheap. Assessing these costs is "sobering". In addition, it is very difficult to quantify the potential gains brought about by this product connectivity." Quantifying the financial gain of data integration is the second challenge highlighted by companies. Indeed, in addition to the high cost of integrating data into physical products, there is a huge difficulty in assessing the return on investment from integrating sensors and data in design. Thus, the budget for associated services in the company is limited because there is no way to rely on the return on investment. All participants showed confidence in data integration but admitted this difficulty and the lack of "proof of its potential". However, they regret that company policy often seems to tend to stick to the minimum investment in data integration. One participant summed it up: “It is indeed perceived as interesting to recover information via sensors and data. But for the moment, the objective is to do something more concrete with a faster economic return. Our R&D budget is not infinite.” Some companies even mentioned the need for an IoT cost-modelling tool in order to assess automatically the financial costs of data integration. Although the interviewed companies have not yet coped with the price of sensors and data. Several were in the process of restructuring internally to move towards a more data-integrated design process

In this section, an industrial perspective of the challenges related to data integration in product design has been highlighted. Following this research, the next section proposes a synthesis on the challenges encountered in the literature review, the scientifical workshop and the industrial interviews.

6 CHALLENGES SYNTHESIS

Even if researchers, scholars and industrials, all recognise the potentials and promising perspectives of data integration in physical product design, many challenges were identified through our research. This section intends to summarise these scientific and industrial challenges. It highlights and prioritises the challenges to be addressed in order to foster the future development of data integration in the design process. To follow our research scope, we will focus more on the identified challenges affecting the early design of physical products. Thus, some of the challenges mentioned in the previous sections will not be considered. However, it would be of interest to consider them in further research. A summary of the key challenges considered is provided in Figure 6. We have chosen to highlight these key challenges because they appear to be the most mentioned in the literature, the most important to the workshop scholars and the most frequently reported by the interviewed industrials.

