Design Inquiry Through Data

Kun, P.

DOI
10.4233/uuid:7e914dd9-2b53-4b2c-9061-86087dbb93b9

Publication date
2020

Document Version
Final published version

Citation (APA)
Kun, P. (2020). *Design Inquiry Through Data*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:7e914dd9-2b53-4b2c-9061-86087dbb93b9

Design Inquiry Through Data

DISSERTATION

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, Prof.dr.ir. T.H.J.J. van der Hagen,
Chair of the Board for Doctorates
to be defended publicly on
Tuesday 25 August 2020 at 12:30 o'clock

by

Péter KUN
Master of Science in Interaction Design and Technologies,
Chalmers University of Technology, Sweden
born in Jászberény, Hungary

This dissertation has been approved by the promotors.

COMPOSITION OF THE DOCTORAL COMMITTEE:

Rector Magnificus,	chairperson
Prof. dr. G. W. Kortuem	Delft University of Technology, promotor
Dr. I. J. Mulder	Delft University of Technology, copromotor

INDEPENDENT MEMBERS:

Prof. dr. ir. A. Bozzon	Delft University of Technology
Prof. dr. P. Coulton	Lancaster University
Prof. dr. ir. M. S. Kleinsmann	Delft University of Technology
Prof. dr. M. M. Specht	Delft University of Technology
Dr. A. Wolff	Lappeenranta-Lahti University of Technology

The research has received funding from the European Commission under grant agreement H2020-ICT-2015-687818 (Open4Citizens).

Péter Kun – peter@peterkun.com

ISBN:
978-94-6384-154-2

TYPEFACE:
Source Sans Pro, Source Serif Pro

LAYOUT DESIGN:
Péter Kun

Table of Contents

Chapter 1

Introduction

1.1 Problem description

The area of design has expanded rapidly since the late 1960s, both in academic discourse and in industry. While the specific meaning of the word ‘design’ within more narrowly defined particular contexts has not been lost, the concept of design as a whole has become more and more encompassing (Buchanan, 2001). Expanding far beyond beautification and form-giving, or the technical conception and creation of artifacts, processes and organizations, design is increasingly aiming to bring a creative capacity to tackle complex problems – problems without simple short-term solutions, such as environmental degradation, health, poverty, or education. Although, the move towards more complexity has already been reflected in early scholarly work, such as Rittel and Webber’s concept of wicked problems (Rittel & Webber, 1973) and Buchanan’s concept of ill-defined problems (Buchanan, 1992), the debate on design and complexity is still ongoing. With their introduction of DesignX, Norman and Stappers (2015) have added new dimensions to this timely debate: DesignX highlights the need for designing at multiple scales and multiple disciplines. Dorst (2015b) have described the nature of contemporary problems as *“open, complex, dynamic, and networked”*, and suggests that the role of designers for solving contemporary problems is to bring the designerly capacity of framing and reframing to transdisciplinary teams. These two examples indicate that for tackling complex problems, designers are unable to operate in a vacuum – they need the expertise of others involved. Consequently, it can be concluded that to tackle such complex problems, design techniques on their own are not sufficient, new techniques are necessary to achieve sufficient impact.

Today’s context for design can also be seen as a ‘datafied’ world. Datafication (Lycett, 2013) refers to the trend of how many aspects of the world are getting rendered as data in large data infrastructures. To illustrate the increasing ubiquity of digital data in the complex problem domains that designers tackle, sensor networks are often used to track traffic on roads to inform urban environments or to track physiological measures to inform medicine. Furthermore, digital and connected artifacts enable precise logging, collection, and processing of users' actions. Billions of people use instant messaging over the internet to communicate and post on social media. Data in today's big data era is complex, heterogeneous, and ubiquitous in all aspects of life (Mayer-Schönberger & Cukier, 2013; Kitchin, 2014a). In the context of extracting value from such heterogeneous and complex datasets, different data practices have emerged under the field of data science (Cao, 2017). Data science as a field and profession has synthesized decades-long developments from fields such as data mining or information visualization (Card et al., 1999; Fayyad et al., 1996), and today it broadly refers to all the different ways to yield value out from data. These new data practices have inspired expert and non-expert communities to start employing massive datasets as a new lens for understanding the world in their respective domains. The spreading of data-enabled inquiry is wide: fields such as the natural sciences, social sciences, or the humanities have been affected by data-enabled inquiry (Kitchin, 2014b). For example, scientists can observe how people interact with each other at a massive scale on online social networks (Lazer et al., 2009), and use the gained knowledge to design better crisis responses (Bruns & Liang, 2012). In the humanities, computation enables data-enabled inquiry by turning unstructured data into structured data, such as processing the scans of old texts through optical character recognition (OCR) and make them available for quantitative text analysis. Data practices are no longer solely conducted by experts, instead, a growing number of non-expert communities have emerged to extract value from data. For example, data journalists use data storytelling and data visualizations to enhance reporting and to gain deeper insights (Gray et al., 2012). Another example are citizen scientists, who – often by collaborating with designers – use non-expert data practices and tools to collect data as evidence on their cause, and as an input for participatory design work (Coulson et al., 2018). Such emerging data science practices indicate opportunities for designers to develop their own data practices for conducting research, problem framing, and use data as a creative resource throughout the design process.

In this dissertation, we will develop the argument that data science is an important source of expertise for design and that digital data represents a new creative lens for design inquiry. In this dissertation, we build on Dalsgaard's definition of design inquiry as an *"explorative and transformative process through which designers draw upon their repertoire of knowledge and competences as well as resources in the situation, including instruments, in order to create something novel and appropriate that changes an incoherent or undesirable situation for the better"* (Dalsgaard, 2017, p. 24). Inquiry is a fundamental element of design (Nelson & Stolterman, 2012) and with the maturity of design as a field, an extensive repertoire of established design techniques are taught, used and made available for designers, such as running an interview study or using sketching as a way of thinking. While such established inquiry techniques to observe and intervene in the physical world are common, data offers access to scale, level of detail, or timeframes that otherwise would be inaccessible or inconvenient with established methods. In this dissertation, we will argue that there are vast opportunities to expand design inquiry into data and to use data for revealing previously hidden aspects of the physical world.

1.2 Research focus

PROBLEM STATEMENT

Design and data science are two disciplines with different epistemic goals, practices, and methods. Designer and data scientist collaborations are on the rise in the datafied world. Current industry practices indicate opportunities for embedded big data techniques into the design process, for example, by including data scientists into the design team (Dove et al., 2017; Yang et al., 2018). Such an approach is affordable for technology companies with large budgets, but problematic to apply broadly in all contexts and situations design operates. Although exemplary studies such as by Dove and colleagues (2017) or Yang and colleagues (2018) indicate a strong need for designers to leverage data, there is a lack of empirical knowledge on *how* designers utilize and tailor data science practices and methods. Furthermore, current data practices in design are primarily limited to the solution space. For example, A-B testing is a widely used practice in the software industry, which refers to using statistical methods to measure significant differences in user behavior between alternative design solutions (King et al., 2017). While A-B testing shows that data techniques can be valuable at later stages of the design process, it has limited value for design inquiry where the focus is on understanding and framing the problem and identifying opportunities. A designerly approach is generally characterized by open-ended exploration and a continuous reframing of the problem based on new findings, requiring tailored data frameworks and methods. However, at present, there is limited understanding of how the fields of design and data science could intersect and there is a lack of practical data methods and techniques for design.

GOAL

This research aims to **develop theoretical and practical knowledge on the intersection of design and data science to enable designers to use data-rich practices for design inquiry.** The research aim is complemented with a design goal to **develop methodological contributions to support future data-rich design practices.**

These goals lead us to formulate the following overarching research question:

› **Main RQ:** How can designers integrate data practices into design inquiry?

The dissertation will develop answers to five specific research questions towards addressing the main research question:

› **RQ1:** How can design and data science be aligned as mode of inquiry?

› **RQ2:** How do designers appropriate data science practices for design inquiry?

› **RQ3:** How can data science practices be characterized through a creative process lens?

› **RQ4:** How can a design method support design inquiry through data?

› **RQ5:** How do designers adopt a data-rich design methodology?

These research questions guide the research and design approach that will be presented in the next section.

1.3 Research approach

The scientific philosophy behind this dissertation is inspired by John Dewey's pragmatist worldview (Creswell, 2009; Dixon, 2019), which has influenced a number of theoretical and practical conceptions of design research (Dalsgaard, 2014). Pragmatist research uses multiple methods to gain a more complete understanding of a phenomenon in the lights of multiple contexts, for example, social and political contexts. As such, pragmatism opens up the possibility of mixed methods research approaches (Creswell & Clark, 2017), flexibly combining qualitative and quantitative methods as the research unfolds.

In this dissertation, we use Research-through-Design (RTD) as a methodological approach. According to Stappers and Giaccardi (2017), RTD uses design activities that play a formative role in knowledge generation, often through the design and deployment of prototypes or artifacts to enable interactions that then become observable throvugh design. In such a process of design and deployment, the designer inherently faces opportunities and obstacles to make the best judgment of the real-world situation. In RTD, the thinking process put into the creation of an artifact is also used for insights and knowledge generation. RTD is particularly suited for researching the interdependence between designed artifacts and the practices they enable and support. In the context of data and design, we are aiming to understand how novel methods and methodologies enable and support data-rich design practices.

In our RTD approach shown in Figure 1.1, we design workshops, a methodology, and a corresponding method, in order to study emerging data-rich design practices empirically. Such 'intangible' artifacts in RTD (e.g. (Mattelmäki, 2006)) are not without precedent (Stappers and Giaccardi, 2017). In our approach, design artifacts and empirical research are intertwined in an iterative process of gradual refinement. The figure shows how the theoretical framing for RTD design and research is provided by the ‘Design Inquiry Through Data’ framework, which we develop in Chapter 2.

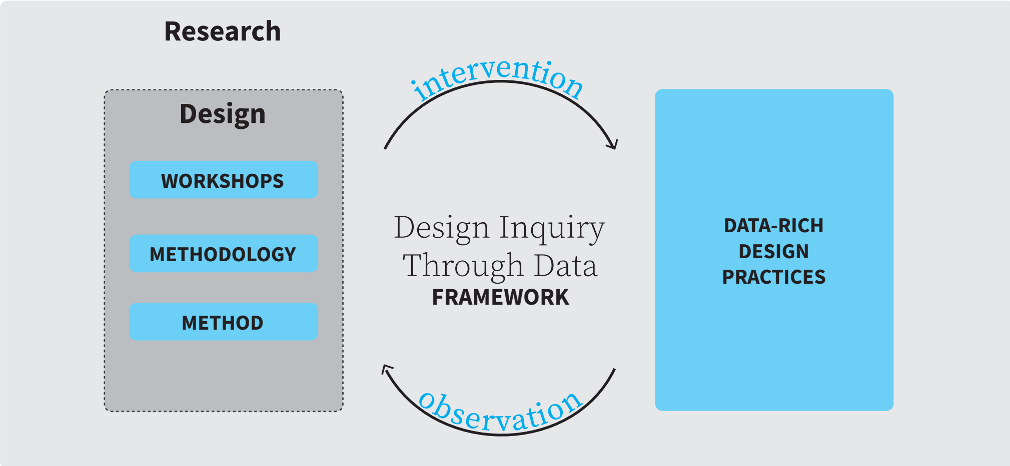


Figure 1.1. Research-through-Design is used to combine the development of methodological support for data-rich design practices. The interventions and observations are operationalized by the Design Inquiry Through Data framework.

The five research questions are addressed throughout five chapters, as shown in Figure 1.2. First, in Chapter 2, we address the first research question in order to frame the relationship between design inquiry and data science. In the pursuit of contributing to the body of knowledge at the intersection of data and design, we motivate a conceptual framework that will be the basis of empirical investigation. The proposed conceptual framework is developed to understand the future data-rich design practice of *Design Inquiry Through Data*, also represented among the contributions in Figure 1.3. Afterwards, Chapters 3, 4, 5, and 6 present empirical investigations. The empirical investigations take a dual role of 1) being sites of data collection for scientific inquiry of the research questions of the thesis, 2) being the sites of interventions where *Design Inquiry Through Data* is iterated upon in an RTD program through multiple chapters. The design process of the RTD program starts in Chapter 3, which can be seen as the initial research into data-rich design practices. As an envisioned design output, Chapter 3 results in an enhanced understanding

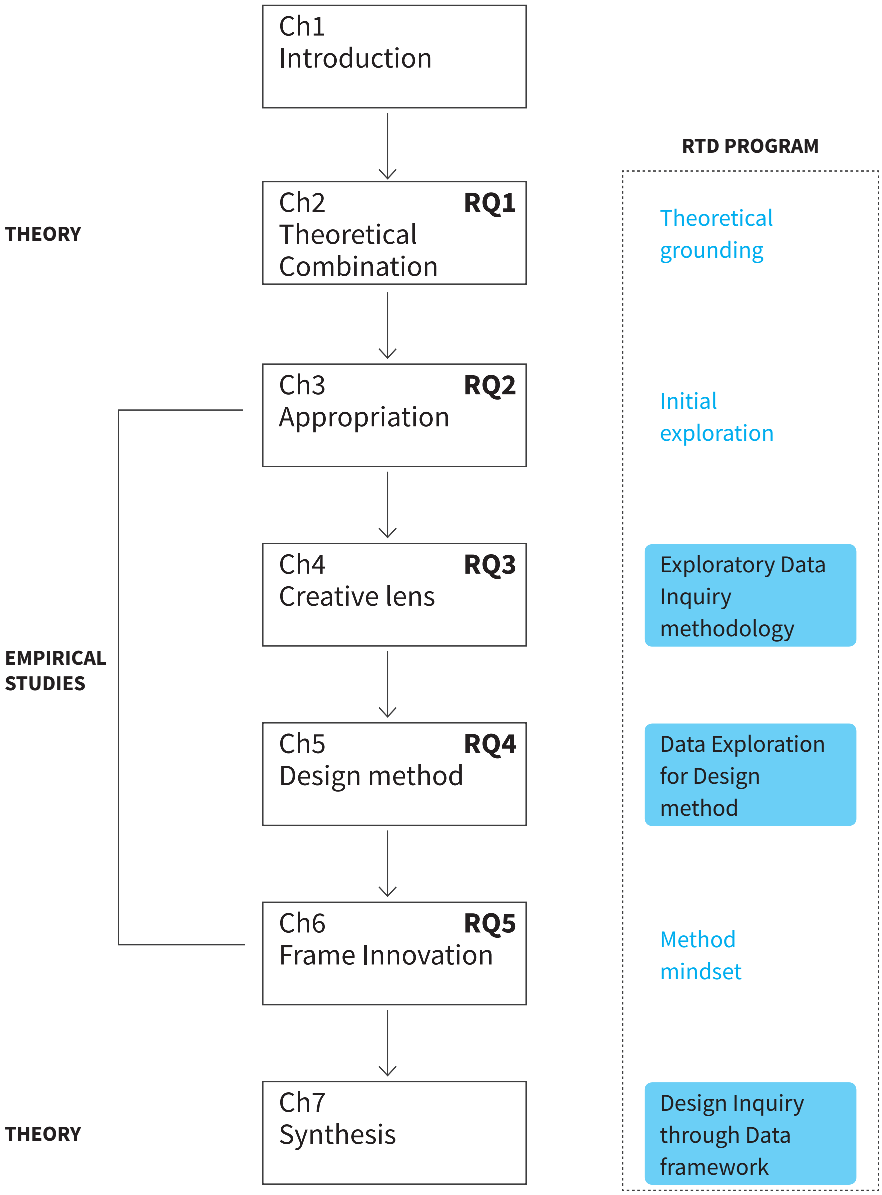


Figure 1.2. Outline of the dissertation. Empirical studies are combined with RTD design activities. The RTD process aims to develop a prototypical future design practice, and to conclude on that, a design framework is generated and then a design method.

of a future design practice where data is used for design inquiry. In Chapter 4, the design exploration results in a methodology – 'Exploratory Data Inquiry' – to inform new methods, tools, and techniques, as depicted in Figure 1.3. In Chapter 5, the *Exploratory Data Inquiry* methodology is put into use to inform the development of the 'Data Exploration for Design' method and accompanying design tools. In Chapter 6, we evaluate the *Exploratory Data Inquiry* methodology and target our investigation on understanding its tacit components. In Chapter 7, we return to the theoretical investigation by reflecting on the overall empirical work and synthesizing the findings into the concluded *Design Inquiry Through Data* framework.

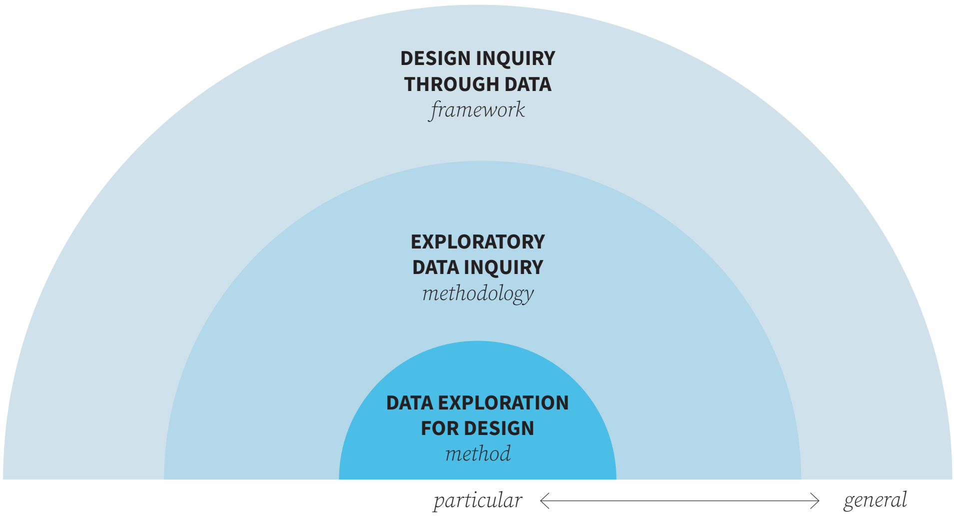


Figure 1.3. The hierarchy between the different design contributions of the thesis. The outer circle represents the design framework of Design Inquiry Through Data, the abstract level and general contribution. The second layer is the methodology of Exploratory Data Inquiry, which describes the combined practice of data exploration in design inquiry. The inner layer is the Data Exploration for Design method, which instantiates the principles of Exploratory Data Inquiry methodology.

1.4 Contributions

The dissertation makes three types of interdependent contributions:

The first type of contribution is a **theoretical deepening** of the apparent cross-section between design theory and data science, providing an improved and shared vocabulary of how these distinct fields can intertwine during design practice. In Chapter 2, the theoretical investigation provides an in-depth perspective into inquiry through the lenses of design and data, and in Chapter 4, we provide an improved framing of data practices from a creative process perspective. In Chapter 7, we interpret and position the findings of the empirical studies.

The second type of contribution refers to **rich empirical insights** on how novice designers use design inquiry through data, providing value for design educators and designers of non-expert data tools (Chapter 3, 4, 5, and 6).

The third types of contributions are the emerged **Design Inquiry Through Data framework** (Chapter 7), the **Exploratory Data Inquiry methodology** (Chapter 4), and accompanying **Data Exploration for Design method** (Chapter 5), as shown in Figure 1.3. The Design Inquiry Through Data framework provides a versatile perspective on how data exploration can be supported within design practice, the Exploratory Data Inquiry methodology provides a formalization of a combined data and design practice, and the Data Exploration for Design method illustrates an instantiation of the Exploratory Data Inquiry methodology as a method.

1.5 Reader’s guide

As already visualized in Figure 1.2, the dissertation consists of seven chapters, each contributing to the goal listed above. The *current chapter* has motivated the research to explore the opportunity of using data techniques for inquiring about the world. The goal and the focus of the research are presented, followed by an elaboration of the research approach.

Chapter 2 first provides a more in-depth background on design and data following a literature study to align theoretical conceptions of design inquiry and data science. The chapter concludes with a conceptual framework for studying design inquiry through data, and a discussion of the setup and methodological considerations for empirical studies.

Chapter 3 presents an exploratory study, which investigates how designers appropriate data science practices, using techniques borrowed from non-expert data communities. This study provides a preliminary understanding of how data techniques are appropriated and in what directions the thinking process of designers change.

Chapter 4 focuses on analyzing data practices through a creative process lens. This study reveals how creativity manifests when designers use data as a mode of inquiry. Moreover, the study informs the presented *Exploratory Data Inquiry* methodology by combining data exploration and design inquiry in the same intertwined practice.

Chapter 5 builds on the *Exploratory Data Inquiry* methodology from the previous chapter. In this chapter, we elaborate upon the *Data Exploration for Design* method based on the *Exploratory Data Inquiry* methodology. The design method, consisting of a process model and design tools to lower the learning curve, is evaluated in a study to assess its capacity for creativity support.

Chapter 6 leverages previous findings into one comprehensive study where design inquiry through data is embedded in a frame innovation research setup. The presented study shows the use of design inquiry through data in a more realistic design situation. It provides deeper insights into how the mindset and thinking of designers change when using data techniques in design work.

Chapter 7 synthesizes the findings of the different studies and provides a general discussion of the dissertation. First, we discuss the process of opportunistic data exploration and the type of hybrid mindset designers assume, and then we offer a reframing of visualizations as prototypes. Then we conclude on the *Design Inquiry Through Data* framework that combines the aforementioned process, mindset, and tools. The chapter closes with a reflection on implications, ethical considerations, and future work on design inquiry through data.

Chapter 2

Relationship Between Design and Data Practices

This chapter frames data in the context of design theory and more specifically design inquiry. In this chapter, we address RQ1 of the thesis, "How can design and data science be aligned as mode of inquiry?" Towards answering this question, we conduct a literature study to contrast and compare theoretical conceptions of design and data science. To achieve this, we first elaborate on what is commonly understood about designing and then zoom-in on design inquiry specifically. We then argue for an interdisciplinary perspective of data and then zoom-in on the practices of data scientists. In the second half of the chapter, we align the interdisciplinary lens of data into design inquiry as a conceptual framework. This conceptual framework will be used as a basis for empirical studies presented in Chapter 3, 4, 5 and 6. We close the chapter discussing the setup and methodological considerations for the empirical studies.

Parts of this chapter are based on:

Kun, P., Mulder, I., & Kortuem, G. (2018). Design Enquiry Through Data: Appropriating a Data Science Workflow for the Design Process. In *Proceedings of the 32nd International BCS Human Computer Interaction Conference (HCI 2018)*. BCS Learning and Development Ltd. https://doi.org/10.14236/ewic/HCI2018.32

Kun, P., Mulder, I., De Götzen, A., & Kortuem, G. (2019). Creative Data Work in the Design Process. In *Proceedings of the 2019 ACM Conference on Creativity and Cognition.* ACM. https://doi.org/10.1145/3325480.3325500

Kun, P., Mulder, I., & Kortuem, G. (Under review at *Interaction Design and Architecture*). Developing a Design Inquiry Method for Data Exploration.

2.1 Design

Designing is a core practice of innovation and has been applied historically to form-giving in the industrial production era, and increasingly to designing interactions in a digitalized world, to services and increasingly complex systems (Buchanan, 2001). How the scope of design practice has escalated can be referred back to tackling wicked problems (Rittel & Webber, 1973), and applying a design-specific flavor of problem-solving to ill-defined problems. Cross has coined such a problem-solving practice as “designerly ways of doing” (1982), which – among other characteristics – refers to a highly iterative trial-and-error process, the use of sketching and prototypes as a mode of thinking, and approaching ill-defined problems with a limited amount of information.

Since Cross (1982) summarized the unique characteristics of designing decades ago, design research has further described how design practice works. In their seminal paper, Dorst and Cross (2001) opened up how the co-evolution of problem and design space happens throughout the design process. As they describe, throughout the process of designing, two conceptual spaces are evolving, one that is concerned with the problem being solved (i.e., the problem space) and another one that is concerned with the potential solutions to that problem (i.e., the solution space). Dorst and Cross (2001) based their work on the co-evolution model of Maher et al. (1996). In their paper, Maher et al. reflected on the difference between search and exploration, originally in the context of representing generic design processes and rationalizing these concepts for genetic algorithms. As Figure 2.1 shows, moving from an ill-defined problem takes place through exploration, and this exploration generates a well-defined problem (or at least, a better-defined problem) and a solution. Dorst and Cross (2001) showed that in the case of design practice, this exploration is a continuous co-evolution of the designer increasingly understanding the problem space and the design space. In the next section, we detail what this exploration process entails.

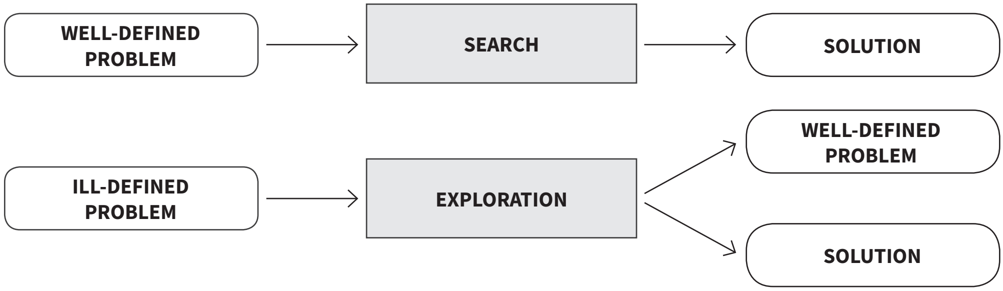


Figure 2.1. Difference between search and exploration according to Maher et al. (1996).