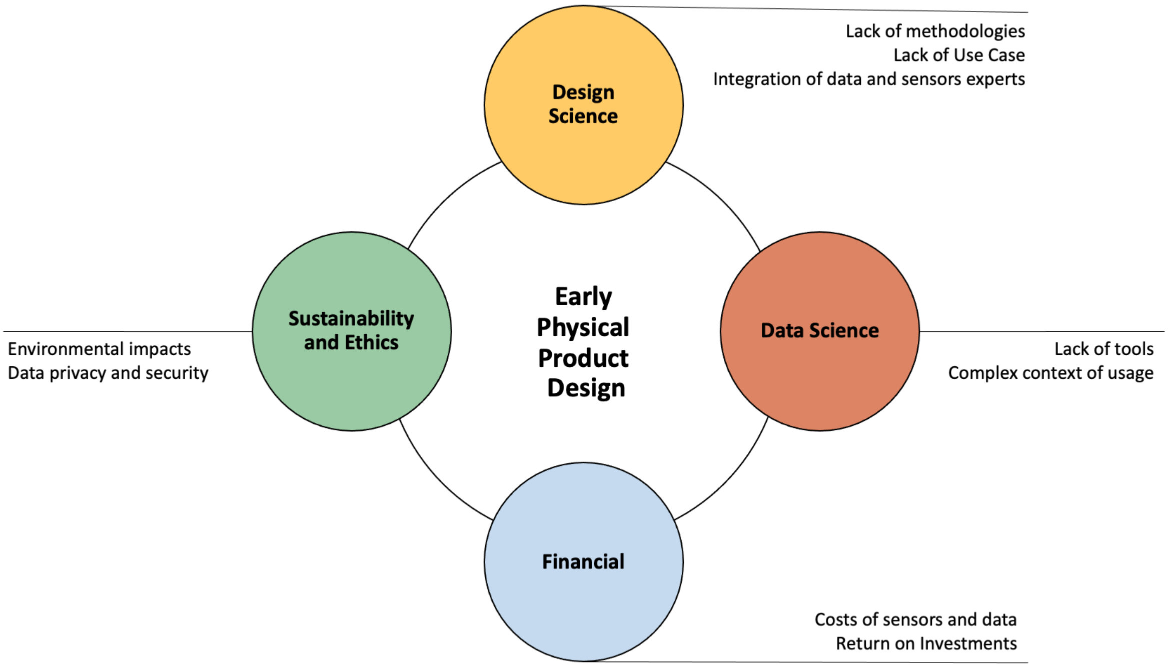

Figure 6. Synthesis of the key challenges identified in the interview campaign

The lack of a structured methodological approach was evident in our research. It was a major concern expressed simultaneously in the research literature, the scientific workshop and the industrial interview campaign. Without a structured methodological approach, designers are not taking full advantage of data potentials and are missing out on design opportunities. This lack is due to the novelty of data integration in the design of physical products. During our research, some participants even showed frustration at the waste of potential, time and energy that this lack induced. Moreover, practices are held back by a reluctance to invest in data for design, as it is difficult to clearly quantify its potential and returns on investment. Thus, Design Science is slowly approaching data integration in physical product design. To date, data is not considered at the forefront of product design, only as a secondary tool to assist designers. There is a need for a paradigm shift in physical product design. To progress toward this new paradigm, methodologies are needed to properly address data-driven design and maximise its potential. In this regard, the following subsections present the pathways towards the establishment of solid data-driven design methodologies thorough use cases, guidelines, sustainable, ethical and organisational aspects.

6.1 Use Cases

The first major challenge in the product development process toward the change in design paradigm is providing successful and robust use cases. In our research, use cases emerged as the gateway to integrating data into the design of physical products. Firstly, it presents an opportunity for testing. Indeed, from a scientific point of view, use cases allow building and testing different methodological proposals. Designers could experiment with data-driven design methods or tools and evaluate their potential. Design science would thus progress towards robust and verified methodological proposals. From an industrial point of view, case studies provide concrete illustrations of the potential of data. They provide a simple way of presenting the opportunities and limitations of data integration to uninitiated stakeholders, especially to the departments that make decisions about the research budget. In this sense, they are seen as a lever to encourage investment in research on data integration within companies.

Finally, from an industrial as well as a scientific point of view, the use cases have an important potential for inspiration. They allow to inspire the stakeholders in the design of correlations between data and knowledge. The challenge is that the uniqueness of the use cases and the data associated with them makes reliable knowledge transfer difficult. However, recent research may help to fill this gap. Wilberg et al. (2018) have compiled a database of 245 Internet of Things (IoT) product case studies. This catalogue aims to provide design analogies through relevant use cases for designers struggling with the design of connected products.

6.2 Guidelines

Faced with the many possibilities offered by data and sensors, design stakeholders could be lost when integrating data. The need for guidelines in the design of the data-enabled product was expressed several times in our research. First, there is a need for methods. These provide a framework and direction for data-driven design approach, but also establish good practice. Today, as data integration is done rather independently and on a case-by-case basis, there is no systematic approach. In our research, different methods were needed depending on the participants' experience with sensors and data. For example, novice designers were looking for structured methods to guide them in identifying the data to be captured and in integrating sensors. Others, who were a little more experienced, were looking for methods of exploring correlations, to link captured digital data to knowledge about product use. And, the more experienced participants went further, looking for correlations about the environment and the users of the product. In sum, the methods should guide stakeholders to anticipate the exploitation of data and maximise their potentials. Designers will be able to better anticipate the design of next generation products and provide effective solutions for maintenance or improvement of products during the use phase. Secondly, to facilitate the work of designers, tools would complement these future methods. At the same time, the tools would improve the quality, time and costs of the design process. As with methods, the need for tools may vary according to the areas of expertise and experience of the design stakeholders. However, the development of some generic tools could support a majority of stakeholders in product design. For example, a library of sensors, data and correlations to provide automatically designers with the relevant data to collect and the means to deploy was mentioned during our research. The first proposals for methods and tools, if effective and successful, can also contribute to the paradigm shift towards the integration of data into physical products. The exploration of methods and proposals for tools should establish a solid methodological basis for data-driven design. Thus, fill the methodological gap identified in design science research.

6.3 Ethical and sustainability

One of the major challenges raised by our research for the future of data-driven design is the establishment of ethical rules. Problems related to data collected on the internet and its uses also threaten physical connected products. They appear to be even more sensitive as they go beyond the virtual to the physical domain. It is important to anticipate these issues in order to protect the privacy of users. It is therefore necessary to be able to judge the ethical aspect of the information accessible through the interpretation of the data. Judging whether information is ethical or not is beyond the scope of this research. However, it seems essential to define a set of ethical design rules to be observed during product development to ensure future user privacy (Baldini et al., 2016). The recent general regulations on data protection on the Internet could serve as inspiration. It is also a question of securing product data so that it cannot be captured by malicious persons. Similarly, it is not the subject of this research, but it seems important to think of cyber security solutions during the early design of connected products (Porter and Heppelmann, 2015). In addition, there is also the ecological aspect. The collection and use of data are very energy-intensive, whether in the integration of the on-board electronics and its operation or in the storage of the data collected. It is necessary to be able to justify the need and interest of the research carried out before implementing the technological solutions necessary to exploit the data. Moreover, if the approach is valid, the environmental impact of each phase of the product's life cycle can be minimised through eco-design tools and methods (Rossi et al., 2016). In the same way as for questions of privacy and data security, it is by tackling these aspects from the upstream design stage through dedicated methodologies that designers can best anticipate environmental imp acts.

6.4 Organisational

The potential of data integration for physical product design is recognised on both the scientific and industrial sides. At the time of the interview campaign, several companies were in the process of making internal organisational changes that bring the data and sensor professions closer to the design profession. The integration of these new professions to support the upstream design phases is a challenge. Integrating these new specialised experts would allow their specific expertise, methods and tools to be included in the product development process. In this respect, it is important to organise an efficient workflow between the different stakeholders to ensure the best integration of these new skills. For example, in the early stages of design, these specialists can propose capture solutions that are adapted to the needs of the designers and that maximise the potential of the data. Furthermore, in the use phase, they can analyse and structure the data for designers to translate it into useful insights.

7 CONCLUSION

The objective of this paper was to explore the challenges of integrating data in the early stages of physical product design. To answer the research question: “what are the challenges of data-driven design research in the early phases of the physical product development process?”, this paper takes a dual approach combining both a scientific and industrial perspective. First, it presents a literature review for data-driven design practices in early design and its challenges. Then, based on this preliminary study, a scientific workshop and an industrial interview campaign were conducted. The different participants, stakeholders in the design process with experiences in data, were able to express themselves on the challenges of data integration. Finally, a ranking and synthesis of the challenges is proposed. The challenges are then discussed and directions for future research are outlined.

Our findings indicate that the first priority is to work towards a paradigm shift in the use of data in the early design of physical products. The reluctance to integrate data is mainly due to the cost of the technologies involved and the unquantifiable returns on investment. Relevant use cases could be explored to highlight the potential of this integration to foster research into data-driven design of physical products. Secondly, methodological gaps have been identified that should be filled by design research. Design stakeholders lack guidelines to properly integrate data and seize new opportunities. Thus, the propositions of methods and tools should help research to move forward and build a solid methodological basis for the integration of data in the design of physical products. In addition, data can be a sensitive issue with regard to the privacy and security of user data. It can also be an energy-intensive area and contribute to negative environmental impacts. These aspects need to be integrated from the beginning of the development process. For example, a set of ethical design rules for connected products could be proposed. Finally, a framework should be developed to effectively organise the integration of sensors and data professionals as new stakeholders in product design.