In an iterative design process, the design space evolves through prototypes of increasing fidelity, while the problem space is getting increasingly defined too. Although design can be seen as a specific approach to creative problem solving, it is non-trivial how to solve ill-defined and wicked problems. Schön (1988) introduced the term 'problem setting' as a precursory step of solving a problem, where a designer develops skills in reducing 'mess' around a problem situation to identify the characteristics of the problem that will be tackled. In the context of design practice, this notion was further expanded to the notion of 'problem framing', or more generally referred to as framing and reframing (Dorst, 2011). As Dorst (2015b) explains, the process of framing and reframing makes designing a unique type of problem-solving, and framing is the most differentiating competence between expert and novice designers. Depending on the specific design problem, the thinking patterns designers assume can be deductive, inductive, or abductive (Dorst, 2011). In Dorst's words, deduction refers to the reasoning pattern where we predict an outcome based on the 'elements' of a situation. Such as, if we know that there are stars in the sky and we know the natural laws that govern their movements, we can predict where the stars will be at a given moment. Induction, on the other hand, refers to the reasoning pattern where we know what are the different 'elements' involved in a phenomenon and what the outcome will be, but we do not know what governs the elements towards the specific outcomes. Induction is the reasoning pattern behind much of scientific knowing, where people generate models and increasingly detailed descriptions to explain a phenomenon. While deductive and inductive reasoning patterns are pervasive in science or other analytical professions, abductive reasoning is typical and specific of design practice (Dorst, 2011; Kolko, 2009). Abductive reasoning stands for a problem-solving reasoning pattern, where we know the outcome to achieve, and we know the patterns (i.e., models, laws of physics, descriptions), but we need to define what 'elements' will form the solution. Such an approach for problem-solving is referred to as design abduction. Design abduction is the type of abduction in the context of the co-evolution of problem and design space, where the problem-to-be-solved changes with the desired outcome (Dorst, 2011). With the changing problem though, the solution space is also evolving, making the 'elements' involved and the patterns governing their relationships also a subject of creative work.

Abduction, as mentioned above, leads to an open-ended changing of the problem being solved, resulting in an **exploratory** design practice. The continuous learning about the design problem by an exploratory design practice makes the process resemble a bricolage (Louridas, 1999; Vallgårda & Fernaeus, 2015). A *bricoleur* designer explores opportunistically; if there is a way to learn more about the problem domain, designers use any information that can lead them towards better solutions or well-defined problem space. Designers in practice have also been found to diverge opportunistically from a structured plan or methodical process (Cross, 2004). In other words, while designers follow a conscious process, design practice can be seen as opportunistic, where the choosing of techniques to progress the problem and design space are strategic choices of "designing the design process" (Guindon, 1990). In practice, the strategic choices are often about the selection of appropriate design methods, which we will further unpack in the following section.

DESIGN METHODS

Design methods have long been used to codify designers' best practices (Jones, 1970, 1992), as a way of rationalizing the design process, standardize best practices, and enabling designers to collaborate better with colleagues from other professions. Design methods have become essential in design education (van Boeijen et al., 2020) and to open up generic innovation processes for the masses. This magnitude is well-illustrated with the Organization for Economic Co-operation and Development's (OECD) Observatory for Public Sector Innovation (OPSI) project, that enlists over 165 toolkits within public sector innovation (Toolkit Navigator, 2020). While these toolkits have fostered opening up the design process to demographics traditionally not trained in design, they may suggest the oversimplification of design methods. In researching how design students and expert designers use methods, the mental component of methods has been becoming central in recent works, as shown in the following examples. Design methods are used as 'mental tools' rather than prescribed recipes towards a specific design outcome (Daalhuizen, 2014). Moreover, it can be said that design methods integrate into a designer's mindset as tools to answer different questions (Gray, 2016). Furthermore, design methods evolve and adapt to circumstances in design practice (Schønheyder & Nordby, 2018). These findings urge developers of design methodologies not only to attend methods as 'process descriptions', but also to consider the corresponding **mindset**; the tacit component of how designers grow together with their methods and how methods become an integral part of designers' thinking patterns. In the development of a method, **step-by-step guides can be made to support novices**. However, it appears to be more crucial to consider the higher-level design activity goals a designer wants to achieve by using a method, and therefore to develop the method with the intended mindset in mind. In this way, designers of methods need to take into account that the users of methods grow expertise and open-endedly adapt methods in use.

We started this chapter from a high-level description of the design process and then illustrated the nuances of designing in practice being exploratory, opportunistic, and following a continuous framing and reframing of the design and problem space. Next, we further narrow down the scope and focus on how designers understand phenomena and 'bring back' their insights to designing.

DESIGN INQUIRY

In keeping with Maher et al.'s (1996) model that has been introduced in Figure 2.1, we view *exploration* as a key activity in moving from ill-defined problems towards well-defined problems. However, ‘exploration’ is still a generic concept, which can be re-interpreted with the concept of ‘design inquiry’. Dalsgaard defined design inquiry as an *“..explorative and transformative process through which designers draw upon their repertoire of knowledge and competences as well as resources in the situation, including instruments, in order to create something novel and appropriate that changes an incoherent or undesirable situation for the better”* (Dalsgaard, 2017, p. 24). This definition suits the exploratory, opportunistic and continuous framing and reframing of the design and problem space that we highlighted before as characteristics of designing. Nelson and Stolterman (2012) see design inquiry as a compound of three forms of inquiry, composed of “ideal”, “true” and “real” approaches to gaining knowledge (Nelson & Stolterman, 2012, p. 34). Under *ideal*, they mean inquiry seeking a desired state; under *true* they mean the scientific inquiry of seeking facts, and under *real* they mean inquiry seeking the ultimate particular (i.e., a *specific* artifact that exists in context). As they put, the three forms of inquiry work together in concert during designing towards seeking knowledge. Moreover, they also introduce a plurality of outcomes of inquiry (Nelson & Stolterman, 2012, p. 39). They alter the types of knowledge as outcomes of inquiry, namely an outcome can be a reason (conscious knowledge), or intuition (unconscious knowledge), or imagination (subconscious knowledge), but also the product of conscious not-knowing, that they refer to as ‘design thinking’. In their framing for design thinking, it is the quality of mind that is open to what is emergent in the moment of designing. This quality of mind can be interpreted as exploratory and opportunistic characteristic of designing from before, underlying the connection to inquiry.

Nelson and Stolterman’s understanding of inquiry illustrates the richness of design inquiry that goes beyond scientific inquiry. We will put their rich description of inquiry aside now and boil it down to ‘moving from an *unknown* state towards a *known* state’, which enables us to align different interpretations of inquiry. Dalsgaard’s (2017) interpretation of inquiry refers to a move from uncertain situations towards stable situations, as shown in Figure 2.2.

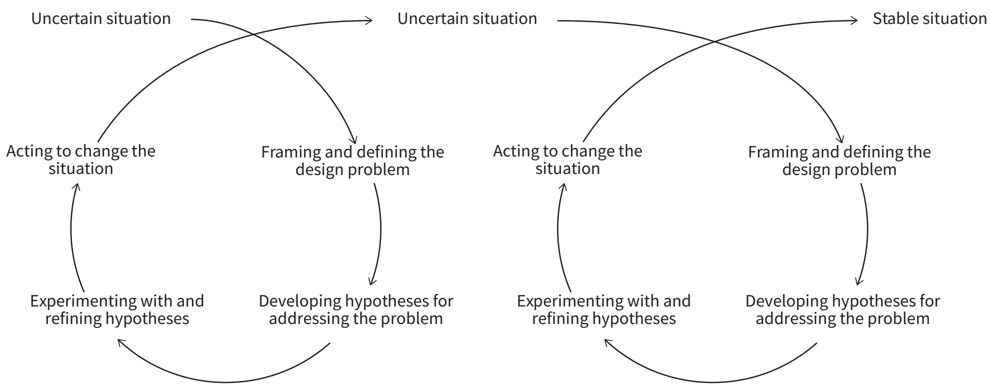


Figure 2.2. Dalsgaard's (2017) model of iterative designerly inquiry.

Such understanding of design inquiry bears resemblance with the transitioning moves between ill-defined and well-defined understanding of problem spaces as presented by Maher et al. (1996), which was further expanded to the co-evolution of problem and design spaces by Dorst and Cross (2001). Figure 2.3 summarizes the shared underlying concepts of design inquiry as the transition moves between unknown and known states of a design situation.

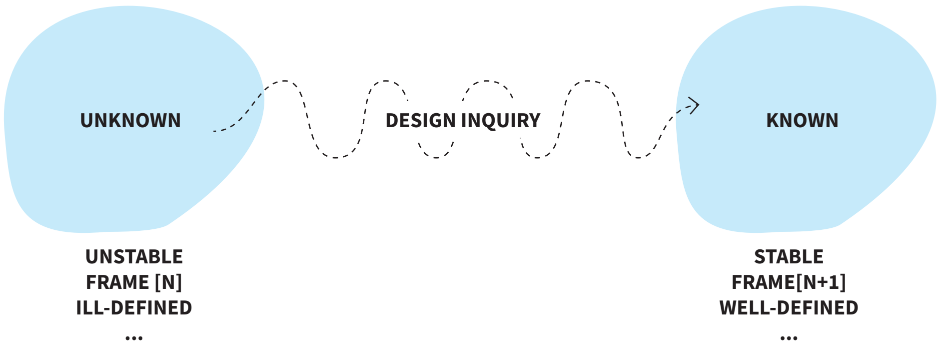


Figure 2.3. Design inquiry as a move between the unknown and the known (unstable/stable (Dalsgaard, 2017); ill-defined/well-defined (Maher et al., 1996); frame [n] / frame [n+1] (Dorst and Cross, 2001).

So far, in this chapter we have discussed the open-ended exploratory and opportunistic nature of design practice, which serves design inquiry of transitioning from an unknown state to a known state. We also highlighted that design inquiry is often performed through design methods and concluded that for the development of design methods, both mindset and step-by-step guidance needs to be addressed. Towards our goal to develop a conceptual framework, in the following of the chapter we will provide an interdisciplinary perspective on data and then zoom-in on the practice of data science. Afterwards, we will highlight the cross-sections of design and data practices and then conclude on the conceptual framework and the emerged research questions and methodological considerations.

2.2 Data and data science practices

Traditionally, data has been understood as quantified numbers from a sensing device, a definition widely used in engineering or science. However, in today's big data era, when data is complex, heterogeneous and ubiquitous, and data has permeated everyday life, the traditional notion of data refers to 'numbers' is not sufficient anymore (Mayer-Schönberger & Cukier, 2013). The plurality of understandings informs the next section showing an interdisciplinary perspective on data. Afterwards, we elaborate on data science practices to show connection points to the design process later on.

DATA AS A BOUNDARY OBJECT

The rise of big data phenomenon established interdisciplinary interests (boyd & Crawford, 2012; Mayer-Schönberger & Cukier, 2013), growing the framing of big data beyond a purely technical viewpoint. Framings of data such as 'subjective and objective data', 'ethical data collection', or 'political and economic value of data' carry loaded meanings around data that might appear in the contexts of designers, yet such meanings around data require new vocabularies to discuss them. Kitchen describes this as data assemblages: *"Data [...] do not exist independently of the ideas, techniques, technologies, people and contexts that produce, process, manage, analyse and store them. Indeed they are organised and stored in databases and data infrastructures that* *form the core of complex sociotechnical assemblages."* (Kitchin, 2014a, p. 185). Kitchin (2014a, p. 24) argues for considering data as the center of different data assemblages, which "frame what is possible, desirable and expected of data". The data assemblage lens is foremost a reminder on acknowledging the various understandings around data in the context of design. A data assemblage resembles the concept of a 'boundary object' (Star & Griesemer, 1989). Boundary objects refer to an object's (abstract or tangible) equipped meaning, that is built by communicating about, or categorizing the object. In the context of interaction design and HCI, 'data-as-boundary-object' has been especially prominent to problematize around data practices, such as data collection and sharing (Vertesi & Dourish, 2011), personal data management (Crabtree & Mortier, 2015; Mortier et al., 2014), interacting with local data (Taylor et al., 2015), or algorithmic model development (Passi & Jackson, 2017). These examples illustrate the contextual considerations designers need to take into account when dealing with data. However, such investigations provide little account to help how to incorporate data into design work. Despite the large variety of such examples illustrating the contextual considerations of how designers need to take data into account, only a few papers, like Feinberg's (2017), build on a data-as-boundary-object lens applicable to designing. As Feinberg puts it: "design projects in HCI can omit the work performed on data, making it seem as if data were a stable material to be 'used'" (Feinberg, 2017). In other words, although there is HCI work out there that has centered around data, we know little about the practical data practices that take place in the design process. To conclude, an **interdisciplinary understanding of data** that goes **beyond 'numbers'** is necessary to observe and approach the large variety of data practices in the design process embedded in socio-technical contexts, both in terms of what a dataset contains, and in terms of considering the contexts and practices around data.

The following section will elaborate on data science practices to open up the nature of practical data practices. Parallel to Section 2.1, we start with a high-level overview of data science, and then zoom-in to data analyst and data exploration practices.

DATA SCIENCE

Throughout the history of computer science, data practices have evolved continuously as computation power and storage became cheap and wide-spread. In the past decades, significant advances have taken place in knowledge discovery and data mining (Fayyad et al., 1996), information seeking (Marchionini, 2006) or information visualization (Card et al., 1999). Such sub-fields of computer science have produced their own conferences, methodologies, education, or curriculums, forming a mature body of knowledge. The 'big data era' (Mayer-Schönberger & Cukier, 2013) and datafication trend across industries (Lycett, 2013) introduced new emergences of how data is utilized in business. Using the earlier mentioned discoveries in subfields of computer science, new kinds of data practices have emerged to utilize data to gain new inferences and epistemologies (Kitchin, 2014b). A clean outcome of such emergence is the field of 'data science' solidifying in the past years (Cao, 2017). Data science unifies emerging practices, data techniques, and know-how in the big data era. The unifying characteristic of data science can also be seen in the light of its demographics, observing how people of varied backgrounds 'convert' to data scientist roles, from biologists to software engineers, and statisticians (O'Neil, C., & Schutt, R., 2013). Despite the massive tractions of data science recently, commonly accepted definitions are still lacking. Some industry roles focus on data analysis or statistical inferences, while others fuel machine learning models. In the following, from all the varieties and 'flavors' we focus on data science that is concerned with the dissertation's goals, namely inquiry and exploration. Furthermore, we also elaborate on the non-expert practices that may have relevance for designers as data non-experts.

DATA SCIENCE PROCESS

Similar to the design process, different characteristics of data science can be highlighted with different models. Baumer (2015) elaborates on a data science process with a focus of inquiry. In Baumer's (2015) holistic approach of teaching data science at an undergraduate level, the core of the process is to address questions through the acquisition, transformation and methodical exploration of data to reach answers in the form of inferences and representations (see Figure 2.4).

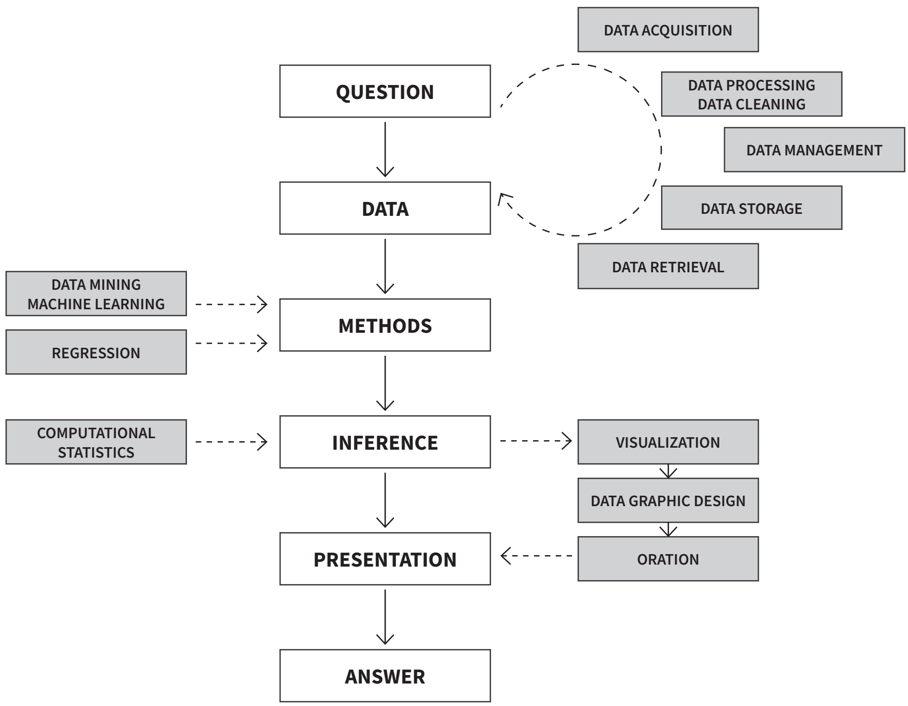


Figure 2.4. The data science process of an undergraduate level course, from Baumer (2015, p. 336). The process starts from a question that is answered through a variety of data practices.

The industry-based model of O'Neil and Schutt (2013, p. 41) in Figure 2.5 illustrates the variety of steps to handle data embedded into an otherwise cognitive process (i.e., defining a question and inferring answers). Although the process model of Baumer comes from educating undergraduate students that are not specialized in data and statistics, O'Neil and Schutt's process highlights an 'ideal' scenario that seems to apply to a variety of domains, and in essence similar to the visualized data science process in Figure 2.5. O'Neil and Schutt (2013) highlight a process where raw data from real-world phenomena is the starting point of steps to clean the data, and then explore it, model it, implement in a data product, or use the inferences for decision-making.

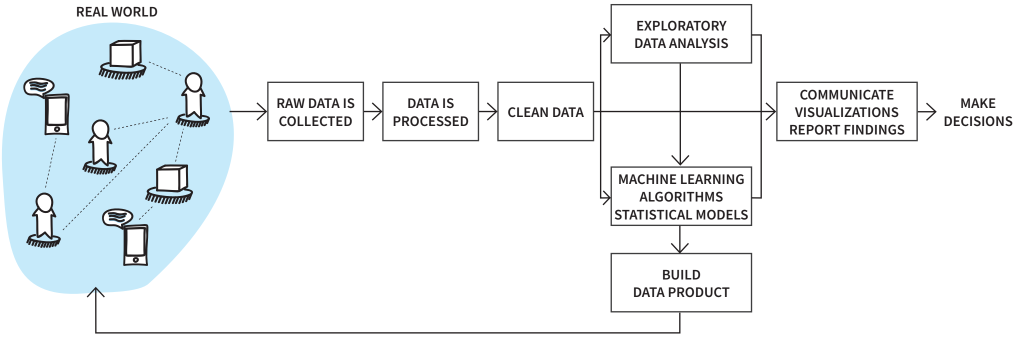


Figure 2.5. An industry-based data science process, based on O'Neil & Schutt (2013, p. 41). The process starts from the real world, and ends in decision-making or an intervention in the real world with a data product.

The comparison of the two presented models highlights that steps performed on the data are similar, both for an inquiry-focused process (i.e., Baumer's case) as well as for a process focused on data products or decision-making. One key milestone in these processes is to arrive at the point of having a dataset. Once there is a dataset, the nature of steps changes. In the following, we will focus on the methods to apply on the dataset (in the phrasing of Baumer) or the exploratory data analysis, as discussed by O'Neil and Schutt.

In 1962, Tukey introduced the term Exploratory Data Analysis (Tukey 1962, 1977) to refer to the use of statistical tools to describe and explore numerical datasets to make inferences from the data. Since then, EDA has taken a more expansive meaning and now includes a broad array of approaches and methods for the exploration of data. Alspaugh et al. (2019) elaborate on a more contemporary view on emerging data exploration strategies. They define data exploration as an *"open-ended information analysis, which does not require a precisely stated goal"*. Alspaugh and colleagues have considered exploratory data analysis on a spectrum between exploratory and directed analysis, with the following description what they see as exploration: *"Exploration is opportunistic; actions are driven in reaction to data, in a bottom-up fashion, often guided by high-level concerns and motivated* *by knowledge of the domain or problem space"* (Alspaugh et al., 2019, p. 22). The characteristics of exploratory and opportunistic resemble the design practice, as discussed in Section 2.1, highlighting an opportunity to consider matching data science practice with design practice. Alspaugh et al.'s (2019) model in Figure 2.6 describes the process of starting discovering a dataset to wrangle and profiling the data and then explore and report it. They unpacked exploration by highlighting the iterative nature of opportunistic looking, which can be interpreted as 'finding something interesting' as well as stressing the process of inquiring the data. To conclude, data exploration can be seen as an opportunistic practice where questions leading the exploration are continuously being generated, resembling the abductive nature of design practice as well.

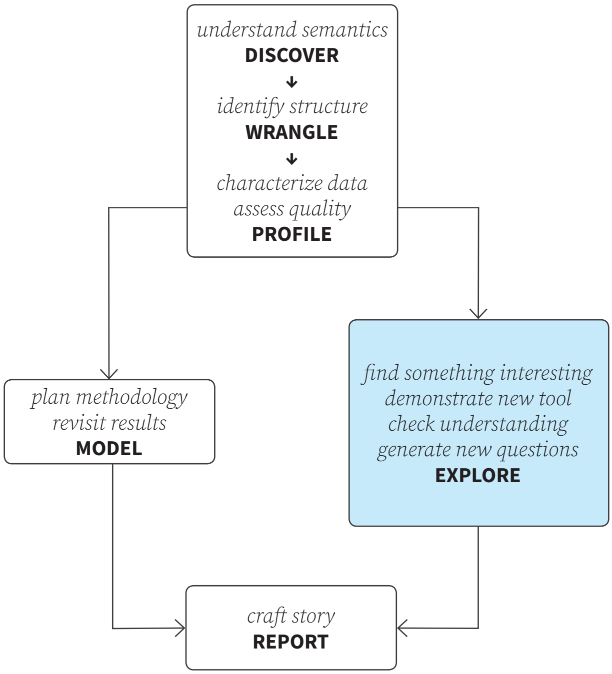


Figure 2.6. The data exploration model by Alspaugh et al. (2019, p. 23). Data exploration is a set of iterative steps between technical processing of the data model and crafting a story of the findings.

So far, in this section, we discussed data expert practices. In the next section, we present the non-expert perspective.

NON-EXPERT PERSPECTIVE

The different models above are based on empirical studies of data experts, such as data analysts or data scientists. However, for the research aim of how designers can utilize data in their design practice, the non-expert perspective can be seen especially valuable, assuming the limited data expertise of designers. Baumer's model in Figure 2.4 describes a process of teaching data competencies in undergraduate education, emphasizing to teach a whole spectrum of tools to prepare students working with data in real settings. The core of Baumer's inquiry process starts with asking a question and ends with an answer gained from data, and then communicate the findings. The tactic of using data for the whole inquiry helped students to learn how to 'think with data'. Outside traditional curriculums, Hill and colleagues (2017) explored teaching data science as a way of 'democratizing data science' for community empowerment. Their approach has been based on teaching basic programming to remain as closest possible to expert data science practices. They particularly emphasized to be able to ask questions that will be investigated from data, and in this process, be able to acquire data from online sources (such as capturing data Wikipedia), then analyze it and develop a visualization to communicate the findings.

While the approach of Hill and colleagues (2017) provided a flexible set of skills and tools, it also came with a price of a steep learning curve. D'Ignazio and Bhargava (2016) approached this problem from a more learning-centered angle. They created a set of learning tools for data literacy, designed to avoid programming explicitly, and targeted data skill acquisition through tailored, single-purpose data tools – DataBasic (D'Ignazio and Bhargava, 2016). These tools can be used with actual datasets and for actual visualization and analysis work, but they are primarily designed to be learning tools, scaffolding more complicated data operations. In another work, D'Ignazio (2017) added to the work on DataBasic tools from her experiences with applying and teaching data literacy positioned in creative work, such as design. Both the programmatic way of Hill et al. (2017) and D'Ignazio's and

Bhargava's (2016) learning tools approached data practices through the use of a set of tools, rather than a focus on one single tool. Such use of smaller tools to perform the different elements of a data workflow is a standard best practice, with roots in software engineering. Non-expert data communities, such as data journalists or digital humanities scholars, have summarized their know-how publicly as methods or tool libraries (e.g., (Digital Methods Initiative – Tools Database, 2020; *School of Data*, 2016; Gray et al., 2012). Data journalists are especially relevant community for designers, as their goals of inquiry can be similarly rich to find 'interesting stories' in a dataset, sometimes lead by the data, instead of a prescribed agenda and question. Data journalists' and digital humanities scholars' tool libraries are curated non-expert tools that support professional data practices that go beyond spreadsheet software (e.g., Excel) without needing to program. In conclusion, non-expert practices can make data science practices accessible to designers without advanced programming skills, and suitable for the opportunistic inquiry that describes design.

In the previous sections, we first elaborated upon the opportunistic and exploratory characteristics of designing and then zoomed-in on design inquiry as the transitioning move between unknown and known states. Furthermore, we presented data from an interdisciplinary perspective. Next, we zoomed-in on data science practices, where we highlighted how data exploration shares opportunistic characteristics with designing. In the next section, we will look deeper at the intersection of design and data, first elaborating on practices in literature how these two fields have been combined and then zoom-in on using data approaches specifically for design inquiry.

2.3 Data in design

In design, 'data' has gained popularity in the last decade with work such as personal informatics (Li et al., 2010), using visualizations as part of co-design (Dove, 2015) or data physicalization (Jansen et al., 2015). Such examples illustrate how data became part of the designed experience of the devices people interact with. For instance, personal informatics capitalized on the growing possibilities of wearables and self-tracking technologies, which were capable of measuring various aspects of the life of their users. Designers have turned their attention towards personal data and designing the user experience of interacting with personal data, as well as what longitudinal tracking of personal data enabled. Dove and Jones (2014) show how designers can bring data representations, such as visualizations, into co-design processes, using the visualizations as a boundary object to talk about people's personal experiences.

Next, we will discuss lenses that design researchers and practitioners have established to inform ways how to combine design and data practices.

LENSES OF DATA AND DESIGN

In this section, we discuss three lenses that attempt to structure the connection between data and design, a theoretical lens by Speed and Oberlander (2016), an industry practice-based one by King, Churchill, and Tan (2017), and one from architecture by Deutsch (2015). We introduce these three lenses as there seems to be no single wide-spread nor commonly-agreed mapping between design and data. Moreover, these three selected lenses illustrate the plurality of views implied around data and design. Speed and Oberlander (2016) present a theoretical lens to categorize different uses of data, primarily focused on how to utilize data-collecting artifacts in the design process. Their lens uses the Latin 'ablative' case to distinguish between designing from/with/by data, and illustrate design research case studies of each combination. Designing *from* data is the use of data as a way to base design decisions in the design process from collected data, such as measured features of people, artifacts, and contexts. Designing *with* data is the use of data when data is an essential part of the designed 'form', such as real-time data streams of an internet-of-things artifact. Last, designing *by* data refers to the use of data when a system designs with algorithms, for example, by parametric or generative algorithms, which are used in architecture.