However, it is important to note that this study is not without limitations. Firstly, due to the novelty of this design area, this research adopted a qualitative approach. Indeed, the scholars of the workshop shared their personal reflections on the challenges of data integration. This means that other challenges could be mentioned by different participants. The same is true for the industrial practitioners of the interview campaign. Despite the effort to represent different industrial fields of connected products, their views on the challenges may also be related to their personal experiences. Furthermore, due to the exploratory nature of our research, the proposed results are not exhaustive and are subject to the authors' interpretation. This means that further workshops and interviews might reveal additional challenges, as well as conducting the same research with a different group of researchers. In this regard, a quantitative approach among academics and industry could be carried out to confirm the leads identified and refine the results of this research.

The objective of our research was to shed light on the main obstacles to the integration of data in design. To go further, future research could explore more precisely the specific needs of certain industry sectors or investigate the needs according to the data maturity of companies. Moreover, from a different point of view, future research could also explore the integration of data to assist other phases of the product lifecycle where data potential is not fully exploited such as end of life.

REFERENCES

Bae, J.K., Kim, J., 2011. Product development with data mining techniques: A case on design of digital camera. Expert Systems with Applications 38, 9274–9280. https://doi.org/10.1016/j.eswa.2011.01.030

Baldini, G., Botterman, M., Neisse, R., Tallacchini, M., 2018. Ethical Design in the Internet of Things. Sci Eng Ethics 24, 905–925. https://doi.org/10.1007/s11948-016-9754-5

Bertoni, A., 2020. Data-Driven Design in Concept Development: Systematic review and missed opportunities. Proceedings of the Design Society: DESIGN Conference. Cambridge University Press, 1, pp. 101–110. https://doi.org/10.1017/dsd.2020.4

Bertoni, A., 2018. Role and Challenges of Data-Driven Design in the Product Innovation Process. IFAC-PapersOnLine 51, 1107–1112. https://doi.org/10.1016/j.ifacol.2018.08.455

Bogers, S., Frens, J., van Kollenburg, J., Deckers, E., Hummels, C., 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. Presented at the DIS ’16: Designing Interactive Systems Conference 2016, ACM, Brisbane QLD Australia, pp. 301–311. https://doi.org/10.1145/2901790.2901855

Boyatzis, R. E., 1998. Transforming qualitative information: Thematic analysis and code development. Sage Publications, Inc.

Briard, T., Jean, C., Aoussat, A., Véron, P., Le Cardinal, J., Wartzack, S., 2021. Data-Driven Design Challenges in the Early Stages of the Product Development Process. ICED21, 2021, Gothenburg, Sweden. pp.851-860. https://doi.org/10.1017/pds.2021.85

Brown, T., 2008. Design thinking, Harvard Business Review, n° 86, p. 84-92.

Bstieler, L., Gruen, T., Akdeniz, B., Brick, D., Du, S., Guo, L., Khanlari, M., McIllroy, J., O’Hern, M., Yalcinkaya, G., 2018. Emerging Research Themes in Innovation and New Product Development: Insights from the 2017 PDMA- UNH Doctoral Consortium. J Prod Innov Manag 35, 300–307. https://doi.org/10.1111/jpim.12447

Camburn, B., Viswanathan, V., Linsey, J., Anderson, D., Jensen, D., Crawford, R., Otto, K., Wood, K., 2017. Design prototyping methods: state of the art in strategies, techniques, and guidelines. Des. Sci. 3, e13. https://doi.org/10.1017/dsj.2017.10

Cantamessa, M., Montagna, F., Altavilla, S., Casagrande-Seretti, A., 2020. Data-driven design: the new challenges of digitalization on product design and development. Des. Sci. 6, e27. https://doi.org/10.1017/dsj.2020.25

Chen, L., Wang, P., Dong, H., Shi, F., Han, J., Guo, Y., Childs, P.R.N., Xiao, J., Wu, C., 2019. An artificial intelligence based data-driven approach for design ideation. Journal of Visual Communication and Image Representation 61, 10–22. https://doi.org/10.1016/j.jvcir.2019.02.009

Chiarello, F., Belingheri, P., Fantoni, G., 2021. Data science for engineering design: State of the art and future directions. Computers in Industry 129, 103447. https://doi.org/10.1016/j.compind.2021.103447

Chiarello, F., Bonaccorsi, A., Fantoni, G., 2020. Technical Sentiment Analysis. Measuring Advantages and Drawbacks of New Products Using Social Media. Computers in Industry 123, 103299. https://doi.org/10.1016/j.compind.2020.103299

Chiarello, F., Fantoni, G. and Bonaccorsi, A., 2017. Product description in terms of advantages and drawbacks: Exploiting patent information in novel ways, In DS 87-6 Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 6: Design Information and Knowledge, Vancouver, Canada, 21-25.08. pp. 101–110.

Chien, C.-F., Kerh, R., Lin, K.-Y., Yu, A.P.-I., 2016. Data-driven innovation to capture user-experience product design: An empirical study for notebook visual aesthetics design. Computers & Industrial Engineering 99, 162–173. https://doi.org/10.1016/j.cie.2016.07.006

Chiu, M.-C., Lin, K.-Z., 2018. Utilizing text mining and Kansei Engineering to support data-driven design automation at conceptual design stage. Advanced Engineering Informatics 38, 826–839. https://doi.org/10.1016/j.aei.2018.11.002

Chowdhery, A.S., Bertoni, M., Wall, J., Larsson, T., n.d. A data-driven design framework for early stage PSS design exploration 16.

Dalzochio, J., Kunst, R., Pignaton, E., Binotto, A., Sanyal, S., Favilla, J., Barbosa, J., 2020. Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges. Computers in Industry 123, 103298. https://doi.org/10.1016/j.compind.2020.103298

Deng, Q., Wellsandt, S., Hribernik, K.A., Thoben, K.-D., 2019. Towards Understanding the Role of Product Usage Information in Product Design Improvement, in: Fortin, C., Rivest, L., Bernard, A., Bouras, A. (Eds.), Product Lifecycle Management in the Digital Twin Era. Springer International Publishing, Cham, pp. 369–378. https://doi.org/10.1007/978-3-030-42250-9_35

Design Council UK. 2005. The Design Process.

Faste R., Roth B. and Wilde D.J., 1993. Integrating Creativity into the Mechanical Engineering Curriculum, ASME Resource Guide to Innovation in Engineering Design, American Society of Mechanical Engineers.

Ghosh, D., Olewnik, A., Lewis, K., Kim, J., Lakshmanan, A., 2017. Cyber-Empathic Design: A Data-Driven Framework for Product Design. Journal of Mechanical Design 139, 091401. https://doi.org/10.1115/1.4036780

Gill, P., Stewart, K., Treasure, E., Chadwick, B., 2008. Methods of data collection in qualitative research: interviews and focus groups. Br Dent J 204, 291–295. https://doi.org/10.1038/bdj.2008.192

Gorkovenko, K., Burnett, D.J., Thorp, J.K., Richards, D., Murray-Rust, D., 2020. Exploring The Future of Data-Driven Product Design, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Presented at the CHI ’20: CHI Conference on Human Factors in Computing Systems, ACM, Honolulu HI USA, pp. 1–14. https://doi.org/10.1145/3313831.3376560