The 'Designing with Data' book by King, Churchill, and Tan (2017) is especially focused on digital products and services, such as online platforms. Their book's subtitle, "Improving the user experience with A/B testing", suggests the use of data for optimizing an existing design and not informing the research phase of the design process. Nevertheless, they categorize three ways to think about data, data-driven, data-informed, and data-aware (see Table 2.1). As they put, *data-driven* design refers to a practice where collected data determines design decisions. *Data-informed* design refers to a practice where a design team uses data as one input in their decision-making process. In their description, data-informed maybe in situations when the problem space is not fully explored yet, and further research iterations are needed. Their third term, *data-aware* design, refers to the case when designerly practices are not only led by data but also used in data collection practices. In such a case, designers and data scientists need to collaborate with developers and business stakeholders to develop ways to collect data that focuses on answering the right questions. As King, Churchill and Tan (2017) describe, a design practice starts as data-aware and continuously goes towards data-informed and then data-driven; in parallel, the problem space becomes narrower, and the design decisions become less about what problem to solve, but optimizing the solution. The dissertation's perspective of design inquiry is exploratory and opportunistic, which suits most the data-aware design of their terminology.

Table 2.1. The categories of data and design by King, Churchill and Tan (2017).

Type	Definition
Data-driven	Collected data determines design decisions.
Data-informed	Data is used as one input in decision-making (among many).
Data-aware	Designerly practices are not only led by data, but also used in data collection practices.

Deutsch (2015) describes a spectrum from data-enabled, data-informed, data-driven practices as data-centric practices in architectural design (see Table 2.2). While architectural design is not within the scope of this thesis, nevertheless, it is the field that has perhaps the longest traditions in using data in the design process, and therefore valuable to look at. In the three practices described by Deutsch, 'data' refers to measurement data of various components in architectural design. He describes *data-enabled* as being aware of the data but not necessarily leveraging it, because decisions are also based on other subjectivities such as emotions, or organizational values (which resonates with a human-centered design mindset). As described by Deutsch, data-driven is when data is the primary priority, and the architectural practice invests in making the majority of design decisions based on metrics, leaving little space for intuition. In-between the two extremes, data-informed refers to using data as a factor in the decision-making process, but not the only one, or only for certain aspects of the design. Overall, we find these definitions limiting for design inquiry, as we explicitly strive for a richer perspective on data than only 'metrics', which is the dominant type of data in architectural design as shown by Deutsch (2015).

Type	**Definition**
Data-enabled	Being aware of data but not necessarily leveraging it (decisions are also based on other factors).
Data-informed	Using data as a factor in decision-making, but not the only one.
Data-driven	Data is the primary priority; metrics are invested in; little space for intuition.

Table 2.2. Data-centric practices in architecture by Deutsch (2015).

While these lenses illustrate the various ways how to combine data into design, especially regarding an approach or mindset for making design decisions with data, they have two limitations. First, they consider primarily numeric data coming from sensors or logging, which does not cover all types of data used by designers. Second, they naturally assume the involvement of data scientists in the design team, a resource that not every design team can afford. With these limitations in mind, King et al.'s (2017) *data-aware* design and data practice can describe the open-ended exploration and opportunistic design practice for design inquiry used in the dissertation. Although the lenses are helpful to categorize and describe certain design and data practices, they do not provide further guidance on *how* to combine these practices. The next section, therefore, will zoom-in on more practical examples of how data practices have been used for design inquiry to access insights otherwise hidden through qualitative methods.

DATA FOR DESIGN INQUIRY

As we discussed earlier in the chapter, design inquiry can be seen as an open-ended and opportunistic practice to gain a better understanding of a design situation using the designer's repertoire of knowledge and competences. Recently in design research, several works have explored new ways of using data as a way of inquiry. In the remainder of this section, examples are shown that draw on data as a resource to gain a better contextual understanding of the users' lifeworlds, as well as to build on the introduced data collecting prototypes not only as sensor devices but touchpoints in larger artifact ecosystems. Bogers et al's (2016) data-collecting technology probes have shown novel ways to gain rich and contextual data using sensors. In their follow-up work, their approach has expended out to probes, toolkits, and prototyping (Bogers et al., 2018; van Kollenburg et al., 2018). Their approach, coined data-enabled design, is a combination of design methodology and technical system that enables design explorations situated in real contexts. Their prototypes can react with the users to test different value propositions using real-time data collection. In their described projects, they combine sensor data with qualitative data collected via traditional methods, such as interviews. Giaccardi et al. (2016) have taken a different direction inspired by anthropology. By equipping everyday objects with a camera – a device capable of rich data collection – they used the collected data to feed into ethnographic inquiry. Their approach combines the sense-making of sensor data and qualitative data. In a follow-up project, Giaccardi (2019) has expanded on this approach using data of sensor-equipped objects and data-as-fact as patterns to inform machine learning models.

While these examples illustrate what near-future possibilities exist for bringing in customized technology into the design team, the resources they require are often beyond what is available for a design team. This problem is overcome by integrating data scientist collaborators into the design team (Dove et al., 2017; Yang et al., 2018); however, such an approach is not necessarily viable for designers outside the technology industry. Furthermore, the examples above heavily utilize customized software and hardware solutions, which limit their utility for broader designer populations.

The previous sections discussed the fields of design and data, as well as the intersections of these two fields, while focusing on inquiry. Next, we develop these findings into a conceptual framework guiding the empirical studies elaborated in Chapters 3 to 6.

2.4 Conceptual framework

The previous parts of this chapter explored the intersection of data and design, highlighting key insights such as the necessity to define data broader than merely ‘numbers’, the open-ended exploration and opportunistic practice of design inquiry, and design inquiry’s resemblance to data exploration. Based on the key insights, the last part of this chapter develops a conceptual framework to set the focus for the empirical studies, for which we discuss the corresponding methodological approach in more detail last.

In Section 2.1, we concluded on design inquiry as the transitioning between an ‘unknown’ state and a ‘known’ state and an open-ended exploratory and opportunistic process to navigate in-between these states. Figure 2.7 shows the combination of these elements, and allows us to focus our investigation on the middle transitioning phase of design inquiry and explore how design inquiry can be approached **through data**.

To elaborate on what we mean by ‘through data’, we revisit the highlights of Section 2.2. First, in Section 2.2, we argued for an interdisciplinary understanding of data that goes beyond numbers. Second, we argued that non-expert data practices might be accessible for designers as well to use data practices. Such non-expert data practices could serve data exploration needs, which resemble the open-ended exploratory and opportunistic process of design inquiry. Therefore, what we mean with design inquiry through data is a type of design inquiry, where the inquiry process utilizes heterogeneous and complex data dispersed in the domain of design inquiry, and such data is leveraged upon through data practices. In such an inquiry, data generates value for the designer towards a better-defined problem state, or an improved frame. Such value can be implicit, like an improved understanding of the problem domain, or explicit, like a visualization elaborating a phenomenon.

Figure 2.7. Design Inquiry Through Data, where data practices are used to leverage data to fuel design inquiry.

In the next sections, we expand further by decomposing design inquiry through data as elements of 'mindset', 'process', and 'tools' to make design inquiry through data operational.

MINDSET OF DESIGN INQUIRY THROUGH DATA

The data practices and design inquiry are conducted by a designer, with her own sensemaking processes, informed by the broader design situation, as well as the existing explicit or tacit knowledge of the designers (Kolko, 2009). Daalhuizen (2014) describes these cognitive aspects as the 'method mindset', composed by the theoretical knowledge about a method or knowledge about the method's use and belief in added value, trust in applicability, and preference for using a method. In other words, a large part of the mindset is informed by a learning curve (how well a designer knows a method and its encapsulated theory), but also more situated knowledge informed by the design situation itself and knowing what method to use for a specific reason.

We use *mindset* to make the cognitive part of design inquiry through data operational, where the sensemaking process can be described, focusing on the thinking patterns of designers. A designer's thinking about data exploration is unlikely the same as a data scientist's thinking, and the lens of mindset allows the observation of nuances in this regard.

PROCESS OF DESIGN INQUIRY THROUGH DATA

The variety of the introduced data science models share a strong process perspective, detailing how different steps can be identified with their own goals and practices. The design process is similarly formalized through iterative steps that continuously arrive at a well-defined solution. Dorst and Cross (2001) illustrate the design process with the co-evolution model, where the ill-defined problems are continuously being framed as the understanding of the problem develops. Such continuous unfolding of the framing of the problem to solve and the right solution for it leads to ways to describe the process without the clear 'boxed' process models data science has. Nevertheless, by focusing on *process* as an element of the conceptual framework, the different process understandings between design and data science can be understood better.

We use *process* to make the procedural part of design inquiry through data operational, where different steps are taken, and the decisions to take different steps can be described while transitioning from an 'unknown' situation to a 'known' situation. We not only focus on the continuous unfolding of the problem and design space in a design situation, but also the process of data science with its different steps and practices.

TOOLS OF DESIGN INQUIRY THROUGH DATA

Tools are integral parts of both the repertoire of designers and data scientists. Without software tools, it would be impossible to leverage the computational aspects of data. Tools are integral parts of a designers' repertoire that can take forms such as design methods, physical tools such as pen and paper, or software tools (e.g., for sketching or prototyping). As discussed earlier, non-expert data practices seem suitable for designers, and a primary way of how non-experts approach data practices is through the use of tools tailored for their needs.

To conclude, we highlight *tools* to make the hands-on part of design inquiry through data operational, where the taken data and design actions can be described, focused on tools to understand the requirements designers have for them, and the way how tools are performed with.

The three intertwined methodological elements of *mindset*, *process*, and *tools* are chosen to make design inquiry through data operational. In the next section, we will combine these three elements into a conceptual framework and set up a series of empirical studies. The gained insights will be used to inform the development of the methodological contributions as stated in the goals of the dissertation.

CONCEPTUAL FRAMEWORK

Figure 2.8 shows our conceptual framework and illustrates how *mindset*, the *process*, and *tools* can be embedded into the design inquiry setup. The three elements decompose the transitioning from the unknown situation to the known situation. In it, the designer follows a *process* that sets the transitioning. The process is rationalized by a *mindset* that the designer follows as she makes sense of the design situation momentarily. The designer's thinking process is leveraged by the *tools* involved in order to interact with the data.

The conceptual framework depicted in Figure 2.8 has a double role. First, it is used as a lens for understanding data-rich design practices, and second, it is used to inform the design of corresponding methodology. In the next section, the two roles of the conceptual framework will be used for setting up a set of empirical studies that lead the investigation of the research questions.

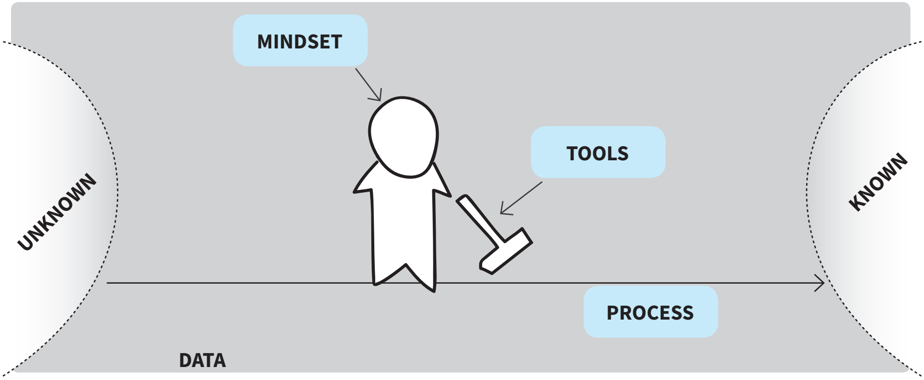


Figure 2.8. Conceptual framework for Design Inquiry Through Data. Process is the basis of design inquiry through data, where mindset represents the thinking and sensemaking of the designer, and tools represent the artifacts to leverage the process.

2.5 Setting up the empirical studies

The conceptual framework is used for setting up empirical studies for understanding data-rich design practices, and in the meanwhile, for developing methodological contributions for data-rich design practices as part of an RTD program. The research questions formulated in the first chapter of the dissertation are guiding this search for both theoretical and practical knowledge on the intersection of design and data science:

- **RQ2:** How do designers appropriate data science practices for design inquiry?
- **RQ3:** How can data science practices be characterized through a creative process lens?
- **RQ4:** How can a design method support design inquiry through data?
- **RQ5:** How do designers adopt Exploratory Data Inquiry in design practice?

Chapter 3 starts with the first study investigating the appropriation of non-expert data practices for design inquiry, focusing on **RQ2**. The second study in Chapter 4 focuses on the framing of data practices through a creative lens, answering **RQ3.** The next two research questions are directly concerned with the design goal of developing methodological contributions to support data-rich design practices. First, in Chapter 5, a design method is developed and evaluated to answer **RQ4**. Then, in Chapter 6, the adoption of a methodology is investigated to answer **RQ5**. The summary of these studies and connected research questions can be found in Table 2.3.

Table 2.3. Overview of empirical studies and incorporated RTD activities. Numbering of studies refer to their related chapters.

Study	Setting	Research Question	Sub-research questions	Focus of RTD activity
3A (Master thesis records 1)	1-day master-level elective workshop	RQ2: How can non-expert data science practices be appropriated for design inquiry?	How are data practices appropriated for the design process when the starting point is a design brief and a dataset?	Early exploration of a future data-rich design practice
3B (Service Design Tourism)	3-days workshop embedded in master-level semester project		How are data practices appropriated when used as a complemen-tary method in design inquiry?	Eliciting the scope of the problem and design space
4ABC (Service Design Mobility)	3-days workshop embedded in master semester project	RQ3: How can data science practices be characterized through a creative process lens?	How does creativity manifest when using data practices in the design process? How to describe data practices in designerly terms to be useful for designers?	Extracting the Exploratory Data Inquiry methodology that will provide design ra-tionale for design methods and design tools.
5 (Master thesis records 2)	1-day master-level elective workshop	RQ4: How can a design method support design inquiry through data?	What data tools and techniques can support creative work with data? What is the nature of creativity support in the context of data exploration in design inquiry?	Developing a design meth-od based on Exploratory Data Inquiry methodology. Evaluation of the design method.
6AB (Frame Innovation + data exploration)	5-days workshop with master students	RQ5: How do designers adopt Exploratory Data Inquiry in design practice?	How can we combine a problem framing framework with design inquiry through data?	'Stress test' the Exploratory Data Inquiry methodology and investigating the mindset of use.

For the empirical studies, methodological considerations have been made to shape the data-rich design practices along with the RTD research approach. The research is focused on intervening in the design process, more specifically to a core component to doing design (i.e., design inquiry), hence it is essential to establish research setups where designing can be observed in required aspects. In more practical terms, design inquiry needs to be situated in actual design processes conducted in realistic scenarios. We address this need by setting up studies where 'complexity' is addressed in terms of the ecological validity of the design situation, but also by used datasets or scope of data and design activity. In our approach, we use studies that are contained in design workshops, and we address the learning curve involved, discussed in the next sections, respectively.

DESIGN WORKSHOPS AS SITES OF INQUIRY

The consequence of collecting data from situated and realistic scenarios is to focus on collecting rich insights from different cases. For observing realistic and situated design processes, our approach is to use design workshops as sites of research inquiry. The use of design workshops as the primary source of data collection and primary sites of research inquiry necessarily involves working qualitatively and with a smaller set of participants. Design workshops (in this case, 1-5 days long pressure cooker setups) enable conducting a reasonably long and complex design inquiry process with incorporating multiple iterations, attending a learning curve, and giving space for reflection. The workshops are also an environment to iterate on the future data-rich design practice we are aiming to support. These workshops are continuously iterated following a design process, on the one hand, to tailor the workshops to the specific cases, but also integrating the learnings that take place along the process. Reflecting on the design workshops on the lines of *mindset*, *process*, and *tools* will produce implications for design that are further synthesized in Chapter 7. Showing implications for design are a common way to generalize knowledge that is generated throughout designing (Dourish, 2006), and the designed outputs (i.e., design workshops) are also valuable for design practice as a way of encapsulating a new method of conducting design inquiry through data.

ADDRESSING THE LEARNING CURVE OF DATA TECHNIQUES

As the core of the research project is focused on intervening in the design practice by introducing techniques, approaches, know-how from another field (i.e., data science), the learning curve of such 'new' techniques and approaches need to be taken into account in the methodological choices. As designers are relatively inexperienced with hands-on working with data, yet well-informed about the potentials of data techniques, the learning curve should address hands-on skill development as well as removing black-boxing.

LIMITATIONS

The methodological choices create limitations in the approach. The main limitation is the designers' experience level. Experienced designers (such as designers with years in training) have professional knowledge about designing that designers in training (e.g., master-level design students) do not possess yet. First, a practical consideration is the difficulty of claiming professional designers' time to participate in studies lasting multiple days, so that sufficient time is allocated to the learning curve while the design task at hand is realistic. Second, the rich insights are to be gathered about how designing happens when using a new approach for design inquiry. As with experience, designers also become more 'rigid' in their choices, and develop strategies to follow; their experience as much becomes a hindrance as would contribute to the research goals. Vermaas (2016) challenges this notion by questioning whether only expert design practice should be accounted for in design method development.

The following of the thesis presents empirical studies tackling these research questions in Chapters 3, 4, 5, and 6, respectively. In Chapter 3, we focus on how designers appropriate data science practices. In Chapter 4, we conduct a study where we observe data science practices through a creative process lens, concluding the chapter with a design methodology. In Chapter 5, we develop a design method based on the design methodology from the previous chapter and evaluate it. In Chapter 6, we observe the adoption of the design methodology in design practice. Chapter 7 starts with synthesizing the different studies reported in Chapters 3–6, to reflect on the findings and extract the answers to the research questions.

Chapter 3

Designers Appropriating Data Practices

In the previous chapter, we unpacked the cross-section of design and data science practices and focused the research on design inquiry through data. The review of existing work indicates that design and data science practices can be constructively combined. However, it remains unclear how such a combined practice takes place. In this chapter, we address RQ2 *of the thesis, "How do designers appropriate data science practices for design inquiry?" In order to answer this question, we conduct two studies to analyze how designers appropriate non-expert data tools. One study focuses on the appropriation of data practices when the starting point is a dataset and a design brief, and the other one focuses on the appropriation of data practices when data practices are used as a complementary method in design inquiry. The results indicate that designers use their creative capacities in defining what data to acquire, and in appropriating non-expert data practices driven by designerly sensemaking.*

This chapter is based on:
Kun, P., Mulder, I., & Kortuem, G. (2018). Design Enquiry Through Data: Appropriating a Data Science Workflow for the Design Process. In *Proceedings of the 32nd International BCS Human Computer Interaction Conference (HCI 2018)*. BCS Learning and Development Ltd. https://doi.org/10.14236/ewic/HCI2018.32

3.1 Introduction

As has been concluded in Section 2.3, designers lack methodical guidance on how to bring data into design practice. The use of data practices by data journalists and digital humanities scholars indicates similar yet largely unexplored opportunities for incorporating data techniques into design. Understanding how designers can appropriate existing data practices will help us to gain an initial understanding of the boundaries of future data-rich design practices and to inform the following steps by revealing potential tensions or design opportunities.

In this chapter, we address **RQ2** of the thesis, *"How do designers appropriate non-expert data science practices for design inquiry?"* To answer this question, we set up an exploratory study in which we have conducted two design workshops, study 3A and 3B. More specifically, we observe how designers appropriate data and data practices in two conditions: 1) from data to problem space; and 2) from problem space to data. These two empirical design workshops are tightly controlled experiments where design context and data are given. In the following, we further detail the two studies we conducted with master-level design students using existing non-expert data tools for design inquiry with data. After presenting the studies and the results, we evaluate the impact of such an approach on the design process. Our results indicate that existing non-expert data tools can be incorporated into design inquiry, and designers can use their creative capacities in hypothesis forming of data collection and data exploration of digital data.

3.2 Research approach

The review of data science practices in Section 2.2 indicated two contexts where data science practices could be appropriated for the design process; one where the starting point is a dataset for exploration, and a second one, where the starting point is defining a phenomenon to collect data about. These contexts bear similarities with each other, but have underlying assumptions about the role of datasets. The first context reflects on the ubiquity of existing datasets, such as accessing datasets from open data portals. It is unclear how designers could work from a provided dataset in finding the right problem to solve. Finding the right problem requires analytical work, such as data exploration of a public or received dataset, to extract value from it. The second context hints to the increasingly effortless ways to collect and store data. In this context, designers use data science practices to augment their research process in capturing and analyzing data to answer inquiries. In this case, designers can use the capturing and analysis of digital data to complement qualitative techniques, such as interview studies or ethnography, to gain additional insights from the data.

These two contexts informed the setup of two exploratory studies:

1. Study 3A - master thesis records: Designers analyzing a **provided dataset** to identify a problem space for design concepts;

2. Study 3B - tourism: Designers with identified problem space to **capture datasets**, and analyze them.

Both of the studies were limited to the conceptual phase (i.e., ‘fuzzy front end’) of the design process (Sanders & Stappers, 2008), focusing on using data in gaining an understanding of the world in order to inform design concepts. To be able to better navigate through data science practices, we adapted the process described by Baumer (2015) to match a generic design process. Baumer's process as shown in Figure 3.1 resembles the scientific inquiry process by starting with defining a question to investigate. In a next step to answer this question, data is collected. It is most likely necessary to transform and clean this data to prepare it for exploration. When a dataset is available for exploration, different analytical methods are applied on it, such as statistical analysis or visual analysis. The exploration generates an inference, such as insights. These insights can then contribute to the designer’s understanding of the problem or can be further communicated as visualizations, reports, or design concepts.

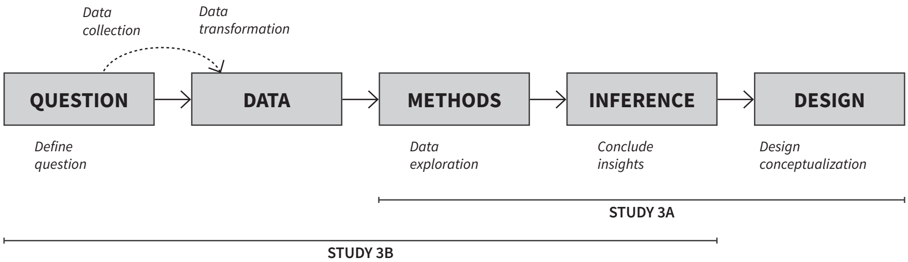


Figure 3.1. Schematic of a generic data and design process for Study 3AB, which also served as the basis of the conducted studies. The scope of the two studies is indicated in relation to the process.

Both exploratory studies aimed to answer the following sub-research questions:

› What are the conditions that enable data science practices to be integrated into a design process?

› Can non-expert data tools support designerly work?

› How do the design process and design reasoning change when using digital data?

More specifically, for Study 3A, our research objective was to explore how master-level designers – inexperienced with data practices – use an unknown dataset for a specific design goal, using non-expert data tools without a prior tutorial. For Study 3B, our research objective was to see how novice (i.e., master-level) designers – inexperienced with data – appropriate data science practices in design inquiry, using non-expert data tools after trying them through homework prior to the study. Table 3.1. shows the setup of the studies, which are elaborated below. Both studies were promoted as learning workshops to teach designers data competencies and tools, by hands-on working on a design problem with data. The studies were similar in several aspects.

Following D'Ignazio's (2017) guidelines on using familiar datasets, Study 3A featured a dataset relatable for the participants' personal experience, while the participants of Study 3B collected data themselves in a problem space established prior the study. We provided open-source or freely available non-expert data tools, as shown in Table 3.1. Throughout the studies, the participants worked

Table 3.1. The setup and methodology overview of Study 3A and Study 3B.

	Study 3A (master thesis records)	**Study 3B (tourism)**
Research questions	How are data science practices appropriated for the design process when the starting point is a design brief and a dataset?	How are data science practices appropriated when used as complementary methods for designerly inquiry?
Setting	One-day elective class.	Three consecutive days workshop, part of a semester-long project.
Partici-pants	First year master design students (n=20, 13 females, 7 males) from Service Design, Interaction Design and Product Design. Participants worked in pairs.	First year master design students (n=26, 20 females, 6 males) from Service Design. Participants worked in groups of 4-5.
Materials		
Dataset	1884 master thesei records with complete metadata from Delft University of Technology. Scraping and moderate cleaning was done by us.	No provided dataset (the participants captured data as part of the study).
Software tools	Microsoft Excel, Google Sheets, RAWGraphs, OpenRefine, Google Fusion Tables	WebScraper, Microsoft Excel, Google Sheets, RAWGraphs, OpenRefine, Carto
Design tools	Worksheets for Activity 1: dataset column titles, process reflection sheet. Worksheets for Activity 2: Data design canvasdesign reflection sheet.	No additional design tools provided.
Procedure		
Pre-study task	No pre-study task.	Homework a week before the study on scraping a page (with WebScraper), and to extract one insight from the Titanic dataset with RAWGraphs.
Study	Basic introduction to data processing and tools. Activity 1 (Data exploration): Processing the provided dataset and analyzing it towards concluding 3 insights and make a presentation. Activity 2 (Conceptualization): Based on one insight from Activity 1, generate a data-inspired design concept and make a presentation.	Basic introduction to data processing and tools and debriefing the pre-study task. Activity 1 (Question definition): Related to the semester project, defining three research questions to be answered with data. Activity 2 (Data collection): Capture data (by scraping or downloading) for the questions from Activity 1. Activity 3 (Data transformation): clean, prepare, transform the captured data from Activity 2. Activity 4 (Data exploration): Make sense of the dataset from Activity 3 by analysis or visualization. Conclude on three main insights gained. Iterate from Activity 1, if necessary. Prepare a presentation about the process and the insights.
Follow up	Post-study questionnaire (fill rate: 75%) about learning goals, individual reflections and impact of the learning on participants' future work.	Post-study questionnaire (fill rate: 50%) about learning goals, individual reflections and impact of the learning on participants' future work.
Research data	Content analysis of participants' worksheets and presentations from Activity 1 and 2, post-study survey and observations.	Content analysis of presentations from Activity 4, ethnographic field notes throughout the study, post-study survey and observations.

towards tangible outcomes (as insights and concepts) captured during mid-term and final presentations. In the following section, we present the two studies and their respective results in detail.

3.3 Study 3A (Master Thesis Records)

With Study 3A, we aimed to observe how data science practices are appropriated for the design process, with the conditions of novice designers – master-level design students – facing an unknown dataset without prior experience in data. Part of their education, design students had previous coursework on basic statistics, programming, and design research methods. This background led us to assume that design students with bachelor's degrees from technical universities would have some tacit data knowledge that can inform their approaches. Furthermore, we assumed a basic level of familiarity with spreadsheets software (e.g., Excel), and familiarity with typical visualization techniques (e.g., charts, graphs).

PARTICIPANTS AND SETUP: Twenty students (13 females, 7 males) participated in the current study, as a one-day elective class. The students were first-year master students of Delft University of Technology in product, interaction and service design orientations of design. Participants worked in pairs during the study.

MATERIALS: The participant pairs were provided with a dataset, several software tools, and worksheets to use. The dataset was a complete database of all master's thesis records of the internal repository of Faculty of Industrial Design Engineering at Delft University of Technology. At the time of the study, the dataset contained 1884 rows and 28 columns of various metadata, including the theses' Title, Abstract, Mentors, Keywords, etc.

Additional worksheets supported the participants' processes; these worksheets were used for collecting research data, but also to guide the process for the participants. Activity 1 (Exploration) was supported with a printout of the column titles of the dataset if the participants wanted to take notes about it. Activity 2 (Conceptualization) was supported with a data design canvas worksheet, containing guiding questions for the process: Data (‘What are the available data?’), Model (‘How will it work?’), Experience (‘How will it look like?; What will it do?’), Problem, Added value. The reflection sheets contained an empty timeline to visualize and describe the process of the participant pairs (an example of the filled reflection is shown in Figure 3.2).

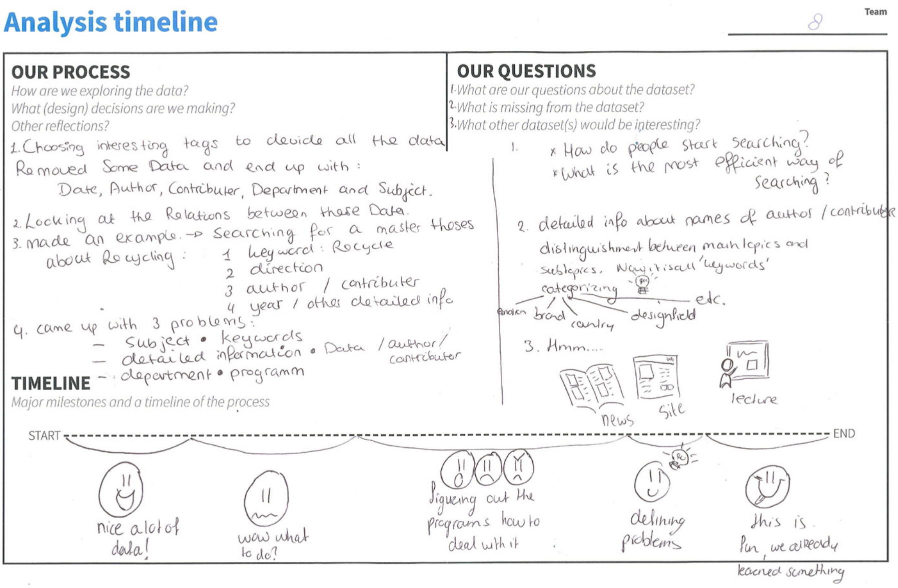


Figure 3.2. Worksheet from Study 3A, a participant pair reflecting on their data analysis process.

PROCEDURE: The study started with a basic introduction to data processing and the provided tools (the participants worked on their own computers) and the different worksheets. The first half of the study was Activity 1 (Analysis); the pairs received a task to explore the provided dataset and extract three main insights that they found as design problems to solve. The pairs could use the provided additional worksheet (dataset column titles), but it was not compulsory. They received minimal guidance on how to open the dataset in spreadsheet tools and to do basic data cleaning and transformations in OpenRefine. For the visual analysis of the dataset, participants were provided basic guidance to use RAWGraphs (Mauri et al., 2017), and Google Fusion

Tables. At the end of the activity, participants needed to fill up the process reflection sheet. The second half of the study was Activity 2 (Conceptualization); the pairs received the task to develop a design concept based on one selected insight from their output of Activity 1 (Exploration). The procedure was based on the process from Figure 3.1, and Table 3.1 provides an overview of the operationalization during the study.

DATA COLLECTION AND ANALYSIS: Throughout the study, observations, notes, and photos were captured. For both activities, we provided worksheets to capture the participants' self-reflections on their process (see Materials). Both activities were concluded with the participants preparing a short visual presentation with three insights and a design concept, respectively. Following the study, we analyzed the presentation materials, self-reflection worksheets, and observations to identify patterns, similarities, and differences. The study was also followed by a questionnaire sent to the participants to collect immediate data about their learnings and reflections on the impact of the workshop on their future work.

RESULTS

EXAMPLE PROJECT: We first present one concept generated by one participant pair throughout the study to illustrate the kind of complexity and novelty achieved by a one-day setup. The dataset contained 1884 records of different master thesis entries. All of these thesis entries had multiple keywords (such as: 'design', 'sustainability' or ,'Internet of Things'). The average keyword count per thesis was 4.50 (SD=2.34, min=1, max=29). Similar to the keywords, all thesis entries had multiple mentors (mostly faculty members). The average mentor count per thesis was 2.32 (SD=0.67, min=1, max=6). Participant pair 3A.6 argued, that based on the characteristics of the keywords and mentors, it is possible to explore the most common keywords for a given mentor, and vice-versa, which mentors are most common for given keywords (i.e., keywords and mentors formed a bipartite graph). Following this insight, this participant pair presented a concept to find the perfect mentor based on keyword interests (see Figure 3.3). Table 3.2 highlights how different data attributes were used for similar and different concepts.

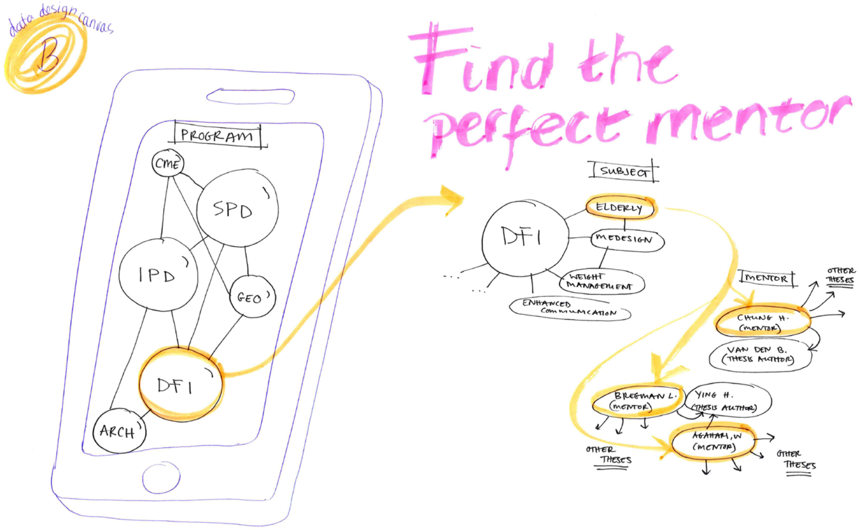


Figure 3.3. A concept of finding the right mentor for a certain master thesis (Study 3A).

PROCESS: We observed participant pairs during Activity 1 (Data exploration) daunted by the initial task of taking a previously unknown dataset and extract valuable insights out of it. They performed this task without formal training in working with datasets, following a hands-on learning process. An example reporting of this process from 3A.4:

1. *"Start with repository and identify users and use-cases.*
2. *Looking at the dataset, trying to understand.*
3. *Trying out the tools: without any questions behind, just exploring.*
4. *Visualizing random columns [with RAWGraphs]*
5. *Seeing some patterns? [pointing back to point 2.]*
6. *We looked back at the dataset and started to ask ourselves some questions*
7. *Trying to simplify the dataset to our needs using OpenRefine"*

The participant pairs generally followed a similar process:

an unstructured, ad-hoc process of data analysis, where they continuously gained a better understanding of the dataset, learned the usage of the tools and got familiar with the techniques of working with data. The main used data techniques were cleaning the dataset to remove inconsistencies, such as character capitalizations or spelling errors, and transforming the dataset in various ways so it can be provided as input into the used tools. In the end, 9 out of 10 participant pairs succeeded in presenting three insights based on the dataset; one pair misunderstood the task.

Based on one insight from Activity 1 (Exploration), during Activity 2 (Conceptualization) each participant pair developed a design concept. Table 3.2 provides a detailed overview of the developed concepts. Most

Table 3.2. The design concepts generated during Study 3A, and the used data attributes informing the concepts.

Group	Concept description	Data properties used							
		Keywords	Mentors	Departments	Program	Title	Author	Date	Abstract
3A.1	Finding the right mentor for your graduation	x				x	x		x
3A.2	Network visualization of finding the right topic for your research	x	x						
3A.3	Tinder for finding the right thesis subject	x				x			
3A.4	Finding the right subject for your graduation	x	x	x					
3A.5	Personalized search based on user data								
3A.6	Finding the right mentor for your graduation		x	x	x				
3A.7	Finding the right mentor for your graduation	x	x	x					
3A.8	Finding the right subject for your graduation	x	x		x		x		
3A.9	Connect people around the same interests	x		x	x		x		
3A.10	Showing trends in graduations				x	x	x		x

participant pairs focused on a few data attributes from the dataset, namely the thesis title, abstract, graduation mentors, departments, and keywords. Three concepts focused on finding the right mentor and four concepts focused on finding the right graduation subject. One concept targeted improving the overall search experience, one concept aimed at connecting people with similar interests based on the subjects, and one concept focused on showing trends in graduation projects.

PARTICIPANT REFLECTIONS: In the post-study questionnaire, the majority of participants primarily valued learning about the tooling to work with data. In detail, they found learning about the generic workflow of working with data as something new. Furthermore, they found the provided non-expert tools approachable to integrate data into their design process.

The participants also reflected on the thinking process shift necessary to utilize data. As one participant phrased his main learning: *"Asking the right questions at the beginning of the data, what do we want to know, helps to understand what to look for."*(participant from 3A.4). Another participant phrased it differently: *"The importance of a research question or hypothesis for structuring and processing the data"* (participant from 3A.1).

3.4 Study 3B (Tourism)

With Study 3B, we aimed to see how the appropriation of data science practices while using non-expert tools could complement the design research process. We assumed that novice designers inexperienced with data would need to do multiple iterations of the activities to get comfortable with data capturing and analysis for designerly insights. Similarly to Study 3A, the design students had previous coursework on basic programming and quantitative and qualitative research methods for design. Therefore, we expected that the design students with backgrounds from technical universities have basic familiarity with spreadsheets software, and familiarity with typical visualization techniques (e.g., charts, graphs).

PARTICIPANTS AND SETUP: 26 students (20 females, 6 male) participated in the study, which ran for three consecutive days as a part of the participants' semester project. All students were first-year master students in service design from Aalborg University Copenhagen. During this study, participants worked in groups of 4-5.

MATERIALS: The provided software tools are summarized in Table 3.1.

PROCEDURE: Prior to the study, participants received a Pre-study task (with two sub-tasks) to get familiar with data capturing and data visualization for analysis. The participants were instructed to scrape a specified webpage (their university library's search results page), and to visually explore a sample dataset from RAWGraphs (Mauri et al., 2017) and extract three insights from it. The study started with a basic introduction to data practices and the provided tools (the participants worked on their own computers) and a debriefing of the pre-study task. The beginning of the study was Activity 1 (Question definition); the participant groups needed to define research questions based on their semester brief (which was focused on tourism in Copenhagen, Denmark). Activity 2 (Data collection) continued with the research questions from Activity 1 (Question definition), with the task to capture data in relation to the research questions. The task of Activity 3 (Data transformation) was to clean, prepare, and transform the captured dataset from Activity 2 (Data collection). The end of the study was Activity 4 (Data exploration), during which the participant groups needed to make sense of the dataset by analysis and visualization and prepare a presentation about their process and outcomes. The participant groups could iterate from Activity 1 (Question definition) to Activity 4 (Data exploration), if necessary. The procedure was based on the process from Figure 3.1, and an overview of how it was operationalized can be found in Table 3.1.

DATA COLLECTION AND ANALYSIS: Throughout the three days of the study, observations and photos were captured by the researchers.

Following the study, we processed the presentation materials, and the observations to identify patterns, similarities, and differences. The study was followed by a questionnaire sent to the participants to collect immediate data about the learning goals and self-reflection on the impact of the workshop on their future work.

RESULTS

EXAMPLE PROJECT: The problem space of this study was centered around tourism in Copenhagen (Denmark). For the current study, the participants were told to utilize a data science workflow to further their research about the problem space. In order to illustrate the kind of problems and what complexity the participant groups operated on, we first present the work of participant group 3B.2. This group focused on a specific neighborhood from the lenses of tourism. Their leading research questions were:

- Which places are recommended in [certain neighborhood]?
- Where do locals and visitors spend their time in [certain neighborhood]?
- What do people search about [certain city] abroad on Google?

For example, in their approach, participant group 3B.2 analyzed social media hashtags for a specific neighborhood and especially looked into the less common hashtags from slang and subcultures.

PROCESS: Prior to the study, the participants received two pre-study tasks as homework. The task to visually explore a dataset (to be done individually) was done by all participants, while the task of scraping a webpage (to be done as a group) was done by half of the groups. During the debriefing, the participants reported difficulty in extracting interesting findings from the sample dataset without background knowledge and knowing what would be interesting to know about this dataset.

The participant groups started with Activity 1 (Question definition): the groups first considered their project and defined initial research questions to be answered with data. Moving forward to Activity 2

(Data collection), the groups captured data from online resources, primarily by scraping and downloading existing datasets. Scraping was mainly daunting for participants without extensive programming skills; nevertheless, by the end, most participants managed to develop non-trivial scrapers, tackling pagination, and similarly complex problems. All scraping was done using browser extensions. The groups ended up capturing data about tourism, primarily by scraping publicly accessible data from social media (e.g., Twitter and Instagram) and tourism websites (TripAdvisor, etc.), as shown in Table 3.3. As the next step, the participant groups worked on Activity 3 (Data transformation). The main needs of data cleaning were to eliminate inconsistencies, hidden characters, and similar string operations. As a significant portion of the captured data was location-specific (e.g., addresses), some groups used OpenRefine to enrich their datasets with GPS coordinates. This was accomplished by following an OpenRefine recipe calling an external API with the address input to enrich the data with GPS coordinates. The participant groups finished the study with Activity 4 (Data exploration). The groups explored their dataset through visualizations in RAWGraphs and Carto.

Group	Problem area	Data sources	Tools used
3B.1	What are the places locals visit and how to provide local experiences to visitors?	Crowdsourced review sites (2), curated travel sites (1), social hospitality site (1)	WebScraper, OpenRefine, Google Sheets, RAW-Graphs
3B.2	Focused on a specific neighborhood, what are the recommended places and places of interest for locals and visitors?	Crowdsourced review sites (2), curated travel sites (2), social media (1), qualitative interviews	WebScraper, OpenRefine, Google Sheets, RAW-Graphs
3B.3	What places are recommend-ed by locals? How far visitors go from the hot spots?	Crowdsourced review sites (1), curated travel sites (1)	WebScraper, OpenRefine, Excel, Carto
3B.4	In detail comparing the different neighborhoods.	Crowdsourced review sites (1)	WebScraper, OpenRefine, RAWGraphs, Google Mapmaker, Carto
3B.5	Can data-driven technologies support providing visitors the experience of locals?	Social media (2)	Twitter API, WebScraper, RAWGraphs, Carto
3B.6	How can the visits of business travelers be extended?	Crowdsourced review sites (1), Open weather data (1)	WebScraper, OpenRefine, Excel

Table 3.3. The problem areas under investigation during Study 3B, and an overview of the data acquired by participant groups and their tool usage.

Throughout the three days of the study, all groups went through several iterations of Activity 1 (Question definition) to Activity 4 (Data exploration). Table 3.3 shows each participant groups' main research direction, the data sources, and the tools used. In the end, all participant groups managed to find valuable insights for their semester project. To better illustrate the kind of research questions the teams attempted to answer, an example: one team focused on approaching how seasons influence tourism. When they found that the correlation of seasonality and tourism is probably low for their target group, they focused on comparing the target city with similar cities, based on weather and other predictors.

PARTICIPANT REFLECTIONS: In the post-study questionnaire, the majority of participants' reflections were unanimous: all responses noted data acquisition as primary learning, followed by visualization of data and an increased general understanding of data, its processes, and its potentials for the design process. Besides three respondents with more technical background, the participants were also unanimous in reporting how challenging it was to scrape data.

Participants emphasized the transition from Data collection to Data exploration: *"[...] the moment we visualized the data using the tools provided to us. Finally all those lines of data were converted into a visual representation of the three days of hard work."* (participant from 3B.1). Some responses further reflected on the necessity for visualization to see the data in context: *"[...] visualizing the data. For me it first really makes sense and is useful, when I can see it visually, since this makes the data more concrete. Finding out that there were many different ways and different tools to visualize it, was nice."* (participant from 3B.3).

There were also other comments given and issues raised in the responses. A participant with a technical background reflected on demystifying working with data: *"[the study] helped me to understand that there is no need of any deep technical knowledge, to start playing with data and applying it [in the design process]"* (participant from 3B.2). More participants noted that the study helped them better understanding the phenomenon around big data and increasing their awareness of the online data traces: *"[the study] also made me more aware of the digital footprints I leave online, everyday. Many people are warning about this, but I had not quite understood it until now."* (participant from 3B.6).

3.5 Discussion

The current study explored the appropriation of data science practices by two groups of master-level design students into the design process. In this section, results from the two studies are positioned in HCI and design literature, highlighting further research opportunities.

GAINING DOMAIN-KNOWLEDGE

The two studies differed in working from provided data (Study 3A) and capturing data (Study 3B), and the participants familiarized themselves with the datasets differently. For Study 3A (master thesis records), we followed guidelines by D'Ignazio (2017) to work with familiar datasets and messy data. The participants were familiar with the general domain of the dataset as being enrolled in programs that require writing a master thesis to finish the study. However, several data properties were unclear for them (having one more year before starting their thesis project). The dataset was not entirely clean (Wickham, 2014), requiring the participants to do data cleaning on it. This 'friction' work turned out to contribute to gaining a more detailed understanding of the dataset. For Study 3B, as the participants worked on their ongoing semester project and had done research prior to the study, gaining domain-knowledge was less pronounced.

The importance of domain-knowledge has long been acknowledged and researched in data mining (Anand et al., 1995) and later in data science (Waller & Fawcett, 2013). Gaining domain knowledge needs to be considered when pursuing data science practices in the design process; access to a dataset, such as stumbled upon open data or a design process at a hackathon, still requires building up the understanding of what is inside the dataset. Additional description of the dataset can foster this understanding, sometimes called a data dictionary, to describe the different properties in the dataset. Designers can also use other, qualitative data inquiries for gaining domain knowledge, or can collaborate with a domain expert too.

NON-EXPERT DATA TOOLS AS AN ASSEMBLAGE

The steps of the data science process – such as capturing online data or cleaning a dataset – were followed through non-expert data tools selected appropriately for the needs and skill levels of the participants. Learning about these different data tools was highlighted as a major takeaway from the studies. The approach of using multiple different data leads to non-expert data tools forming a system assemblage (Kling & Scacchi, 1982), where the different tools enable different actions to be taken on the dataset. Following through the multiple steps of such a data science process happen by using non-programmatic tools. The system assemblage has positive and negative consequences. The assemblage enables designers to optimize their process using different tools for different tasks, choosing more appropriate tools for certain jobs. Furthermore, while some tooling is generic, such as a text editor that can perform basic string operations on a dataset (e.g., find and replace), other tools are data type specific. For example, geo-located data is typically inspected through map-based visualizations, while data with numbers and categories are plotted on graphs. However, different tools can require certain formats and data transformations to prepare the input. Dealing with different tools complicate the learning curve of different non-expert data tools and the assemblage's overall usability.

QUESTION-DRIVEN INQUIRY

Following through data science practices in Study 3B, the participants initially struggled with the computational thinking required by data acquisition through scraping and with understanding what kind of questions they could possibly answer by capturing and analyzing data. This understanding increased through an iterative process in defining better questions, and as a consequence, capturing more targeted data (approximately half the time of Study 3B was spent on doing multiple iterations). This iterative process of refining the research question and collecting data to extract insights applies data science practices of the co-evolution of problem and solution space (Dorst & Cross, 2001).

Designers are exposed to thinking about wicked problems (Buchanan, 1992; Rittel & Webber, 1973) and formulate design questions that generate design spaces (such as 'How might we...?' questions). However, during the study, questions towards falsifiable/provable hypotheses (resembling the 'scientific method') turned out to be more productive. Throughout iterations, participants both continuously learned more-and-more about the domain, and also improved the imposed questions that can be addressed via data inquiry. In our observation, working with digital data for design inquiry requires designers to formulate questions more precisely. A qualitative inquiry such as field observations can be 'forgiving' while being conducted, inquiry through collecting digital data requires precision in instructing a software tool for aforementioned data collection. In this way, the creativity of designers is channeled into hypothesis and research question formulation.

CREATIVE USES OF DATA EXPLORATION

A common data science terminology for the early step of exploring data is "Exploratory Data Analysis" (EDA). EDA was originally introduced for the exploration of numerical datasets using a statistical toolbox (Tukey, 1962, 1977). Commonly during EDA, various statistical techniques are applied to understand the data better, generate various hypotheses, and test those against the data. Yu (1994), following Pierce's pragmaticism explains how deduction, induction and abduction plays a role in EDA: abduction is used to generate a hypothesis, deduction to evaluate the hypothesis, and induction to justify the hypothesis with empirical data. Most commonly, data and visual analytics are targeted at using deduction to analyze data (Wong & Thomas, 2004).

Interestingly, the early phase of design is largely influenced by abduction (Dorst, 2011; Kolko, 2009). We observed the use of data as a source of inspiration, following abductive sensemaking, where an inquiry starts with a specific goal in mind. However, the emerging findings change what the initial question was. The approach of 3B.2 highlights this: they visualized social media hashtags and found subcultural and slang hashtags. From the visual inspection they found insights they did not know beforehand that they could use, and gaining knowledge that otherwise would have been hard to gain from user interviews or field studies. They used their findings not to prove a hypothesis, but for a creative thought-process to explore a phenomenon, otherwise they would hardly access. This is a creative way of using data – using data as a generative design tool – and one where the human abductive sensemaking is necessary to create the right connections.

Designers are trained in making sense of the world following patterns of thought where what is being designed is being informed by a constantly reframed problem space (Dorst, 2011). Our observations indicate that this skillset can be transferred for exploratory data analysis, using an abductive hypothesis generation as a creative process. Further studies in understanding the creative process throughout data science practices could help to inform new data uses and to generate future design methods with data.

LIMITATIONS

Our study contributes with an exploration of how designers appropriate non-expert data science practices for design inquiry. The study nonetheless was conducted with master-level design students. Future research with expert designers and in design practice would support generalizing our findings for designers on all expertise levels and designers working in a range of non-educational settings. Furthermore, the current study was limited to working with data collected from online resources.

3.6 Conclusions

The current chapter presented two exploratory studies elaborating on how data science practices could be appropriated in design practice. The findings demonstrate that existing non-expert data science practices can be combined into design practice. Furthermore, the findings show that designers transfer their creative capacity to hypothesis forming for data collection and use their designer sensemaking to synthesize data exploration of digital data in design inquiry. Elaborating upon this insight, in the next chapter, we look more in-depth into data science practices through the lens of a creative process.

LESSONS LEARNED FOR DEVELOPING METHODOLOGICAL CONTRIBUTIONS

This study shows how data practices can be integrated into design inquiry. We show the learnings from the study in Figure 3.4 on the interplay of designers' mindset lead by abductive appropriation of non-expert data science practices and tools.

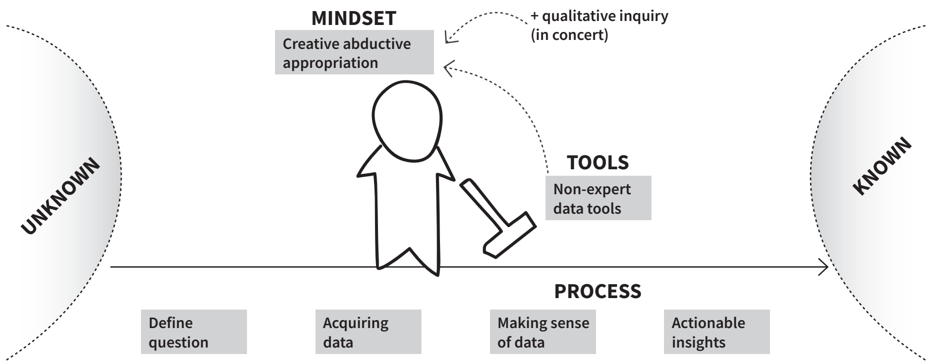


Figure 3.4. Data science process is the basis of the creative abductive appropriation of non- expert tools for design inquiry.

TOOLS: Non-expert data tools are suitable for design inquiry through data in general. We observed a 'creative tension' of designers using non-expert tools – initially made for analytical work. Instead of using the tools made for their specific target groups with specific needs, designers used the tools in creative ways to generate insights in design inquiry.

MINDSET: Even when working with digital data through tools initially made for data analysis, designers use their abductive sensemaking for synthesizing the findings of data exploration, resulting in a creative appropriation of tools. As second learning, design inquiry *merely* through data can offer limited insights and should be used for its strengths. Qualitative methods can provide deep insights into contexts, and inquiry through data could be used in concert with them.

PROCESS: From this study, it is clear that the power of using data science practices lies in *navigating the whole process*. The various steps start from formulating a question, acquiring data, and then making sense of data for actionable insights (see Figure 3.4). Even if designers collaborate with data scientists on these steps, the designer involvement is important to use data for design inquiry.

In the following of the thesis, we will build on the findings from this chapter. Next, Chapter 4 further investigates data practices through the lens of a creative process in order to conclude on the *Exploratory Data Inquiry* methodology, which we will use for developing a design method later.

Chapter 4

Data Practices as a Creative Process

In the previous chapter, a study has been conducted to gain an increased understanding of how designers appropriate existing data science practices. The study highlighted the creative ways of how designers approach data. In this chapter we further investigate how designers use non-expert data practices in design work, and analyze this work through a conceptual framework rooted in creativity and design process. With the presented exploratory study, we address RQ3 *of the thesis, "How can data science practices be characterized through a creative process lens?" This chapter presents the outcomes of three studies where service design teams used design inquiry through data integrated into their problem exploration phase. The findings show that observing data science practices with a lens rooted in assessing design methods enables to frame data science practices as design methodology, and to intertwine data and design practices. The results inform the development of a methodology to structure data science practices methodically and coherently into design processes. We coin this methodology Exploratory Data Inquiry.*

This chapter is based on:
Kun, P., Mulder, I., De Götzen, A., & Kortuem, G. (2019). Creative Data Work in the Design Process. In *Proceedings of the 2019 ACM SIGCHI Conference on Creativity and Cognition.* ACM. https://doi.org/10.1145/3325480.3325500

4.1 Introduction

Chapter 3 provided an early exploration of how designers appropriate data science practices integrated into design practice. The findings of Chapter 3 highlight that it is possible to integrate data practices into design practice and designers integrate data practices in a 'designerly' manner, such as by abductive sensemaking of data, similarly to the uses of design methods. More precisely, we found that designers' creativity manifests in hypothesis forming for data collection, and in the abductive sensemaking how designers synthesize the inferences from data exploration. These insights have informed the focus of the current study to explore data science practices through a creative process lens, to be able to intertwine data and design practices through a 'shared language'.

The current chapter, therefore, investigates **RQ3** of the thesis: *"How can data science practices be characterized through a creative process lens?"*. So far, we have seen that it is possible to conduct data practices as part of design inquiry. In order to understand more in-depth how these practices *intertwine*, a new lens is necessary that enables to examine them with an advanced vocabulary. As a result, the current chapter builds on the idea that if designers appropriate data science practices in a 'designerly' manner, then we could consider data practices as a *creative process* to analyze more precisely how data and design practices intertwine. To study this, we set up an empirical study as a design workshop with master-level design students. Due to the exploratory nature of the research, we aim to observe, in contrast to Study 3A and 3B, a more realistic and naturalistic design process. To allow the setup of the current study that is less controlled and has less prescriptive design activities, we remain to work with master-level novice designers in educational settings, acknowledging the limitations of generalizability of the potential findings, as discussed in Section 2.5. In order to effectively analyze the participants' process of design inquiry through data, we first expand the conceptual framework's *process* aspect to be able to interpret creative processes in more detail.

The contributions of the chapter are two-fold. First, the reported study provides an in-depth description of how novice design teams incorporate data practices into their design process, which is valuable for design educators who are interested in the hands-on learning of data practices through publicly available tools. Second, the proposed *Exploratory Data Inquiry* methodology helps designers and professions using designerly techniques, such as social innovators, product managers, or entrepreneurs, in understanding how to operationalize data practices in the early phase of the design process. Next, we zoom-in on expanding the conceptual framework with interpreting a creative process.

4.2 Interpreting a creative process

Our exploratory study from Chapter 3 and related work have indicated that exploration practices of data practices could be conducted creatively in a design process. However, the findings did not show how creativity manifests when using data practices in the design process. Next in the section, we deepen our conceptual framework's process aspect aiming to clarify the relationship between creativity and data practices.

Contemporary understanding of the creative process generally starts from finding a problem, towards generating ideas and then selecting the best idea (Sawyer, 2011, p.87), while covering the whole spectrum of designing. Design methodologists have used the concepts of divergence and convergence as basic tenets of creative work (Liu et al., 2003), which offers a simple, but illustrative framing of how to analyze creative processes. Jones (1970, 1992) used the divergence-convergence dichotomy as divergence being an act of enriching the options space (i.e., exploring answers, acquiring data), while convergence being an act of narrowing down the option space (i.e., defining a question, filtering the data or reaching conclusions from the data). From a process perspective, it can be concluded that how a designer or a design team moves from diverging to converging is fuzzy. Kaner (2014) calls this interim step as the *"groan zone"*; more precisely with the groan zone, he refers to the time, when the team feels at odds what is going on, how to interpret the outcomes from a divergent process, and how to align as a team. Gero (1996) offers a perspective on the interim step with the concept of emergence in a design process, as the act when the designer refocuses the attention or reinterprets results of the different actions taken. Interestingly, Gero's perspective can be seen as a parallel to the "groan zone". Figure 4.1 shows how the three concepts can be combined to describe one 'loop' of design activity. It is valuable to observe what happens when emergence happens, which can tell about the strategic considerations taken by the design team to conclude a design inquiry, and where that leads. To conclude, deepen the conceptual framework's process aspects with the concepts of divergence, emergence, and convergence, specifically for studying creative data practices in the design process.

Figure 4.1. Divergence, emergence, and convergence in the same loop will be used for analyzing the data and design process of Study 4ABC.

In the next section, we will present our study with design teams using data-rich design practices as part of a larger design brief. The findings of the study will be used to inform developing a design methodology to provide systematic guidance for designers to use data practices in design inquiry.

4.3 Method

To address our research objective about understanding how creativity manifests when using data-rich design practices, we set up a design workshop that was situated in the early phase of a larger design project. During the workshop, we observed three design teams who were in the process of reframing an ill-defined problem area towards a more articulated problem space.

CONTEXT OF STUDY 4ABC

Study 4ABC were part of a 4-months long service design semester project, during which we held a 3-days long 'data workshop' in the fourth week of the project. This particular moment for intervening with a data workshop was chosen, because by then the design teams already had done initial explorations into the problem domain. Differently put, our data workshop was embedded into the design teams' inquiry processes to reframe an ill-defined problem into more focused problem areas before engaging in creating a service design concept by the end.

This setup was selected on the basis of providing high ecological validity to observe the creative use of data practices as part of a real design project. It was important for the case selection to embed creative data practices into a design project that resembles design practice to be able to observe the rationalization of real designerly actions (as opposed to a lab study). Despite observing the work of novice designers, the design teams worked on real-world projects with real-world stakeholders, using a large variety of methods common in service design practice.

DESIGN BRIEF

The design projects' initial problem theme was 'mobility', specifically focused on Copenhagen, Denmark, referred to as 'city' in the following. Within the theme of mobility, the teams were steered towards focusing on the context of different neighborhoods to explore new value propositions in shifting urban mobility issues from the production of infrastructures towards people's own ability to change their environment through a change in their behavior. The participating design teams took different directions within this brief, as elaborated upon in Table 4.1.

Table 4.1. Overview of Study 4ABC.

	Study 4A – 'Reframing mobility'	Study 4B – 'Harbor'	Study 4C – 'New neighbor-hood'
Participants	n = 4 (1 male, 3 female), age between 20-25	n = 5 (3 male, 2 female), age between 20-25	n = 4 (1 male, 3 female), age between 20-25
Backgrounds	BA Industrial Design (2x), BA Digital Design, BSc Media Technology	BSc Industrial Design, BA Digital Design, BSc Media Technology, BSc Transportation Engineering	BA Digital Design (2x), BSc Communication Design, BSc Media Technology
Problem domain	Within the larger 'mobility' theme, the group did not reach to a more clear problem area yet. They focused the exploratory data inquiry on more closely examining two neighborhoods of the city, and to investigate a broader question on what 'mobility' means for people living in the city.	Within the larger 'mobility' theme, the group's focus was to investigate oppor-tunities of using the harbor of the city for new services or design interventions. In this regard, they focused on one, rather industrial, area of the city, to make it more sustainable and livable.	Within the larger 'mobility' theme, the group's focus was on a recently developed neighborhood of the city, which lacks social cohesion. The investigation was focused on better understanding the situation for designing a service or design intervention for citizens.
Data acquisition	Prior to the workshop: Twitter data collection based on 21 users and 76 keywords (n = ~60,000 tweets). Also, using interview transcripts from earlier inquiry. During the workshop: scraping Instagram posts based on location POI.	Prior to the workshop: Twitter data collection based on 51 users, 61 keywords (n = ~200,000 tweets). During the workshop: scraping Instagram posts based on location POI. Downloading one open data dataset from the city's open data portal.	Prior to the workshop: Twitter data collection based on 4 users, 15 keywords, and a geographical bounding box around one neighborhood (n = ~200 tweets). During the workshop: scraping Instagram posts based on location POI. Downloading three open data datasets from the city's open data portal.

PARTICIPANTS

The three design teams consisted of 4-5 first-year master-level design students in service design, with mixed backgrounds of engineering and design. The design students (n = 13) had prior coursework in user experience design, programming, and design research methods. The teams worked on a semester-long design brief within a studio-based learning environment. While working on the design problem, the education goals were to master different design inquiry and prototyping methods. The design teams received education on the ethical dimensions of research and were expected to uphold the same standards during the case study period as well.

MATERIALS

CURATED LISTS FOR CAPTURING TWITTER DATA: Two weeks before the design workshop, the groups were asked to think of inquiry that they would explore through data. To help the teams in formulating these hypotheses, they were asked to provide a curated list of hashtags, usernames and (optionally) a geographical bounding box to capture tweets from Twitter. We used the curated lists to set up Twitter data collection starting twelve days before the design workshop through Twitter's Streaming API using DMI-TCAT (Borra & Rieder, 2014). The Twitter data collection continued throughout the design workshop as well. As the data capture from Twitter was set up by the authors, the groups only needed to access DMI-TCAT for the analysis and export of datasets. In this regard, the groups only needed to curate the search queries for data collection.

ACQUIRING ADDITIONAL DATA: As described later, during the design workshop, the design teams became interested in capturing data from Instagram locations. For the scraping, the design teams used a Python script found online, with the help of the authors. Some design teams also used open data accessed from the local open data portal, downloading the datasets in comma-separated values (CSV) formats.

RECOMMENDED SOFTWARE TOOLS: The design teams could freely choose how to inspect and analyze the acquired data, but we provided recommendations for anticipated types of analysis. We suggested the use of standard spreadsheet software (Microsoft Excel, Apple Numbers, or Google Spreadsheets) for basic data operations or OpenRefine (OpenRefine, 2020) for more advanced data cleaning and transformation work. For visual analysis, the design teams received basic training in RAWGraphs (Mauri et al., 2017), and were also encouraged to use the common charting options in spreadsheet software. For more advanced visualization, the teams were suggested to use Voyant tools (Sinclair et al., 2018) for text analysis and Gephi (Bastian et al., 2009) for network analysis.

DATA COLLECTION

Throughout the design workshop, multiple interim presentations were held (middle and end of the first day, end of the second day, and final presentation on the third day), and the presentations were audio-video recorded by us. All design teams presented at these occasions, resulting in sum 12 recordings of 20-30 minutes footage each, totaling about 5 hours of material.

We audio-recorded also focus group conversations with each group to capture their reflections on data practices generally, on the workshop more specifically, and to learn what steps followed the workshop, 7-10 days after the workshop. The groups' final project reports were also collected at the end of the semester, which were analyzed to evaluate the data workshop's influence on their overall design process.

DATA ANALYSIS

The recordings were used to reconstruct the processes of the design teams. During the analysis, the cases were systemically coded, based on categorizing the different acts during the process following the description of divergence, emergence, and convergence from Section 4.1. We particularly followed the characteristics of what design rationalization took place when emerging, thus switching between divergence and convergence. Figure 4.2 shows this emerging process of each of the three groups.

In the following part, we will elaborate on the results of the data analysis and provide a rich description of the three cases.

4.4 Study 4A - 'Reframing Mobility'

At the beginning of the data workshop, Group 4A was working on debriefing 'mobility' and what this theme means to citizens. This broad aim was represented in the Twitter keyword lists the group provided (around 50 hashtags to track). One week before the workshop, the group decided to focus their inquiry on two districts of the city, and their expectation from the data workshop was to improve their framing of problem areas (i.e., design opportunities) in these two districts and to identify actors from these neighborhoods as potential collaborators.

Figure 4.2 shows that the group spent the first workshop day on getting familiar with the dataset collected from Twitter. For the second half of the first day, the group decided to split into two pairs; one that filters the data and a pair that explores the data in general. During the second day, the two pairs kept working in parallel. One pair started filtering and cleaning the Twitter data, and the other pair was experimenting with network analysis on the Twitter dataset, and then did a text analysis on the interview data. During the third day, the team explored Instagram briefly and focused on synthesizing their findings.

TWITTER DATA COLLECTION AND EXPLORATION

Two weeks before the workshop, for the keyword lists the group provided a long list of hashtags closely related to different aspects of mobility in an urban context (e.g., biking, public transport, etc.) both in English and the local language (a diverging act). Besides the hashtags, they provided Twitter user handles to track accounts of public actors related to mobility (e.g., activist groups).

When the workshop started, the dataset was about 60,000 tweets. The size of the dataset initially baffled the group, not knowing how to process it. Throughout the first day, the participants started to filter down the initial dataset (a converging act). Their filtering strategy was to eliminate tweets not made in the proximity of the city; since only a partial number of tweets consisted of the proper geolocation, they decided to use the tweets' timezone metadata. A further filtering direction was to focus on their two neighborhoods of interest, but that was harder to operationalize and was put into the background.

During the second day, the group explored the dataset through network analysis using Gephi, to find relationships in the tweets (emergence-convergence). They pointed out that working with network graphs produces intriguing visualizations, however, it is not trivial how to interpret and use the graphs in the design process. In the end, they used the graphs to explore the connections and see which hashtags go together more and less frequently (convergence).

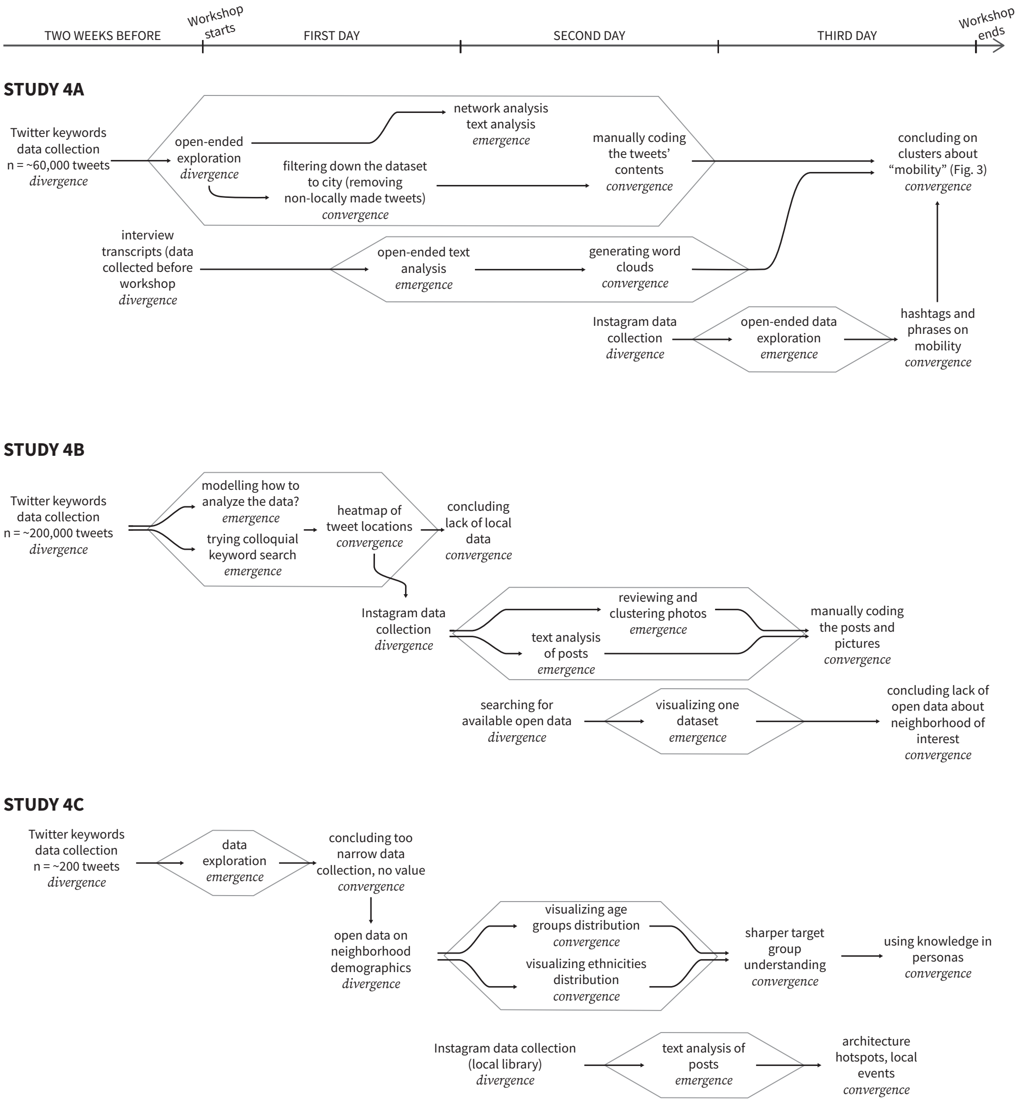


Figure 4.2. Overview of the group processes of Study 4ABC, visualized following the coding activity with the process-focused conceptual framework.

Parallel to the network analysis, the team also manually coded the filtered down dataset (emergence-convergence): they read the content of the tweets and coded the tweets depending on what mobility cluster the content of the tweet referred to (such as, bikes, cars, trains), as shown in Figure 4.3.

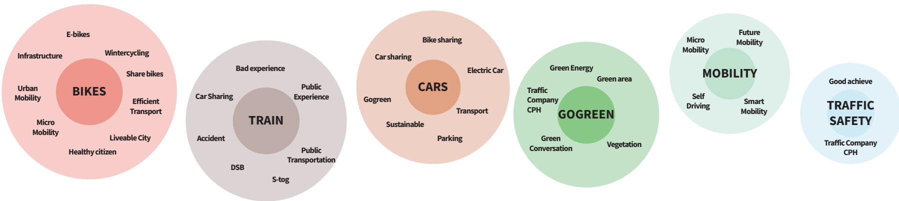


Figure 4.3. Clusters of Group 4A.

TEXT ANALYSIS OF INTERVIEW DATA

In parallel, the group also investigated what they could extract through data practices from the interview studies that they conducted a week before the data workshop. As the group rationalized it:

> *"We don't have an exact certain problem, so we tried to use the data we already had. One is interviews data (we did very broad interviews) and what we tried is to use the tools that we learned to dive into the interview data."*

The team had the interviews fully transcribed prior to the workshop, and this text corpus was entered into Voyant Tools (divergence). Not having a background in quantitative text analysis beforehand, they first explored word clouds as a simple and easily comprehensible way of exploring and visualizing the interviews' content. Later, the team investigated the functionalities offered by the tool, and tried to make sense of bi-grams and tri-grams how they provide new learnings about the domain (emergence).

EXPLORING INSTAGRAM

Towards the end of the second day, inspired by the other groups' experiences with Instagram data, the group decided to capture data from their district of interest (as results from **#district** hashtag) from Instagram (divergence). The captured posts were filtered down manually, based on relation to mobility (does the post relate to mobility - in, does not relate - out). This narrowed dataset was then analyzed as a text corpus in Voyant Tools (emergence). It revealed trending hashtags and phrases in relation to mobility at that time (convergence).

AFTER THE WORKSHOP: the group used the insights gathered from the data workshop to narrow down their problem space. By combining the findings from the workshop with the qualitative data they gathered previously through interviews and surveys, they decided to focus on bikes and sustainability, investigating possible service solutions for abandoned bikes. Interestingly enough, they used creative data practices also later on in their project, when they deliberately collected data in the form of CSV files to be analyzed with RAWGraphs to more closely explore the chosen specific issue of abandoned bikes. It was also interesting to see how the datasets collected in the workshop and during the design project were iteratively re-visited when, through field research, the students found out other possible questions to investigate from datasets.

4.5 Study 4B - 'Harbor'

By the beginning of the data workshop, Group 4B had already narrowed their focus on the city's harbor and one surrounding neighborhood. Their interests were to improve sustainability and livability in this area, which only recently started to be used for residential purposes compared to its past of being a large industrial zone. This focus resulted in a Twitter keyword list related to the harbor, and their interests, such as **#sustainability, #green mobility, #smartcity**, both in English and the local language.

The first day of the workshop was spent on exploring the Twitter dataset and trying to define how to ask questions to make sense of the data (see the process overview in Figure 4.2). Concluding that Twitter does not yield sufficient results, the group opted to capture data on Instagram focused on their neighborhood of interest on the second day, as well as to look into potentially interesting datasets on the local open data portal. The second and third days were spent on exploring and analyzing the Instagram data with different tools.

TWITTER DATA COLLECTION AND EXPLORATION

The group provided a list of over 50 hashtags, and user handles to follow two weeks before the workshop (divergence). By the start of the workshop, about 200,000 tweets were captured. When exploring the data, it became soon apparent to the group that the keywords based on such formal or academic language did not yield tweets about people's actual mobility experiences; tweets captured with such academic terminology were more part of professional conversations around the themes (convergence). As a consequence, the group tried to explore the collected dataset with more colloquial terms (such as '**delay**' or swearwords), but this approach did not yield useful results either (emergence). Driven by curiosity to try out the different tools and to see how these tools would afford new insights, the group explored making a heatmap visualization and network graphs. For the heatmap, the group combined all the datasets from the three groups and made a heatmap using the geolocations of the tweets that had this property. When reading the heatmap, they realized that no tweets came from their neighborhood of interest, and the harbor area only had tweets from the touristic zones. The finding made them conclude that the data collection from Twitter would yield only limited insights (convergence).

To approach the network analysis, the group first had an extensive discussion trying to different mental models on how to interpret the data (divergence). For example, they considered the question: "Who are the most influential users?" Such a question led them to consider what influentiality may mean and what properties could describe it in the dataset; for example, the number of followers or the number of retweets. Defining these questions to ask from the dataset then informed the procedures they considered following for exploration and analysis (emergence). The groups ended up creating some network graphs in the end (see Figure 4.4), but did not consider them too valuable for pursuing further.

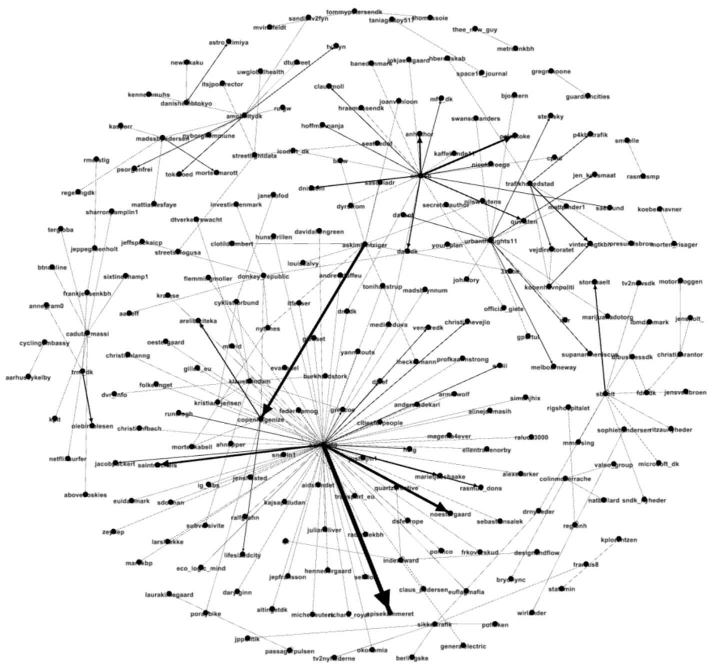


Figure 4.4. Network visualization of Group 4B.

EXPLORING INSTAGRAM

After realizing that the Twitter dataset especially lacks tweets about their neighborhood of interest, the group focused on Instagram and capturing posts with photos based on geolocation (divergence). Their query was based on the neighborhood of interest (both searched as a string as well as a location point-of-interest). They collected 1,000 posts, including pictures and metadata into a tabular data file. They explored the text corpus from the texts and comments of posts with Voyant tools, looking for the most popular hashtags that were combined with the **#district** one (convergence). The filtering enabled them to have a better sense of the hashtags people use. As the group described, “We started to find some more personalized hashtags”, which was contrasting to their experience with Twitter’s formal or academic hashtags in tweets (in their dataset). Under personalized hashtags, they meant local culture related ones, or postal codes.

Furthermore, the group also spent some time looking at the pictures with the posts, with which they also noticed the prominence of selfies, fitness content, and marketing items (i.e., brands). To address this bias, they manually categorized the pictures based on categories such as shops, food or sports (emergence). The group noted that a selection of the pictures could be useful in future co-design setups as discussion starters about the neighborhood or the themes they are interested in, as these pictures are not stock photos but based on everyday life.

OPEN DATA

The group also had a smaller inquiry with open data. They explored the local open data portal to check what kind of open data may be available about the neighborhood (divergence). Interestingly, the group found that their neighborhood of interest is not part of the open data portal's districts. They found one dataset in relation to the harbor and their interest (where are the registered parking areas for boats), which they visualized using Carto. Their main conclusion regarding open data was that their project could potentially focus on fostering more mobility-related open data to be captured and shared in this neighborhood (convergence).

AFTER THE WORKSHOP: the group returned to creative data practices when a relevant stakeholder shared its data about the harbor bus. With a narrower problem space and a specific dataset, the group could get useful information about people's behaviors in the harbor, which informed their design space too.

4.6 Study 4C - 'New Neighborhood'

Group 4C focused on a specific neighborhood from early on in their project. The group's interest was to explore daily life in the neighborhood and to identify design opportunities for their project. Two weeks before the workshop, for the Twitter data collection, the participants provided a narrow list of hashtags and users, all focused on the neighborhood. As a consequence, the data collection only resulted in a small dataset.

The beginning of the workshop was spent on looking at the small Twitter data collection, which was found too little to extract any significant insights (see the process overview in Figure 4.2). The low number of tweets made the group reflect on what data they would like to capture optimally; they were interested in a large Facebook group where the residents of the neighborhood organize themselves. Given the location specificity of their interests, the group explored Instagram data the second day, as well as open data getting to a better understanding of the demographics of the neighborhood.

TWITTER DATA COLLECTION AND EXPLORATION

Two weeks before the workshop, for the keywords list, the group was already focused on one specific neighborhood, which made them defining a keyword list that turned out to be rather narrow (divergence). By the start of the workshop, the tweet collection resulted in over 200 tweets, and the group found this dataset too small to extract substantial insight from it (convergence). They also emphasized geolocated tweets (over 30 tweets at the start of the beginning), but these tweets were all related to a large international concert held in the venue in the neighborhood. This event skewed the results, and the group noted that such biases in the dataset made it useless for them. The group's expectation was capturing daily life in the neighborhood, which was not delivered through collecting tweets.

OPEN DATA

From the Twitter dataset, the group concluded that they would like to learn more about the people that live in group's neighborhood-of-interest (emergence). To explore this further, one pair from the group explored the local open data portal for relevant datasets (divergence). The group found several datasets which they found instrumental in defining the target group of their project better. They looked at age distribution, ethnicity, and male/female ratio in the neighborhood. By using Excel for analysis, they found that the primary age group in the neighborhood is between 21 and 35 age, and relatively more children than older people, confirming their hypothesis that the neighborhood is primarily populated with young families. Additional insights on ethnicity revealed that over 68% of the residents are from the country, and the rest were born abroad. The group found these findings from the available open data valuable to inform their personas and target group profiles (convergence).

INSTAGRAM ON LOCAL LIBRARY

Towards the end of the first day, after the lack of success with the dataset from Twitter, the group did some exploratory searches on Instagram with some of the hashtags, and the findings seemed promising to further this inquiry the following days. On the second day, the group collected over 500 posts from Instagram searched on the neighborhood library (divergence). The group explored the collected data as a text corpus in Voyant tools, identifying the most common hashtags in the dataset (emergence). Hashtags on 'architecture' stood out in their exploration. As the neighborhood was built in the past 20 years featuring famous contemporary architecture landmarks, the group figured there might be more opportunities to explore this direction further in the next steps (convergence).

ONLINE ETHNOGRAPHY OF FACEBOOK GROUP

The group found a local Facebook page that collected residents of the neighborhood. In this group, people organized themselves for social events, swapping items, asking the public for their tips on everyday matters (such as, where are nice dog-walking paths). The group wanted to capture data from this Facebook group, but it turned out to be technically impossible. The group was interested in identifying reoccurring patterns that emerge from looking at the data on a longer timeframe, and for that to capture data from this group. Despite efforts to figure out an effective scraping method to get the data, it turned out to be technically too complicated, and they ended up observing the activity in the group as a non-participatory online ethnography (Kozinets, 2002).

AFTER THE WORKSHOP: the data workshop inspired the group to explore heatmaps in order to understand better how people moved inside the neighborhood of their interest. They only found data available about sports activities, and families' daily routines were not represented. They decided then to create the missing dataset through a participatory process, gamifying the data collection on a physical heatmap.

4.7 Discussion

This section aims to interpret our observations on how creativity manifested in the groups' processes. We start with considerations particular to the domain of the case, how learning about users through social networks took place. Afterwards, we elaborate on more domain-general points on how designers' empathy takes place in data collection and how creative data practices is conducted in concert with other design activities. After the section, we will distill our findings as a design framework.

LEARNING ABOUT PEOPLE THROUGH SOCIAL NETWORKS

The three groups engaged in inquiring people through social networks. While the participants' demographics are generally familiar with social networks, they faced domain-specific challenges due to lack of experience, not as a user but researcher of social networks. Designers can tap on decades of HCI research investigating behavior on online platforms of the past and today (e.g., (Gilbert & Karahalios, 2009; Litt & Hargittai, 2016)). In addition, social scientists and media scholars have analyzed how different social phenomenon unfolds on online platforms (e.g., (Stieglitz & Dang-Xuan, 2013)). For example, earlier work on social networks also established key metrics that can help in critically assessing a network (e.g., how to measure influence or engagement) (Bruns & Liang, 2012). Furthermore, these fields have also innovated on developing non-expert data tools for people with non-technical backgrounds, such as DMI-TCAT (Borra & Rieder, 2014) for Twitter data extraction, Gephi (Bastian et al., 2009) for network graph analysis, Voyant Tools for quantitative text analysis (Sinclair & Rockwell, 2012).

Learning about people through social networks also has caveats to be taken into account. First, collecting tweets with tools like DMI-TCAT does not provide capturing tweets from the past, as historical data can be acquired by purchasing access to it. This limitation was initially hard to accept for the groups. Therefore, to ensure a rich and 'large enough' dataset collected on a particular social phenomenon, designers need to take into account that data collection needs to run for some time. Second, the bias in data collection on social networks needs to be taken into account. As the groups also pointed out how people use different social networks and what content they share varies to a large extent, and designers also need to take into account their personal preconceptions about certain platforms. Biases need to be accounted for about data derived from social network sites, as the user base and ways of use of such sites are not random (Hargittai, 2015). While these limitations are present for other design research methods as well, these issues are non-trivial when designers emerge from *participating* on social networks to *using* social networks for research.

However, after their initial learnings of exploring data collected from Twitter, the groups approached social networks more informedly, addressing the limits in the second round of data collection. Groups B and C needed location-specific data approached Instagram targeting the geolocations Twitter failed to deliver. Group 4C focused on a specific online community in a Facebook group followed by digital ethnography. These examples illustrate that properly framing what data to collect is especially crucial for creative data practices. While social networks are common and familiar platforms in the life of designers working with digital technology, the necessary domain-expertise cannot be neglected to effectively social networks for creative data practices. These findings can be generalized for other domains as well; designers should investigate prior research and expertise in the domains they explore and be aware of biases and other factors that may influence their data collection.

EMPATHETIC DATA QUERY DESIGN

Search queries play a significant role in defining what data to collect from large data infrastructures. While search queries are widely used in everyday life on finding information, their use becomes less trivial for defining queries for data collection. The workshop procedure started with the groups composing curated lists of search keywords with Twitter user handles (**@user**), hashtags (**#hashtag**), or simple text phrases ('**text phrase**'), and their combinations, including boolean operations of AND and OR. The composition of curated lists generated extensive discussions and reflections throughout the case study. The design teams recognized that for using data practices creatively, a different, designerly mindset is required to define the queries in curated lists. Initially, their search queries were approached from information search aspects, often using precise terminology from the domain of interest (e.g., **#sustainability, #green mobility**). However, after exploring the datasets from 12 days of continuous recording of tweets, it became clear that formal terminology is not how people talk in real life or on social networks.

Rogers (2017) discusses the considerations for query design for media researchers. In his claim, search queries lead to certain languages people use, and that carries a statement of how different media outlets (professional media like newspapers or social media individuals) position their messages on controversial public discourse. While this perspective has been focused on journalistic and media practices, it applies to designers too. Designers need to incorporate sensitivity to people beyond mere information seeking. During the case study, the design teams needed to appropriate their training in being empathetic with the people they research and use their terminology in defining their search queries, resulting in a mindset shift how queries were approached. The teams needed to explore the taxonomy of *colloquial language* around their observed phenomenon. For example, people rarely talk in terms like *'mobility'*, but as of *'delay'*, or swear words; sharing stories that happened, such as the bus was late again.

Reinterpreting search queries through a more-empathetic query design indicates new creative ways of how designers' sensitivities manifest when working with digital data. In practical terms, we recommend to designers to develop a curated list, that 1) supports to infer insights from the investigated question; 2) generates a 'large enough' data collection, so there is a substantial dataset to explore; 3) the signal-to-noise ratio is good, the dataset is relevant and captures valuable content. These recommendations are domain-general, however in specific contexts, further criteria are potentially necessary.

Iterating the query design can be done similarly as design work, which keeps the activity familiar and compatible with common design activities.

CREATIVE DATA PRACTICES AS DESIGN INQUIRY

Throughout the workshop, the opportunistic use of different inquiries resembled the bricolage practice (Louridas, 1999; Vallgårda & Fernaeus, 2015) of designers around problem-framing. The design inquiries enabled designers to extract valuable insights from different data sources through different analytical tools. In other words, data practices were used for creative framing and reframing of problem and design spaces, as formulated by Dorst (2011). Figure 4.5 shows Dorst's framing lens of *"What? – thing", "How? – working principle", "Aspired value"* as an equation how data practices were reinterpreted as design reasoning.

What? *(thing)*	+	**How?** *(working principle)*	=	**Value** *(aspired)*
some data *(social network posts, interview transcripts, filtered datasets)* ...		open exploration text analysis network analysis ...		answer to question insights quotes ...

Figure 4.5. Creative data practices in framing (based on Dorst (2011)).

Within the three framing ingredients, we recognize three starting points to rationalize creative data practices:

1. Having a dataset to explore and extract value from (What? is fixed), in this case, open exploration of data or an analysis method can lead towards value;
2. Having a method provided (How? is fixed), in this case, the method requires certain types of data which informs the data acquisition for extracting value;

3. Having an aspired value to extract, which can both inform the dataset and its analysis.

Approaching creative data practices following these three starting points makes it possible to use data practices in concert with other design inquiry methods. For example, Group 4A explored interview transcripts through quantitative text analysis, which was different from a ‘regular’ qualitative data analysis. Similarly, groups shortlisted influential Twitter users for a future interview study, creatively appropriating their inferences from data. Such hybrid thinking around data practices suggests that how designers would fluidly combine different types of qualitative inquiry, they do that naturally with digital data as well. The design teams emphasized that exploring digital data is less resource-intensive than a field study, given that no travel and human work hours spent on the field are required, and the data collection can happen without having someone on the field. Such a rationale shows that despite the limits of acquiring and exploring digital data can be a strategic choice of inquiry for initial explorations.

4.8 Exploratory Data Inquiry methodology

While using data is not unfamiliar for designers, it is often unclear how to approach data practices creatively in the design process. The analysis of the three studies and the corresponding discussion highlighted patterns in the process that can be described by the expanded conceptual framework in Figure 4.1. We propose a methodology to structure creative data practices in the design process as a mode of inquiry. This methodology is referred to as Exploratory Data Inquiry, a methodology to guide designers of any experience and technical expertise to make methodical considerations of data practices around a creative process from framing to inferring from the data. Figure 4.6 shows the outline of Exploratory Data Inquiry that consists of three steps, which are elaborated below.

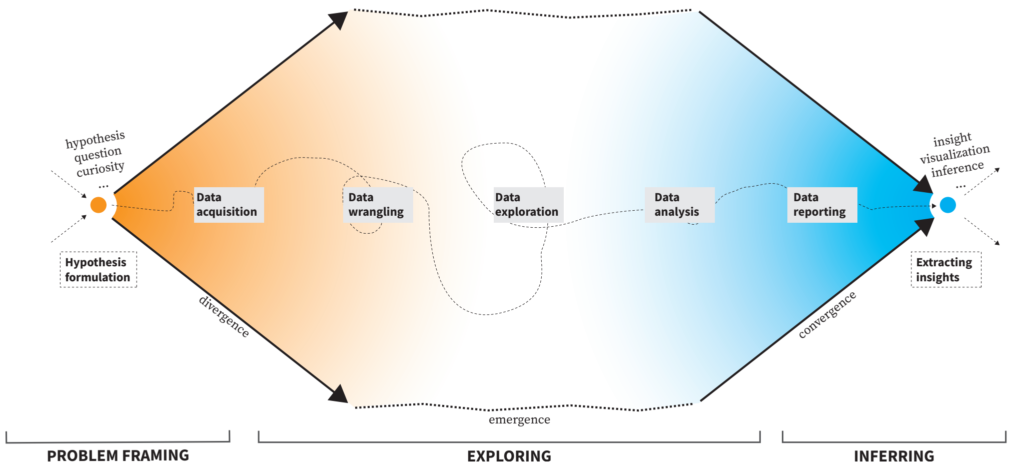


Figure 4.6. The overview of the Exploratory Data Inquiry methodology, combining a divergence-emergence-convergence creative loop and related acts of data practices.

1) PROBLEM FRAMING: Exploratory Data Inquiry starts with framing the problem that is rooted in the design problem's domain. Contemporary understanding of creativity emphasizes the importance of framing the right problem first (Sawyer, 2011, p. 90), therefore framing is often converged into a hypothesis or a question to lead the following Exploratory Data Inquiry steps. This frame is subject to change as the understanding of the data and the problem-design space grows, evolving the frame iteratively.

2) EXPLORING: The exploration process of data happens starts from immersing into the dataset, filter, transform, and clean it (also called data wrangling (Kandel et al., 2011)). This process evolves from an open-ended exploration towards conducting different analyses on the data, following an opportunistic mindset to steer towards findings valuable insights to evolve the problem-design space.

3) INFERRING: The final step of Exploratory Data Inquiry is to reach conclusions by implicit or explicit inferences. Sometimes the conclusions are tacit and personal, building the design team's common-sense about understanding the problem, which can also trigger going back to updating the problem frame. Some more explicit conclusions are inferences as answers gathered for the leading question, or extracting information from the data, or creating a representation of the data as a visualization.

EXPLORATORY DATA PRACTICES

Interwoven with the three steps of *Exploratory Data Inquiry*, different acts of exploratory data practices take place throughout the process (Alspaugh et al., 2019).

DATA ACQUISITION: Data can be acquired in various ways from the problem framing and the leading question. During the case study, data was acquired from social networks through data scraping (a technique to download the content of a website and structure it into a dataset) (Mitchell, 2018), or capturing data from data infrastructures through application programming interfaces (APIs). APIs are primarily meant for machines to connect different sources of data, but they can also be polled for extracting data as datasets.

DATA WRANGLING: Data wrangling is "a process of iterative data exploration and transformation that enables analysis" (Kandel et al., 2011). Wrangling is a technical task to clean the data from inconsistencies and prepare it for exploration.

DATA EXPLORATION: Data exploration is an open-ended process, which often starts with discovering the data in spreadsheet software, to understand the properties and identify the structure within the dataset. As this divergent process progresses, more convergent analyses can be initiated.

DATA ANALYSIS: The previous open-ended exploration enables more focused analyses to take place, targeting different dimensions and characteristics of the data. For example, the case study groups explored the same Twitter dataset through network analysis of users and hashtags, as well as text analysis of the content of the tweets. These analyses can be exploratory in nature, depending on what conclusions are reached at the end.

DATA REPORTING: The analyses outcomes may come as tacit knowledge generation (the design team understands the problem better), but generating visualizations are more common. Visualizations can be used as boundary objects to communicate results within the team and with stakeholders.

The three steps of Problem Framing, Exploring, and Inferring follow a sequential order, and are a general scheme applicable broadly on different design inquiry methods. Similarly, the exploratory data practices are also presented in a sequential order. However, progressing through these steps and practices inherently lead to a continuous familiarization and learning about the problem domain, which can result in iterating back to the initial hypothesis formulation, or to lead to a skipping steps or practices to extract insights.

4.9 Conclusions

The current chapter presented a study on designers conducting data-rich design practices during the early phase of design. The designers' work was analyzed from a creative process lens to see how divergence, convergence, and emergence takes place through data practices during the design process. The study and the analysis revealed how creativity in framing manifested in data practices. We used the observed processes of the study teams as an input for developing the *Exploratory Data Inquiry* methodology to help to structure creative work with data. The framework can help designers and other professions using designerly techniques to operationalize data practices in the early phase of the design process.

LESSONS LEARNED FOR DEVELOPING METHODOLOGICAL CONTRIBUTIONS

This study shows how data science practices can be intertwined with design inquiry through a creative process lens. From the study, the primary contribution is the *Exploratory Data Inquiry* methodology, as it brings together data and design practices under the same process and with a shared vocabulary, as we will explain below as well.

TOOLS: With this study, we have gained insights into the caveats of using social media for data acquisition. First, the tools for acquiring data require more technical competence than the non-expert tools from the previous study. Second, tools for data acquisition are subject to technological possibilities and may introduce unusual preparation steps and timeframes prior to use them in design inquiry.

MINDSET: Connected to the data acquisition tools from above, what data to acquire requires a non-trivial type of work to design the queries for data acquisition. Developing these queries can benefit from the empathetic focus of designers. This study also deepened our understanding of how designers' sensemaking works through exploring data. Framing and reframing practices also take place while using design inquiry through data, leading to data exploration that is more opportunistic in characteristics than data analysis.

PROCESS: The Exploratory Data Inquiry methodology is a major contribution to the framework to explain the process of intertwining data and design practices. From this study, it can be seen that design inquiry through data can be considered as three conceptual stages of problem framing, exploring, and inferring. Figure 4.7 shows these stages as diverging-emerging-converging loops typical in design work.

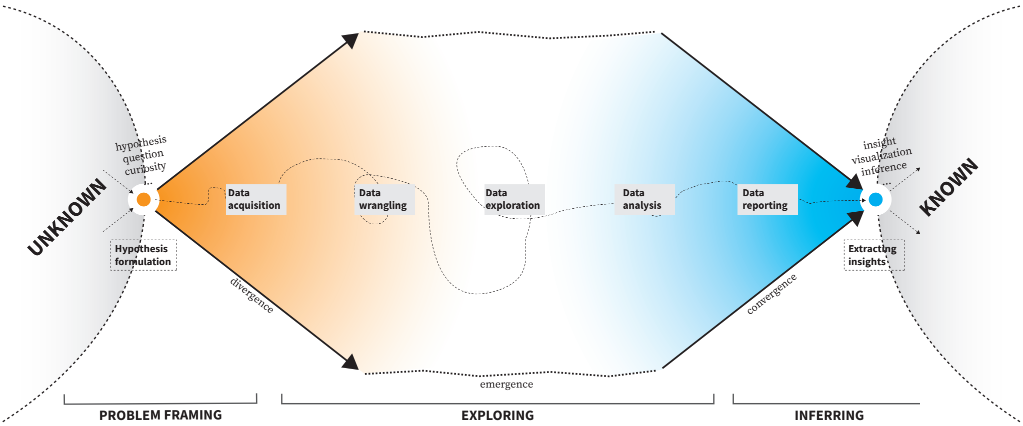


Figure 4.7. The Exploratory Data Inquiry methodology can be used to illustrate how the transitioning in design inquiry intertwines data and design practices.

Our study contributes to the understanding of data practices as a creative process, nevertheless, our approach has limitations as the empirical work was conducted with novice designers. While the Exploratory Data Inquiry methodology is domain-general and presents a high-level perspective on creative data practices, future research with expert designers would support validating our findings. Future work thus points in various directions. First, with a more nuanced understanding of the characteristics of creative data practices through Exploratory Data Inquiry, we see potential in using the methodology as a guide for new design methods to support designers through the different steps, which we will show in Chapter 5. Second, we aim to use the Exploratory Data Inquiry methodology in different domains and with different sources of data, as will also be shown in Chapter 6.

Chapter 5

Developing a Design Inquiry Method for Data Exploration

In the previous chapter, a study was conducted to explore data science practices intertwined in design practice using a creativity process perspective. The findings of the study have been distilled into a methodology that we coined as Exploratory Data Inquiry. The Exploratory Data Inquiry methodology provides a familiar vocabulary for designers to approach exploratory data practices integrated in design practice. In this chapter, we address RQ4 *of the thesis, "How can a design method support design inquiry through data?" In order to answer this question, we build on the Exploratory Data Inquiry methodology and develop a design method for design inquiry through data. We coin this design method as Data Exploration for Design method. The current chapter presents a study to evaluate the method and also to gain a deeper understanding of the mindset and approach of data exploration for design inquiry. We close the chapter by positioning the findings in design theory literature and motivating a set of principles of data exploration for design inquiry.*

This chapter is based on:

Kun, P., Mulder, I., & Kortuem, G. (2018). Data Exploration for Generative Design Research. In *Proceedings of Design Research Society 2018* (pp.1342-1356). London: Design Research Society. https://doi.org/10.21606/drs.2018.565 (Best Paper Award)

Kun, P., Mulder, I., & Kortuem, G. (Under review at *Interaction Design and Architecture*). Developing a Design Inquiry Method for Data Exploration.

5.1 Introduction

Working with data is becoming unavoidable in design practice. As an increasing amount of contemporary life is conducted through digital and connected artifacts, everyday life is becoming facilitated through data as well as captured in large data infrastructures. Different industries and scientific fields have found new ways to inquire about their problem domains through such large datasets. As already introduced in Section 2.3, by exploring digital data, new insights can be found that otherwise with 'old ways' would be unobtainable, but it remains unclear though how *designers* can leverage on such large sets of data in their design practice. In Chapter 3, we explored how designers 'naively' appropriate data science practices into design practice. That study highlighted that designers approach data in 'designerly' ways, as a tool to contribute to the co-evolution of problem and design space exploration. According to this finding, it is possible to approach data science practices as another element in designers' repertoire, such as design methods. We pursued this direction further with another study reported in Chapter 4, in which we analyzed data science practices through a creative process lens of divergence-emergence-convergence. Since these divergence-emergence-convergence loops are common in design practice as well, analyzing data and design work with a shared lens enables us to intertwine these two practices better. Consequently, we concluded the *Exploratory Data Inquiry* methodology combining these two practices.

In this chapter, we address **RQ4** of the thesis, "How can a design method support design inquiry through data?" To answer this question, we use research-through-design as an approach, where we first motivate the development of a design method (i.e., the designed artifact of RTD) based on the Exploratory Data Inquiry methodology. Specifically, we frame data exploration from a design inquiry perspective and contribute to design practice in the big data era by presenting data exploration as a design method. We coin this design method as Data Exploration for Design method. The following sub-research questions guide this particular study:

- › How can data exploration be approached as a design inquiry method?
- › What kind of mindset and expectations do designers assume while using data exploration as a design inquiry method?

As highlighted earlier in Section 2.2, the contemporary understanding of design methods goes beyond step-by-step process guides, and the *mindset* designers assume while using a method is just as much important as the process. Consequently, the current study also investigates such a mindset, as shown with the second research question above. In the continuation of the chapter, first we will present our design rationale for developing the design method, and then we present the *Data Exploration for Design* method and a corresponding study we conducted to learn about the mindset and expectations of creativity support in the context of design inquiry through data. We finish the chapter with a set of principles to follow data exploration as a design inquiry, when data exploration is fundamentally intertwined with design inquiry, beyond the usage of the method.

5.2 Design rationale

First, we will revisit key highlights from Chapter 2 in keeping with the learnings from Chapter 3 and Chapter 4. The review on design inquiry (Section 2.1) and non-expert ways of learning and using data (Section 2.2) has resulted in a few general design principles to approach the development of a design method for data exploration. First of all, it has been concluded that data exploration can intertwine fundamentally with design inquiry, and therefore should be approached in an *open-ended* and *holistic* way. With the term open-ended, we refer to support data coming in various shapes, formats, or topics, catering to the unlimited types of design situations designers face. The learnings from Chapter 3 have contributed to an updated interpretation of designers appropriating data-rich design practices in an *open-ended* way by using non-expert data tools creatively and by combining data and qualitative inquiries. Under holistic, we mean to support the complete data workflow, from asking a question to be addressed by data, to data collection and transformation, to inferences from data. The learnings from Chapter 4 updated our interpretation that designers approach design inquiry through data *holistically* through the three particular steps, which are 'problem framing', 'exploring', and 'inferring'. Furthermore, designers follow data exploration steps embedded into divergence-emergence-convergence loops of design practice.

The *open-ended* and *holistic* design principles of approaching data exploration in design work lead to creative usage of data in the design process. To interpret such creative usage of data in practical terms, we use the four levels of creativity framework by Sanders and Stappers (2008), a practical framework for everyday manifestations of creativity. In this framework, Sanders and Stappers (2008) define *Doing, Adapting, Making, and Creating* as an increasing order of expertise/interest as can be seen in people's lives. They argue that people can be simultaneously on different levels of creativity for different areas of life. Considering designers' relatively low level of data expertise, we assume that most designers today would be on the levels of *Doing* and *Adapting* to utilize data. In Table 5.1, we present an adjustment of their framework for our design rationale, to serve as guidance for the development of our design method. Inspired by Sanders and Stappers' creativity framework, we elaborate upon the levels of *Doing* and *Adapting*, interpreting as *Doing* with data and *Adapting* data techniques for design inquiry.

Level	Type	Description
4	Creating	The highest level of expertise/interest in this spectrum, addressing such cases that fundamentally transforms the design practice intertwined with data.
3	Making	The level of *'asserting own ability or skill'*, which we see as the utilization of data commonly in one's design practice.
2	Adapting	Appropriation of techniques starts to happen at this level. This appropriation can be guided and inspired, by appropriating data thinking and existing data techniques into one's process.
1	Doing	The level of being able to transform a dataset independent of a tool (thus having a sense of how to manipulate a dataset) is part of general technical literacy, at least through basic knowledge of spreadsheets software (e.g., Excel).

Table 5.1. Four levels of creativity defined by Sanders and Stappers (2008) adjusted for interpreting creative use of data exploration in design inquiry.

The principles of *open-ended* and *holistic*, together with the four levels of creativity defined in the creativity framework (Sanders & Stappers, 2008) have been made operational for developing a design method for data exploration following the taxonomy of Sanders, Brandt, and Binder (2010, p. 196). In their terms, *tools* are “material components used in design activities”; *toolkit* is a collection of tools used in combination for a specific purpose; *technique* is a description how tools and techniques are put into action; *method* is a combination of tools, toolkits, techniques put together strategically towards a specific design research plan, and at last, *approach* refers to an overall mindset for conducting the design research plan. In keeping with this taxonomy, we constructed our design method consisting of a workshop procedure, a curated recommendation of existing software tools, and design tools (card decks and booklets), as elaborated in the next section. Furthermore, a design method should not only guide to set realistic expectations about data, but also indicate the potentials of data with growing data expertise. We addressed our assumption that most designers lack data expertise by scaffolding data exploration in the format of familiar design tools, while supporting a dynamic skill acquisition process and open-ended and holistic use of data exploration for design inquiry.

In the following section, we present the resulting design method for data exploration, which we refer to as *Data Exploration for Design* method.

5.3 Data Exploration for Design method

The *Data Exploration for Design* method aims to enable and guide designers to creatively explore and use datasets for design inquiry. The purpose of such a creative exploration of data is to enable extracting valuable inferences for the design process, that otherwise would have been harder to technically infeasible to find by using other design inquiry methods.

In keeping with Sanders, Brandt and Binder’s taxonomy (2010), the presented *Data Exploration for Design* method consists of three disjunct components; a method outline, recommended software tools for data operations, and design tools. A method outline forms the primary basis, which combines data exploration and design inquiry into an intertwined approach through a procedure of conceptual stages. The method outline is complemented with software tools commonly used by other non-expert data communities. Furthermore, we developed a set of card decks and booklets to support the learning curve of novices during the workshop. The next sections present each of these different components of the method, respectively.

METHOD OUTLINE

The method outline has been designed in keeping with the *Exploratory Data Inquiry* methodology earlier introduced in Section 4.8. Figure 5.1 shows how the *Exploratory Data Inquiry* methodology can be framed more directly as an iterative stages. The *Data Exploration for Design* method follows *Exploratory Data Inquiry*'s three conceptual stages of problem framing, exploring, and inferring. The three stages integrate into an inquiry within a design situation. This outline is used to develop a workshop structure for a one-day workshop setting, where the input to the design process is a design brief and an available dataset. The one-day format is not a restricting way of conducting the method, as methods evolve and integrate with individual design practices (Daalhuizen, 2014; Gray, 2016; Schønheyder & Nordby, 2018).

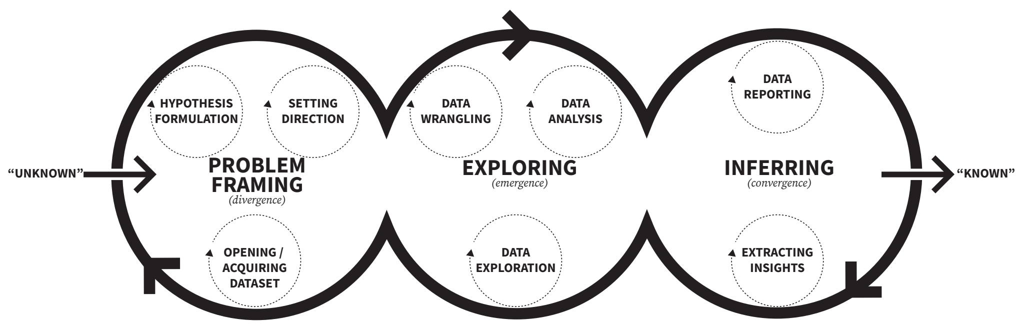


Figure 5.1. The outline of the Data Exploration for Design method, following the three conceptual stages from the Exploratory Data Inquiry methodology. The different conceptual stages proceed in sequential order, but iteratively.

PROBLEM FRAMING: The first conceptual stage of the Data Exploration for Design method is centered around framing the problem to explore through data. During this stage, the designer sets up a data exploration by formulating a hypothesis, opening or acquiring a dataset, and setting a direction for the data exploration. Hypotheses emerge in various shapes; it can be an explicit hypothesis or research question but can also be an opportunistic 'curiosity' or a 'hunch' when the problem formulation is still in the early stages. Data exploration continuously proceeds from implicit hunches towards explicit research questions used for proving a hypothesis. Following a hypothesis or research question, a direction can be set for exploration. The direction bridges how to explore a hypothesis and what data is available for such exploration. If a dataset is already available, it is a much lower effort to set the data exploration strategy that suits the data, such as what type of methods and tools can be used for the given dataset. Similarly, when a specific data exploration method or tool is readily available, then the data acquisition can be defined accordingly.

The three components mentioned above are continuously evolving in the Problem Framing conceptual stage. In other words, if the design process is based on a design brief, then in this stage, the brief is being explored from a data perspective. Typical questions in this stage are: *"What hypothesis do we want to inquire about?", "What datasets are available?", "How will we explore the data?"* The co-evolution process of designing provides answers to these questions, as the design problem unfolds. Therefore, iterating back to this conceptual stage is a natural part of processing through the method.

EXPLORING: The second conceptual stage of the Data Exploration for Design method is centered around the actual exploration of the data and data operations necessary for that. During this stage, the designer is wrangling (transforming and cleaning) the data, exploring it, and conduct different data analyses on it. These steps are attempts to productively process the dataset to explore and analyze it in ways that can fuel inferences into the design inquiry. Data wrangling is an essential step in working with data, as significant proportions of time are spent on cleaning and processing the data. Cleaning and transforming the data are iterative steps, with the aim to decrease the extent of corrupted data and to shape the data for different exploration and analysis tools. The most valuable time to inquire into a design problem is spent in steps of data exploration and data analysis, by increasingly understanding the problem space and finding answers to hypotheses and research questions. The available dataset, the research questions, and the design situation, result in myriad combinations for data exploration and analysis.

Connected to the direction set in the previous conceptual stage, the designer will explore the data pursuing a particular interest (i.e., research question) in mind, however throughout the process itself, as the understanding of the problem grows, the research question may continuously evolve. Thus, iterating between exploring and problem framing conceptual stages is an expected proceeding through the method.

INFERRING: The third conceptual stage of the Data Exploration for Design method is centered around extracting valuable inferences out of the explored dataset. During this stage, the designer extracts insights and works on reporting the findings from the inquiry process. The conclusions from the data exploration process trigger a new iteration of inquiry with the same or a different design method or proceeding further in the design process. The steps in this conceptual stage build on representations and visualizations generated from the Exploring stage. Such outputs can be utilized further in the design process as boundary objects, contributing to the increasing understanding of the design situation and problem space. Beyond visualizations, alternative inferences are different insights, such as answers to a research question. While explicit answers to research questions are often contained in a report or presentation to stakeholders, implicit findings are also generated throughout the data exploration process. Such ‘small insights’ help to build the common sense thinking about the problem domain. These different types of insights can lead to iterating back to the previous conceptual stages, which is an expected proceeding through the method.

Different types of tools support the three conceptual stages. As can be seen from the description of the three conceptual stages, the thinking processes of designing are intertwined with thinking and working with data. These processes are supported by a combination of design tools and non-expert data tools (see Table 5.2). Under design tools, we refer to supporting materials for learning, and under non-expert data tools, we refer to publicly available software tools that are wide-spread and widely supported by non-expert communities.

Table 5.2. Design tools and non-expert data tools used in the Data Exploration for Design method.

Stages	Problem framing	Exploring	Inferring
Design tools	'Basic data types and techniques' card deck 'Questions for data' booklet	'Data techniques' card deck 'Questions for data' booklet 'Working with data 101' booklet	'Questions for data' booklet
Non-expert data tools		Spreadsheet software (e.g., Excel) Data wrangling tools (e.g., OpenRefine) Data visualization tools (e.g., RAWGraphs) Data analysis tools (e.g., Voyant Tools - text analysis, Gephi - network analysis)	Data visualization tools (e.g., RAWGraphs)

In the following, we will present the developed design tools and a curated set of software tools.

DESIGN TOOLS

Although a substantial part of data exploration happens through software tools, the cognitive aspects of data exploration are equally important tacit knowledge to be gained by working with data. The cognitive aspects, such as computational thinking or sense-making of data, are part of an initial learning curve that will become part of a designer's mindset. We addressed this learning curve by developing two card decks and two booklets to scaffold various data best practices. Card decks are ubiquitous design tools (Roy & Warren, 2019), and have also been effectively used for data visualization (He & Adar, 2017). Card deck based tools have also been used to bring theoretical academic work into design practice, using card decks as tools to facilitate workshops (e.g., (Deng et al., 2014; Hornecker, 2010)). Following such examples, we have approached the support of open-ended data exploration in a domain-general and extendable way by the use of card decks and booklets. We have aimed with the card decks and booklets to introduce basic design tools that can be reproduced with a home printer, tailored for specific datasets and design situations. For example, a card deck can be extended with additional cards based on the different types of data in a dataset or domain-specific exploration possibilities. The booklets are eight-page foldouts, which is a limited format to contain focused information. We have deliberately left undesigned how to use the card decks and the booklets to foster creative exploration and intertwining how these design tools can integrate into designers' practices. However, we have expected some regular usage patterns for card decks, such as 'forced pairing' of cards to trigger new ideas by combining different cards or using different cards as a way to 'reverse engineer' and model existing data projects.

The following sections present the basic card decks and booklets prepared for the current study. As specified before and similar to the workshop procedure, we expect derivatives how the *Data Exploration for Design* method is being appropriated. These basic cards and booklets have informed the design of extended versions used in pressure cooker events.

CARD DECKS

In this section, we first present the two card decks developed for the current study and then discuss the extensibility to alter and create new card decks.

BASIC DATA TYPES AND TECHNIQUES: These cards provide a quick overview of the basic types of data and the most common and essential data techniques that are applied on datasets (see Figure 5.2). These cards can be used as a reminder of considering alternative options, as well as a quick reference to browse through a dataset. One part of the card deck is cards summarizing the various types of data commonly found in datasets describing everyday phenomena, such as numerical data, geo-located data, categorical data, or textual data. The other part of the card deck is a collection of fundamental activities one can perform with data, such as: compare or identify data points. These activities are so prevalent that they go unnoticed in most cases. However, when someone is unfamiliar with using computational thinking, these activities do not naturally come up (such as selecting a datapoint – identify).

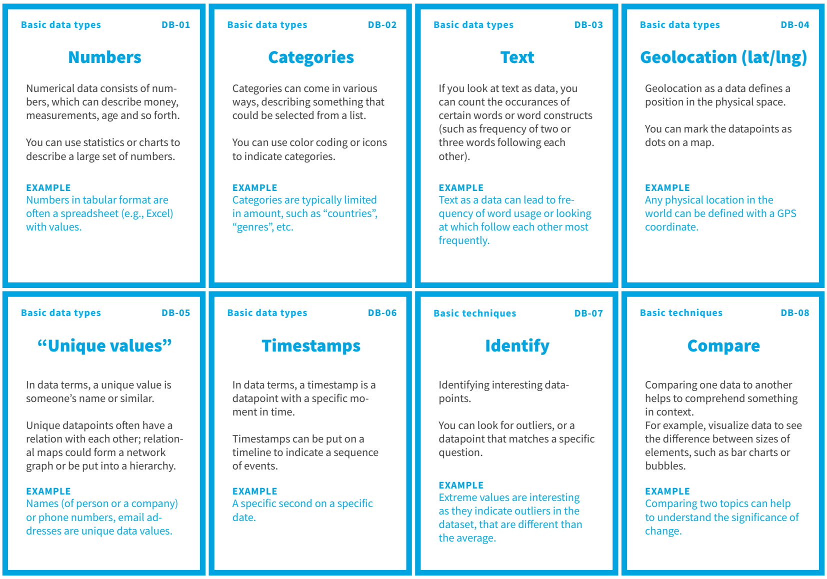


Figure 5.2. The Basics of data card deck summarizes the most elementary data types and data techniques.

DATA TECHNIQUES: This card deck is a summary of typical techniques to apply on a dataset in order to extract further meaningful information out of the data (see Figure 5.3). An example data technique is map visualization, which can easily be accomplished, for instance, when the dataset contains GPS coordinates. The related data technique card provides a basic overview of what kind of input(s) such a technique requires (e.g., GPS coordinates or addresses). One explicit aim of the card deck is to foster different data exploration techniques, i.e., not to fixate on one type of exploration. This also serves to stretch a learning process and to go beyond familiar methods.

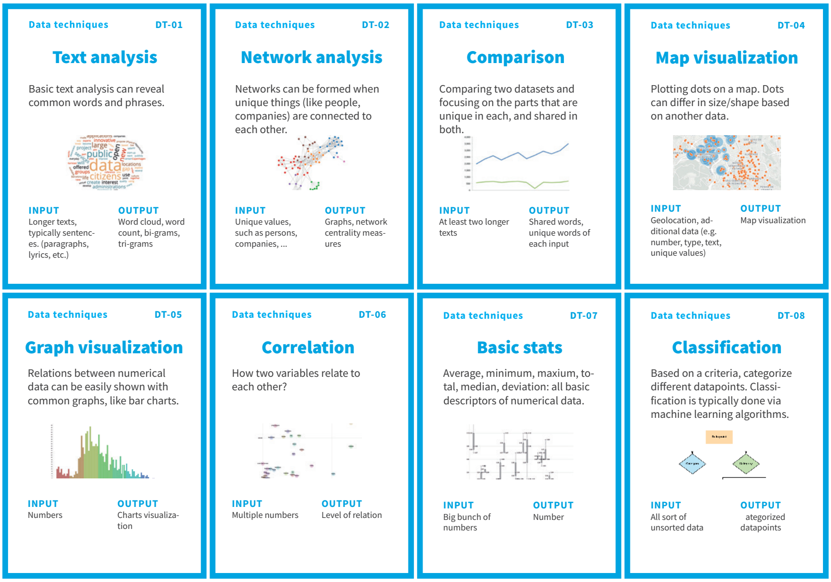


Figure 5.3. The Data techniques card deck summarizes common techniques to extract information out of data.

EXTENSIBILITY: At the core of our design rationale is to tailor the card decks to specific datasets and specific design situations. Datasets from different domains, such as metadata of library records or location coordinates of urban space artifacts, require different data exploration approaches, yet designers can face both examples. We also emphasize that the presented card decks are just initial decks that we created for the reported study in the paper, and tailoring of the card decks should be part of the design work. Furthermore, tailoring card decks can support different layers of abstractions; for example, a card deck that summarizes different visualization charts can be valuable for a dataset containing many numerical and categorical data columns. Such a bespoke card deck would provide more detailed level of visualization choices than the cards from Figure 5.3.

In the next section, we present the two booklets that we developed to complement the card decks.

BOOKLETS

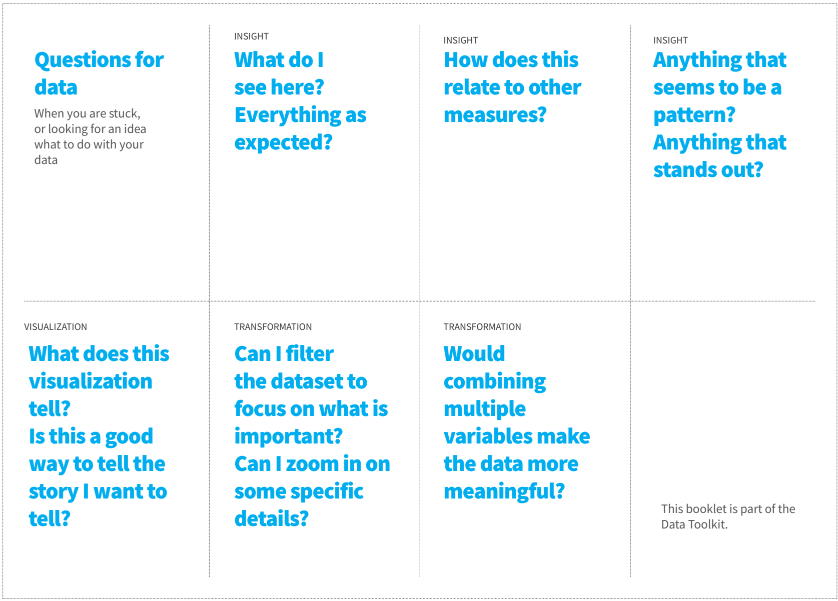
This section presents two booklets we developed for the study, and discuss the extensibility to alter and create new booklets.

QUESTIONS FOR DATA: The aim of this booklet is to guide designers to find a way to get unstuck from a confusing situation (see Figure 5.4). The booklet is based on the insight that, for the first time, it is daunting to open an unfamiliar dataset without knowing its content. The booklet contains questions hinting towards successful strategies to process the dataset and overcome the initial challenges. Depending on different situations, these questions are aimed to:

- Look at raw data and not knowing what is the next step;
- Look at a visualization and not knowing how to read it;
- Looking at data and not knowing how to extract further insights from it.

The questions in the booklet may seem trivial, but having them around in tangible format in a learning process can serve as a spark of inspiration for a sense-making process.

WORKING WITH DATA 101: The aim of this booklet is to provide a practical guide starting from the basics of opening a comma-separated value (CSV) file – the most common format to store and share tabular datasets – towards more advanced data operations on it (see Figure 5.5). The booklet is based on the insight that for a learner, there are some fundamental data operations, such as filtering or sorting data, which knowledge will be acquired early in the learning curve. Until learning these basics, it saves time during the design process to look up how to do these data operations. Furthermore, having the fundamental operations collected in one booklet emphasizes the right terminology in case of searching for further information.



Questions for data

When you are stuck, or looking for an idea what to do with your data

INSIGHT

What do I see here? Everything as expected?

INSIGHT

How does this relate to other measures?

INSIGHT

Anything that seems to be a pattern? Anything that stands out?

VISUALIZATION

What does this visualization tell? Is this a good way to tell the story I want to tell?

TRANSFORMATION

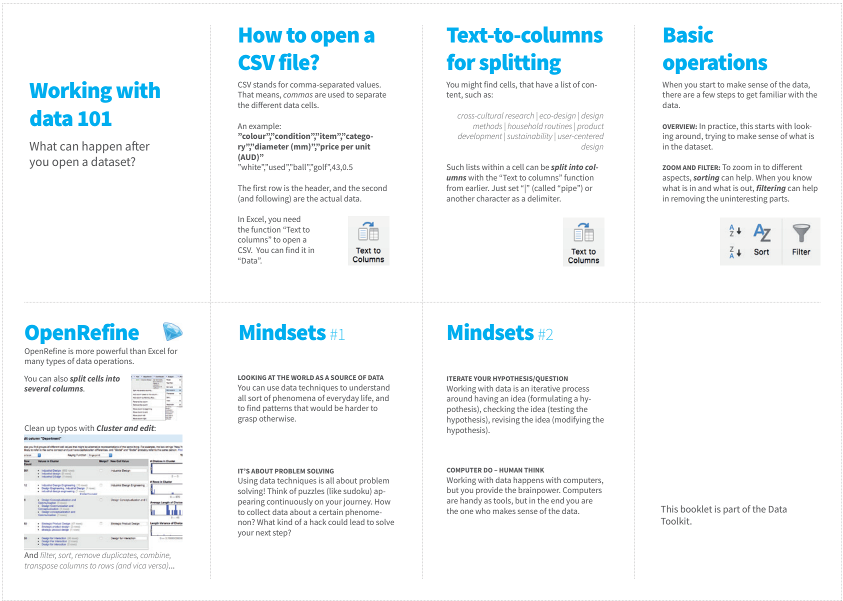
Can I filter the dataset to focus on what is important? Can I zoom in on some specific details?

TRANSFORMATION

Would combining multiple variables make the data more meaningful?

This booklet is part of the Data Toolkit.

Figure 5.4. The 'Questions for data' booklet contains triggering questions to extract insight from a dataset or visualization or to inspire the next steps of the data transformation.



Working with data 101

What can happen after you open a dataset?

How to open a CSV file?

CSV stands for comma-separated values. That means, *commas* are used to separate the different data cells.

An example:
”colour”,”condition”,”item”,”category”,”diameter (mm)”,”price per unit (AUD)”
”white”,”used”,”ball”,”golf”,43,0.5

The first row is the header, and the second (and following) are the actual data.

In Excel, you need the function “Text to columns” to open a CSV. You can find it in “Data”.

Text-to-columns for splitting

You might find cells, that have a list of content, such as:

cross-cultural research | eco-design | design methods | household routines | product development | sustainability | user-centered design

Such lists within a cell can be ***split into columns*** with the “Text to columns” function from earlier. Just set “|” (called “pipe”) or another character as a delimiter.

Basic operations

When you start to make sense of the data, there are a few steps to get familiar with the data.

OVERVIEW: In practice, this starts with looking around, trying to make sense of what is in the dataset.

ZOOM AND FILTER: To zoom in to different aspects, ***sorting*** can help. When you know what is in and what is out, ***filtering*** can help in removing the uninteresting parts.

OpenRefine

OpenRefine is more powerful than Excel for many types of data operations.

You can also ***split cells into several columns***.

Clean up typos with ***Cluster and edit***:

And *filter, sort, remove duplicates, combine, transpose columns to rows (and vica versa)*...

Mindsets #1

LOOKING AT THE WORLD AS A SOURCE OF DATA
You can use data techniques to understand all sort of phenomena of everyday life, and to find patterns that would be harder to grasp otherwise.

IT'S ABOUT PROBLEM SOLVING
Using data techniques is all about problem solving! Think of puzzles (like sudoku) appearing continuously on your journey. How to collect data about a certain phenomenon? What kind of a hack could lead to solve your next step?

Mindsets #2

ITERATE YOUR HYPOTHESIS/QUESTION
Working with data is an iterative process around having an idea (formulating a hypothesis), checking the idea (testing the hypothesis), revising the idea (modifying the hypothesis).

COMPUTER DO – HUMAN THINK
Working with data happens with computers, but you provide the brainpower. Computers are handy as tools, but in the end you are the one who makes sense of the data.

This booklet is part of the Data Toolkit.

Figure 5.5. The 'Working with data 101' booklet contains practical knowledge on how to open and manipulate a dataset in CSV format.

EXTENSIBILITY: Similar to the card decks, the design rationale of the booklets is to customize them for specific datasets and design situations. To enable this, we use the 8-pages 'fanzine' format, which enables tailoring easily. This format has also been chosen to keep the content concise and focused, possible to be printed at home and office printers and to be folded easily. Potential bespoke booklets involve the support of different steps in the design method process, such as guiding data acquisition or data cleaning.

In the next section, we elaborate on non-expert data tools to conduct *Data Exploration for Design* method.

NON-EXPERT DATA TOOLS

In practice, software tools are essential to leverage data, and therefore curating suitable software is especially important. From the perspectives of data expertise and goals with data, data journalists and librarians share a resemblance with designers. Consequently, our curation of tools has been inspired by investigating other non-expert data communities and their recommended tools. Reviewing such communities' handbooks and toolkits (see Section 2.2), we concluded the following set of criteria for software tool recommendations:

- Open source or publicly available for free;
- Available on major operating systems (or working on the web);
- Relatively easy to learn, providing a high-ceiling on functionalities;
- Supporting a non-programmatic workflow with data.

Multiple software tools match these criteria for the different steps identified in the Exploring conceptual stage. In the following, we will present our curation criteria for the core data actions:

DATA WRANGLING (CLEANING AND TRANSFORMING DATASETS): for essential operations on a dataset, we recommend common spreadsheet software, such as Microsoft Excel or Google Sheets. Spreadsheet software is widely available and often part of digital literacy education.

Such software enables direct manipulation of the data and easy sorting-filtering transformations. Furthermore, for cleaning and augmenting a dataset, we recommend OpenRefine (OpenRefine, 2020). This open-source tool provides advanced functionalities to clean and augment a dataset. While a spreadsheet software is capable of these functions as well, OpenRefine is more robust and approachable for non-experts, especially when working with non-numerical (i.e., textual) data.

DATA EXPLORATION AND DATA ANALYSIS: spreadsheet software can be used to explore a dataset and do initial explorations to understand the dataset. Choosing data exploration and analysis tools largely depend on what type of data is contained in the dataset. For the visualization of numerical or hierarchical data, RAWGraphs (Mauri et al., 2017) provides advanced charting options beyond spreadsheet software. This online and open-source tool provides superior charting options over spreadsheet software and is very easy to use. The generated visualizations can be exported in a generic vector format, enabling further editing and additional graphic design work. For design inquiry, we also envision the usefulness of working with textual data and networked data. Voyant Tools (Sinclair & Rockwell, 2012) provide an online environment to conduct text analysis, made for digital humanities scholarly research. Gephi (Bastian et al., 2009) provides an open-source robust network visualization tool, widely used by researchers, including non-expert data non-experts.

The tools mentioned above are recommended based on potential added value for design inquiry, available help online, and active communities around. However, better or more suitable tools may become available in the future. We have chosen easy-to-learn tools developed for non-experts, and thus our workshop procedure does not include formal tutorials on their use. While these tools are also capable of doing advanced data manipulation or data analysis work, such functions require further proficiency (or longer workshop formats to provide time for learning).

The following section presents an empirical study we conducted to assess the applicability of the *Data Exploration for Design* method and to inquire into the creativity support expectations when using data exploration for design inquiry.

5.4 Study 5

A pilot study with novice designers (i.e., design students) has been conducted to assess whether and how the *Data Exploration for Design* method is helpful in using data techniques as a mode of design inquiry. This section presents the methodical setup of the study, which is keeping with the method description introduced in the previous section.

PARTICIPANTS AND SETUP

Thirteen students (female, n=7; male, n=6) participated in the current study. The students could enroll in the study as a one-day elective class offering, without incentives (other than participating in a learning workshop). The students' general interest in participating was to improve data skills that can be applied in their design practice. The students were first-year master-level students from Delft University of Technology, studying different orientations of design (strategic design, n=1; interaction design/user research, n=5; industrial/product design, n=6). All thirteen participants had a bachelor-level degree in design. During the study, participants worked in duos or triads. We assumed that students with a design background would have tacit data knowledge that may inform their approach for design inquiry through data. Under tacit data knowledge, we hypothesized participants to have some familiarity with spreadsheet software (e.g., Excel) from earlier studies, and a general familiarity with general types of visualizations (e.g., charts or graphs). Prior to the workshop, participants filled a self-assessment survey on their skills, as shown in Table 5.3 (Section *Data collection* will provide more details on assessment).

Programming skills (between 1-7, 7 highest)	**Data analysis skills** (between 1-7, 7 highest)	**Technical literacy** (between 1-7, 7 highest)
2.53 (SD: 1.80)	2.46 (SD: 1.05)	3.46 (SD: 2.18)

Table 5.3. Overview of the study participants' skill self-assessment.

MATERIALS

The workshop followed the method outline and tools, as introduced in Section 5.3. At the beginning of the workshop, participants received a design brief, a dataset, suggested software tools to use as well as the design tools (card decks and booklets).

DATASET: The provided dataset was a database of the internal repository for master thesis records of Faculty of Industrial Design Engineering, Delft University of Technology. The dataset contained 2040 rows and six columns of metadata, including the theses' Title, Abstract, Mentors, or Keywords. We moderately cleaned the data to eliminate some distracting inconsistencies. All participants were first-year master students enrolled in educational programs that require to conduct a graduation project (the equivalent of a master thesis) as the final step of their degrees. As such, the provided dataset with earlier graduation projects was personally meaningful for the participants, as they will eventually face the need to define their own project, find faculty mentors for supervision, and so forth. Our intention with providing this dataset was to reduce the domain knowledge acquisition required to understand the dataset.

DESIGN BRIEF: The participants received a design brief in connection to the dataset to define three initial research questions in the context of student graduations and find answers through a data exploration process by the end of the workshop. At the end of the workshop, they were asked to present their findings in a visual format.

DESIGN TOOLS: The participants were provided with the 'Basic data types and techniques' and 'Data techniques' card decks, and the 'Questions for Data' and 'Working with Data 101' booklets.

RECOMMENDED SOFTWARE TOOLS: The participants could freely choose tools to inspect and analyze the provided dataset, but we recommended Microsoft Excel or Google Sheets, OpenRefine and RAWGraphs for anticipated needs (see more in 2.3).

PROCEDURE

The workshop was facilitated by the author as a design workshop to teach design students data competencies, as depicted in Figure 5.6.

Figure 5.6. Impressions from the workshop and the study setup.

The workshop procedure followed the earlier described outline of the *Data Exploration for Design* method (Section 5.3) as the following:

1. **INTRODUCING THE TASK:** At the beginning of the workshop, a basic introduction took place on using data in design and elaborating on a generic data workflow. After this, the participants formed groups (n=2-3). The groups received the dataset, the data toolkit, and a design brief.

2. **OPENING DATASET AND SETTING DIRECTION:** The groups were asked to download and open the dataset to initiate the inquiry process. In connection, the groups read the design brief and defined at least three questions to investigate from the data.

3. **DATA TRANSFORMATION:** The next activity was to familiarize with the dataset, using spreadsheet software or OpenRefine as a suggested software tool, and find answers for the research questions from the previous step. We expected that the questions would evolve as the dataset is continuously further explored. After providing some time for the participants to familiarize themselves with the dataset and realize that the data needs to be cleaned, a facilitator intervention was planned, by showing examples of the capabilities of OpenRefine for data cleaning, as well as a quick tutorial of RAWGraphs, the suggested visualization tool.

4. **DATA EXPLORATION:** Informed by the previous step, the following activity was to explore the dataset primarily by using OpenRefine and RAWGraphs as means to extract insights.

5. **COMMUNICATING THE INSIGHTS:** For the closing of the workshop, the groups were tasked to prepare a short presentation about their exploration process and found insights. They were explicitly asked to make it visual (i.e., present visualizations). The presentations were audio-video recorded.

After the presentations of the student groups, the workshop ended with completing a survey about the learning goals of the workshop and a Creativity Support Index questionnaire (see Data collection Section 5.4). At the end of the workshop, an audio-recorded group discussion took place to capture additional qualitative insights.

DATA COLLECTION AND ANALYSIS

In order to learn how participants used data exploration as a design inquiry method, the participants self-assessed their relevant skills before the workshop. After the workshop, a quantitative tool was used to measure the creativity support of the *Data Exploration for Design* method, as elaborated in the following.

PRIOR TO THE WORKSHOP: At the beginning of the workshop, the participants were asked to self-assess their related skills, using a Likert scale rating from '1 - strongly disagree' to '7 - strongly agree' (for results, see Table 5.3). The questions were as follows:

- My programming skills are great.
- My data analysis skills are great.
- I'm very technology literate.

DURING THE WORKSHOP: Throughout the workshop, we took notes and photos about the participants' process, and audio-video recorded the presentations and the final reflective group discussion. Furthermore, we collected the presentations the groups prepared as tangible process outcomes.

AFTER THE WORKSHOP: For (research) data collection at the end of the workshop, we used the Creativity Support Index (Cherry & Latulipe, 2014), a quantitative, psychometric tool to extract relevant insights into the mindset and expectations of the participants by assessing the design method for its creativity support for design inquiry.

5.5 Results

Results of observing the participants' processes clearly showed that exploring an unfamiliar dataset is not a straightforward task. Even though the context of the dataset was familiar for the participants, they were initially baffled how to approach inquiring the dataset to extract valuable insights for future design steps. After receiving the design brief, the design tools, and the dataset, the groups defined research questions and data hypotheses to set a direction for exploration, and then started with opening the dataset, filtering and sorting the data. After noticing the struggles with the *Data transformation* activity, a facilitator intervention happened to provide a brief tutorial on tips and tricks with OpenRefine. Our approach for facilitating the participants' learning was to let them figure the type of computational thinking required for the process first and then follow with technical tutorials. In other words, we intended to wait with a formal tutorial until 'unknown unknowns' can become more 'known unknowns'. We noted that after the initial learning curve of using new tools, the participants managed to 'zoom in' on their interests in the dataset through filtering and eliminating subsets of the data outside of their inquiry. Some groups even went further in deriving new data from the dataset, namely using the raw data they derived new data columns from counting appearances of keywords. The groups commented that they needed to shift their thinking for transforming the data, indicating their general lack of everyday practice with computational thinking.

The data transformation work was complemented with *data exploration,* for which the primary mean was exploratory visualization of the data, using charts from regular spreadsheet software and RAWGraphs, as shown in Figure 5.7. By introducing a non-expert visualization tool such as RAWGraphs, it was necessary to engage in additional data transformation steps in order to fit the dataset into formats that can be inputted into the tool. While atypical charting options of RAWGraphs going beyond the default charts from spreadsheet software were appreciated, it was also daunting to select appropriate charts suitable for different communication needs.

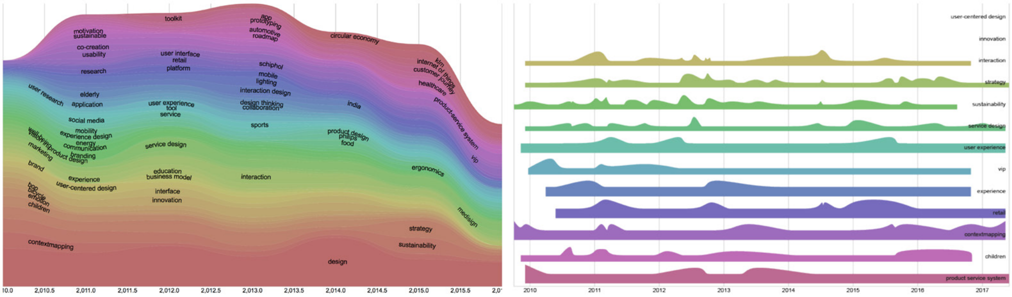


Figure 5.7. Example visualizations from the participants' exploration process. The two visualizations show the most popular thesis keywords per year.

In their process, the groups approached visualizations as 'means-to-an-end' and not as the primary output of the inquiry process.

Following the workshop learning goals of teaching a holistic understanding of data, the inquiry happened both through cleaning, transforming, filtering the data, as well as visualizing certain aspects of it. While the design brief specified to communicate their results at the end of the workshop (and for communicating it, visualizations are quite essential), but the groups did not put much effort into fine-tuning the visualizations.

In the following, we present the outcomes of the creativity support evaluation of the *Data Exploration for Design* method, providing a detailed understanding of how participants perceived the task of data exploration for design inquiry and what are their expectations regarding tools or methods supporting the task.

CREATIVITY SUPPORT EVALUATION

The CSI assessment results indicate an average of 73.85 (SD = 9.44) CSI score for the *Data Exploration for Design* method in this study (n = 13). As argued by Cherry and Latulipe (2014), such an overall score does not tell much about the creativity support performance of the given method. Nevertheless, it can be used to compare the given method with other comparable approaches.

Following the example by Cherry and Latulipe (2014, p. 21:9), in Table 5.4, we report the results with respect to average factor counts, factor score, and weighted factor score for each of the six factors. Average factor counts indicate the number of times participants chose a given factor important (between 0 and 5). In other words, this measure indicates whether the participants find such an aspect important of a creativity support tool for the specific context. Average factor scores indicate how well the *Data Exploration for Design* method scored (between 0 and 20) for the different factors. The high rankings of Exploration and Results Worth Effort indicate that participants found these two factors especially important of a creativity support tool for design inquiry. The average weighted factor scores are most sensitive to the factors that are marked more important (as average factor scores), and for both Exploration and Results Worth Effort factors, the weighted scores were rated higher than the other factors.

Table 5.4. The detailed CSI results from the study show that participants rated **Results Worth Effort** and **Exploration** factors as most important, and the average weighted score for these two categories have also been found highest.

Scale	**Avg. factor counts (SD)** (between 0-5, highest 5)	**Avg. factor score (SD)** (between 0-20, highest 20)	**Avg. weighted factor score (SD)** (between 0-100, highest 100)
Results Worth Effort	3.00 (1.78)	16.15 (1.47)	**48.85** (30.92)
Exploration	3.85 (1.07)	14.62 (1.29)	**55.85** (16.63)
Collaboration	2.08 (1.44)	14.15 (1.92)	28.46 (23.42)
Immersion	1.77 (1.42)	14.00 (2.38)	28.92 (28.15)
Expressiveness	2.31 (1.25)	13.54 (1.66)	30.46 (15.51)
Enjoyment	1.92 (1.44)	15.00 (1.27)	29.00 (21.94)

Overall, the outcomes of the CSI analysis confirm our design decisions that exploration and generating meaningful outcomes that are worth the effort are important, and the design direction is generally validated. In the next section, the results are interpreted and positioned in design literature. We particularly discuss principles of using data exploration tools for inquiry and the methodical use of data exploration in the design process.

5.6 Discussion

Modern everyday life is increasingly facilitated or recorded through large data infrastructures, and at the same time, data is becoming increasingly accessible and present in design practice. Access to data about human experiences shows the potential for gaining new understandings about phenomena other design methods would miss. With the Exploratory Data Inquiry methodology, we proposed to expand designers' repertoire to methodologically use existing data as well as existing data tools for design inquiry. We developed the Data Exploration for Design method by elaborating on established practices and tools from other non-expert data communities. We evaluated the design method for its ways of creativity support, and the outcomes revealed that **Exploration** and **Results Worth Effort** are key characteristics of using data exploration in the context of design inquiry. While a high score of Exploration is not surprising in the context of investigating data exploration as a way of design inquiry, Results Worth Effort can be further interpreted. For interpreting the results, we consider the novelty factor of using data exploration and the promise of gaining a previously-hidden perspective on a particular phenomenon. Since designers are rarely trained in data science techniques, the learning curve needs to be taken into account with the novelty of the approach. Similarly, data science techniques promise access to insights and perspectives of phenomena that otherwise would be hard to extract with more traditional design inquiry approaches. Interpreting these for the current study, generating results that were worth their effort partially acknowledged the learning curve of the different techniques and the unfamiliar approach. With the learning curve in mind, the participants found that the insights that can be gained even with tools that are unfamiliar, hard to use, or not fully designed with a designer workflow in mind, are valuable. In other words, the generated results were worth their effort.

DATA EXPLORATION AS DESIGN INQUIRY PRINCIPLE

The interpretation of the results highlights the importance of the selection of what tools and techniques to use in conducting data exploration as a design method. Previously elaborated in Section 5.3, we selected software tools inspired by other non-expert data communities, such as data journalists, librarians, and digital humanities scholars. Those tools are designed for non-programmers working with data, and therefore are suitable for lowering the learning curve threshold effectively while providing functional capabilities to gain new perspectives about a given dataset. However, what would be the requirements for future tools that are made for the specific needs of designers using data creatively? In this context of designing future data exploration tools and methods, Dalsgaard (2017) offers a more general framework for "instruments of inquiry". This framework considers five main qualities of instruments of inquiry: perception (revealing and hiding facets of a design situation), conception (develop and hypotheses about a design situation), externalization (make imagined design solutions), knowing-through-action (generating knowledge by acting with an instrument), and mediation (mediate between actors and artifacts in a design situation). We will use Dalsgaard's (2017) framework to interpret the findings of the study. Using his framework, we distill a set of principles that can make designers repeatable value in using data exploration as a design inquiry method:

1. **ACKNOWLEDGE BIASES IN DATA COLLECTION:** Designers using data exploration as a design inquiry method need to be aware and perceptive what aspects the data collection shows and hides about a problem area, both working with existing data or when defining what data to collect. Data acquisition can carry built-in biases and limitations, which skew the inferences that can be obtained from the data.

2. **SPEND TIME WITH THE DATA:** There is immense value in spending time exploring the data as a way to build contextual knowledge about the design situation. While entering a new domain, having access to a dataset may speed the initial process of building up the domain knowledge somewhat quicker, in longer time-frames, knowing the dataset and the domain intertwines.

3. **VISUALIZATIONS ARE A MEAN-TO-AN-END:** In the process of working with the data, representations such as visualizations have a two-folded function. First, they help human cognition to understand and contextualize the data, and second, they become shareable units of design work that can be used with other actors. As such, the goal of design inquiry is not to craft a visualization, as opposed to information design and communicating findings.

4. **BE PART OF DATA COLLECTION:** Spending time with the data in the design process means a continuous co-evolution of learning the problem space (Dorst & Cross, 2001). In the unfolding process, new research questions and hypotheses emerge that might not be possible to answer from the initial dataset. Consequently, designers should be involved

in the data collection to be able to iterate on exploratory data inquiry, either hands-on setting up data collection themselves, or defining in-detail what data to collect.

These principles are domain-general and inform the agency of designers both for using data exploration as a design method by themselves or improve their collaboration with data experts.

DATA EXPLORATION IN THE DESIGN PROCESS

Data exploration as a mode of design inquiry seems feasible and valuable to be used in the design process to 'access' the data footprints of human experiences, but when is it a reasonable choice to use data exploration? By comparing data exploration as a design inquiry to other design techniques, more informed choices can be made to suit a specific design situation. Sanders and Stappers (2014) provide an overview of different generative, i.e., exploratory design research techniques. In their work, they primarily compare two modes of *designing with* or *designing for* the users while elaborating upon the traditions of probes, prototyping, and toolkits for the early phase of design. Within this perspective, we can position the *Data Exploration for Design* method primarily in the generative phase of design, following a 'designing for' mindset (see Figure 5.8).

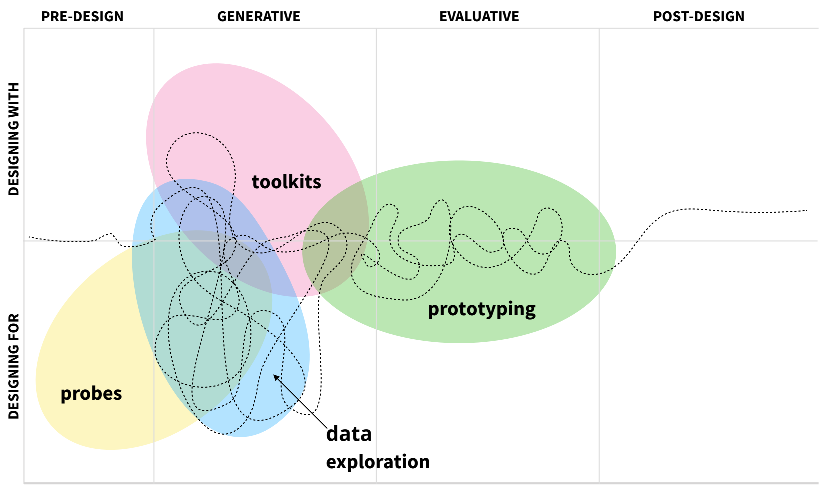


Figure 5.8. The Data Exploration for Design method placed in the co-design map of Sanders and Stappers (2014).

Other data-centered design approaches have also emerged in this space. Bogers et al.'s (2016) data-collecting technology probes have shown novel ways to gain rich and contextual data using sensors. In follow-up work, their approach has expanded out to probes, toolkits, and prototyping (Bogers et al., 2018; van Kollenburg et al., 2018). Similarly, Giaccardi et al.'s (2016) Thing Ethnography was an exploration of ethnographic inquiry through equipping everyday objects with a camera. A camera in this context becomes a data-collecting sensor that is capable of rich data collection. The approaches of these examples share similar technological complexity, often beyond the scope and resources available for a design team. Contrary to the examples, our current approach has focused on using existing data (such as inquiring data from a data infrastructure) and supporting designers to learn the necessary skills to process the data and guide how to extract value out of it. For the approach of using existing data, the main concern is not how to acquire the data, but how to look at the data. Practically, 'looking at data' is informed by the findings of design theory and methodology. First, as Cardoso et al. (2016) highlight, questions are at the core of inquiry and questions make designers explicitly formulate interests from a dataset. Second, designers are opportunistic and use different methods for different inquiries (Guindon, 1990). In conclusion, designers can use other types of inquiry, such as qualitative research, to complement their gained insights from the data exploration process.

LIMITATIONS

The current study aimed to investigate the development of a design inquiry method for data exploration. We derived a design method, referred to as *Data Exploration for Design* method, based on literature and informed by our earlier work, and then evaluated the resulting design method as a one-day learning workshop. A primary limitation of the current study is that the *Data Exploration for Design* method was evaluated with a group of design students (n=13) and by our facilitation. While we see the value for data exploration as a design method for designers across levels of expertise, yet the study participants were master design students. Although master-level design students are quite tech-savvy and data literate, they might not be representative of the whole design profession or designers working in less technical domains. It is also important to note that the study's design brief and provided dataset (metadata of graduation thesis records) set up a limited problem space with specific properties, which does not model all sorts of potential design contexts. With these caveats, it is difficult to assess the effectiveness and added value of data exploration as a design method outside of academic learning environments at the current stage. Overall, the study contributes a design method for data exploration and a set of principles for using data exploration for design inquiry.

5.7 Conclusions

It can be concluded that the *Data Exploration for Design* method enables designers to use data exploration for design inquiry. We outlined a method based on three conceptual stages of *Problem framing*, *Exploring,* and *Inferring*, which stages guide data exploration in different design situations. We also developed two sets of card decks and booklets to support the learning curve of the method and as a way of developing holistic data competence. These card decks and booklets are tailorable and extensible for different design situations and datasets. In the current chapter, the method was evaluated during a workshop and was proven useful in exploring data and in generating valuable outcomes for the design process. Furthermore, using the method contributed to the participants' holistic data literacy, informing how to use data exploration for design inquiry creatively. The study results enable us to extract a set of principles that describe the core mindset of the *Data Exploration for Design* method.

LESSONS LEARNED FOR DEVELOPING METHODOLOGICAL CONTRIBUTIONS

This study shows how a design method and accompanying design tools can structure design inquiry through data. For the goal of developing methodological contributions, the primary contribution is the *Data Exploration for Design* method. From the study, we deepened our understanding of the conceptual stages, as shown in Figure 5.1, and on distinguishing between design and software tools, as shown in Table 5.2. Figure 5.9 shows how the iterating conceptual stages support moving from unknown to known during design inquiry.

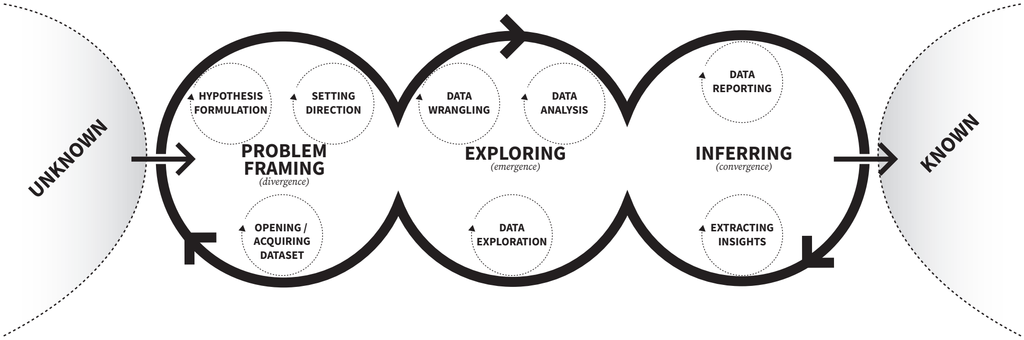


Figure 5.9. The Data Exploration for Design method can illustrate the iterative nature of data exploration in design inquiry.

TOOLS: With the study, we identified two dimensions for tools. One dimension is the divide between design tools and non-expert data tools, see Table 5.2. The former tools support the cognitive aspects of designing, how to think about data. The latter tools are the actual tools to operate on data. The second dimension of tools targets the learning aspects. An important function of design tools is to support the learning curve of the method and to provide scaffolding on how to integrate the method into one's mindset.

MINDSET: Connected to the tools, the design tools especially form how mindset is intertwined with the tools to conduct the process through. Furthermore, based on this study, we also deepened our understanding of how holistic and deep designers should be involved in data collection and spending time with data. Being part of data collection is necessary to infer correctly from the dataset, and in that regard to know what is in the dataset; especially to be able to acknowledge biases. Furthermore, spending time with data cannot be overlooked, the least for building domain and contextual understanding, suggesting that design inquiry through data might not be a 'quick' approach applicable in any given use-case.

PROCESS: Compared to the previous study, the iterative aspects throughout the three conceptual stages are more prominent. Furthermore, the intertwining of design and non-expert tools may also impose additional 'shifts' in the process, changing from one type of tool to the other throughout the design inquiry process.

The previous chapters presented empirical studies that revealed how the appropriation of data science practices happen in design practice (Chapter 3) and then how data science practices can be integrated into design practice through a creative process lens (Chapter 4). In current chapter we built on the previous findings and motivated a design method based on the *Exploratory Data Inquiry* methodology. In the next chapter, we will investigate the adoption of the *Exploratory Data Inquiry* methodology in a frame innovation workshop setup that increasingly resembles real-world design practice.

Chapter 6

Embedding Exploratory Data Inquiry into Frame Innovation

In the previous chapter, we developed and evaluated the Data Exploration for Design method as an approach for data exploration for design inquiry. The design method builds on earlier investigations of how data science practices can be integrated creatively in the design process. In this chapter, we contextualize our earlier empirical work in a more realistic design case. In this chapter, we address RQ5 *of the dissertation, "How do designers adopt Exploratory Data Inquiry in design practice?" In order to answer this question, we will integrate Exploratory Data Inquiry methodology into a frame creation setting, while using a dataset that also resembles 'big data' more closely than previous studies. In the chapter, first, we elaborate on Frame Innovation and existing research methodologies using social media. Then we present an empirical study and discuss our findings on how the mindsets of the participants changed while intertwining a computational approach with qualitative design inquiries.*

The study participants contributed to the description of process in Section 6.4.

6.1 Introduction

The previous studies highlighted that designers are able to integrate data exploration with the core practices of design, such as framing and reframing. Earlier in Section 3.5, we highlighted the characteristics of design inquiry through data of asking suitable questions as part of creatively using data for exploration. Then in Section 4.7, we expanded on this notion focused on how the opportunistic and abductive logic of framing and reframing in design takes place while design inquiry through data. From the previous studies, it was clear that framing and reframing take an unusual role for design inquiry through data, given that they are conducted through tools and techniques that have been made for analytical work using deductive and inductive logic. Following these insights, in the current study, we concentrate further on framing and reframing by *directly* positioning exploratory data inquiry into the overarching design methodology of Frame Innovation (Dorst, 2015b), a formalized methodology of framing and reframing. As the beginning of the dissertation started from designers' need to contribute to solving complex problems with a special skillset unique to designers, the successful combination of Frame Innovation with *Exploratory Data Inquiry* methodology would also mean a potentially new approach for transdisciplinary collaborations between designers and data scientists. To this end, our second consideration for the current study is to 'stress test' the *Exploratory Data Inquiry* methodology in a more contextualized and realistic design situation. Compared to the previous studies that all featured learning workshops either as one-day stand-alone workshops or three-days workshops part of a larger design brief, the current study contains a longer workshop-setting with a significantly larger dataset.

The current chapter investigates **RQ5**, *"How do designers adopt Exploratory Data Inquiry in design practice?"*. In order to answer this, we set up an empirical study that resembles a more realistic design situation than the previous studies. This study is aimed to generate insights towards generalizing the results of the research and thus features a more extended research setup, increased magnitude of data, and assessing the feasibility of how Exploratory Data Inquiry methodology can be intertwined with a Frame Innovation.

The dissertation started with designers increasingly engaged in tackling complex problems (see Section 1.1). In the current chapter, we will use such a design brief and approach that can illustrate complex socio-technical problems, as a way to return to the scale of the initial problem statement. Following a design brief on 'student mental well-being' and an approach that is using the collection of over 11 million tweets about the topic will provide a more contextualized and realistic design situation to explore the research questions.

STUDY RATIONALE

We selected frame creation for our study as a 'container' design methodology for three main reasons. First, frame creation is an overarching design methodology that is used by incorporating different types of design inquiry, and therefore providing space to use our studied mode of inquiry, exploratory data inquiry. Second, we immersed in the core part of frame creation of finding themes and generating frames and ideas, a step that resembles the concept creation of most design activity. Third, frame creation has been used in transdisciplinary settings (Bijl-Brouwer, 2019), and a successful way of incorporating different disciplines into a large design process. We expect that frame innovation could suit incorporating data practices into designing.

We selected social media inquiry about the problem area of 'student mental well-being' as a wicked problem, that is relatable for the participants, a complex socio-technical problem in general, and there is a possibility of capturing a large set of data about it. Participants' relatability was important to enable a quick learning curve about the domain, and focus the learning curve not topically, but on the tools and methods involved in the study. Studying a complex socio-technical problem provided a large solution space, enabling the framing and reframing process to focus on finding the 'right' problem. Last, media and social media discourse on student mental well-being have been rich, enabling the capture of a large number of tweets, although with certain biases such as tweets in English language.

Next in the chapter, we will elaborate on Frame Innovation as a design methodology for reframing complex problems towards problem and design spaces that are more 'fruitful'. We also provide background on inquiry through social media, as tweets are the primary data source for tackling the design situation. Afterwards, we present a case study where we observe a design team working on the design brief using the design approach mentioned above. We analyze the process of the designers in the study and then elaborate on the results. The study results provide further nuances about the key characteristics of the mindset of designers using data exploration as a primary mode of design inquiry. In Chapter 7, we analyze the different studies altogether and extract some more in-depth underlying findings to answer the research questions of the dissertation.

6.2 Background

In the following, we elaborate on frame innovation as a design methodology and research inquiry uses of social media.

FRAME INNOVATION

Framing practices are an essential component in designing. Schön (1984) introduced the concept