Holler, M., Stoeckli, E., Uebernickel, F., Brenner, W., n.d. Towards Understanding closed-loop PLM: The Role of Product Usage Data for Product Development enabled by intelligent Properties 14.
Hou, L., Jiao, R.J., 2020. Data-informed inverse design by product usage information: a review, framework and outlook. J Intell Manuf 31, 529–552. https://doi.org/10.1007/s10845-019-01463-2
Isaksson, O., Eckert, C., 2020. Product Development 2040. The Design Society. https://doi.org/10.35199/report.pd2040
Kim, H. H. M., Liu, Y., Wang, C. C., and Wang, Y., 2017. Special Issue: Data-Driven Design (D3), Journal of Mechanical Design 139, 11. https://doi.org/10.1115/1.4037943
Klein, P., van der Vegte, W.F., Hribernik, K., Klaus-Dieter, T., 2019. Towards an Approach Integrating Various Levels of Data Analytics to Exploit Product-Usage Information in Product Development. Proc. Int. Conf. Eng. Des. 1, 2627–2636. https://doi.org/10.1017/dsi.2019.269
Kusiak, A., 2009. Innovation: A data-driven approach. International Journal of Production Economics 122, 440–448. https://doi.org/10.1016/j.ijpe.2009.06.025
Kwon, H., Park, Y., Geum, Y., 2018. Toward data-driven idea generation: Application of Wikipedia to morphological analysis. Technological Forecasting and Social Change 132, 56–80. https://doi.org/10.1016/j.techfore.2018.01.009
Labrinidis, A., Jagadish, H.V., 2012. Challenges and opportunities with big data. Proc. VLDB Endow. 5, 2032–2033. https://doi.org/10.14778/2367502.2367572
Li, J., Tao, F., Cheng, Y., Zhao, L., 2015. Big Data in product lifecycle management. Int J Adv Manuf Technol 81, 667–684. https://doi.org/10.1007/s00170-015-7151-x
Lin, K.-Y., Chien, C.-F., Kerh, R., 2016. UNISON framework of data-driven innovation for extracting user experience of product design of wearable devices. Computers & Industrial Engineering 99, 487–502. https://doi.org/10.1016/j.cie.2016.05.023
Lützenberger, J., Klein, P., Hribernik, K., Thoben, K.-D., 2016. Improving Product-Service Systems by Exploiting Information From The Usage Phase. A Case Study. Procedia CIRP 47, 376–381. https://doi.org/10.1016/j.procir.2016.03.064
Lyyra, A.K., Koskinen, K.M., 2016. The Ambivalent Characteristics of Connected, Digitised Products: Case Tesla Model S, in: Lundh Snis, U. (Ed.), Nordic Contributions in IS Research. Springer International Publishing, Cham, pp. 57–69. https://doi.org/10.1007/978-3-319-43597-8_5
Ma, H., Chu, X., Lyu, G., Xue, D., 2017. An Integrated Approach for Design Improvement Based on Analysis of Time-Dependent Product Usage Data. Journal of Mechanical Design 139, 111401. https://doi.org/10.1115/1.4037246
Maia, C.L.B., Furtado, E.S., 2019. An Approach to Analyze User's Emotion in HCI Experiments Using Psychophysiological Measures. IEEE Access 7, 36471–36480. https://doi.org/10.1109/ACCESS.2019.2904977
Maia, C.L.B., Furtado, E.S., 2016. A study about psychophysiological measures in user experience monitoring and evaluation, in: Proceedings of the 15th Brazilian Symposium on Human Factors in Computing Systems. Presented at the IHC '16: XV Brazilian Symposium on Human Factors in Computing Systems, ACM, São Paulo Brazil, pp. 1–9. https://doi.org/10.1145/3033701.3033708
Pahl, G., Beitz, W., & Wallace, K., 1996. Engineering Design: A Systematic Approach: Design Council.
Peruzzini, M., Grandi, F., Pellicciari, M., Campanella, C.E., 2019. User Experience Analysis Based on Physiological Data Monitoring and Mixed Prototyping to Support Human-Centre Product Design, in: Rebelo, F., Soares, M.M. (Eds.), Advances in Ergonomics in Design. Springer International Publishing, Cham, pp. 401–412. https://doi.org/10.1007/978-3-319-94706-8_44
Porter, M. E. & Heppelmann, J. E. 2014 How smart, connected products are transforming competition. Harvard Business Review 92 (11), 64–88.

Porter, M. E. & Heppelmann, J. E. 2015 How smart, connected products are transforming companies. Harvard Business Review 93 (10), 97–114

Ranscombe, C., Kinsella, P., Blijlevens, J., 2017. Data-Driven Styling: Augmenting Intuition in the Product Design Process Using Holistic Styling Analysis. Journal of Mechanical Design 139, 111417. https://doi.org/10.1115/1.4037249

Romelfanger, M., Kolich, M., 2019. Comfortable automotive seat design and big data analytics: A study in thigh support. Applied Ergonomics 75, 257–262. https://doi.org/10.1016/j.apergo.2018.08.020

Rossi, M., Germani, M., Zamagni, A., 2016. Review of ecodesign methods and tools. Barriers and strategies for an effective implementation in industrial companies. Journal of Cleaner Production 129, 361–373. https://doi.org/10.1016/j.jclepro.2016.04.051

Rowe, P.G., 1987. Design thinking, Cambridge, MIT press.

Shin, J.-H., Kiritsis, D., Xirouchakis, P., 2015. Design modification supporting method based on product usage data in closed-loop PLM. International Journal of Computer Integrated Manufacturing 28, 551–568. https://doi.org/10.1080/0951192X.2014.900866

Schultze, U., Avital, M., 2011. "Designing Interviews to Generate Rich Data for Information Systems Research" Information and Organization 21 (1), 1–16. https://doi.org/10.1016/j.infoandorg.2010.11.001

Song, B., Luo, J., 2017. Mining Patent Precedents for Data-Driven Design: The Case of Spherical Rolling Robots. Journal of Mechanical Design 139, 111420. https://doi.org/10.1115/1.4037613

van der Vegte, W.F., Kurt, F., Şengöz, O.K., 2019. Simulations Based on Product-Usage Information From Connected Products to Support Redesign for Improved Performance: Exploration of Practical Application to Domestic Fridge-Freezers. Journal of Computing and Information Science in Engineering 19, 031003. https://doi.org/10.1115/1.4042537

van Eck, M.L., Markslag, E., Sidorova, N., Brosens-Kessels, A., van der Aalst, W.M.P., 2019. Data-Driven Usability Test Scenario Creation, in: Bogdan, C., Kuusinen, K., Lárusdóttir, M.K., Palanque, P., Winckler, M. (Eds.), Human-Centered Software Engineering. Springer International Publishing, Cham, pp. 88–108. https://doi.org/10.1007/978-3-030-05909-5_6

van Eck, M.L., Sidorova, N., van der Aalst, W.M.P., 2016. Enabling process mining on sensor data from smart products, in: 2016 IEEE Tenth International Conference on Research Challenges in Information Science (RCIS). Presented at the 2016 IEEE Tenth International Conference on Research Challenges in Information Science (RCIS), IEEE, Grenoble, France, pp. 1–12. https://doi.org/10.1109/RCIS.2016.7549355

Voet, H., Altenhof, M., Ellerich, M., Schmitt, R.H., Linke, B., 2019. A Framework for the Capture and Analysis of Product Usage Data for Continuous Product Improvement. Journal of Manufacturing Science and Engineering 141, 021010. https://doi.org/10.1115/1.4041948

Wilberg, J., Lau, K., Nützel, T., Hollauer, C., Omer, M., 2018. DEVELOPMENT OF A CATALOGUE SUPPORTING IDEA GENERATION FOR INTERNET OF THINGS USE CASES. Presented at the 15th International Design Conference, pp. 1453–1464. https://doi.org/10.21278/idc.2018.0215

Yang, B., Liu, Y., Liang, Y., Tang, M., 2019. Exploiting user experience from online customer reviews for product design. International Journal of Information Management 46, 173–186. https://doi.org/10.1016/j.ijinfomgt.2018.12.006

Yu, C., Zhu, L., 2016. Product design pattern based on big data-driven scenario. Advances in Mechanical Engineering 8, 168781401665680. https://doi.org/10.1177/1687814016656805

Zheng, P., Xu, X., Chen, C.-H., 2020. A data-driven cyber-physical approach for personalised smart, connected product co-development in a cloud-based environment. J Intell Manuf 31, 3–18. https://doi.org/10.1007/s10845-018-1430-y

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: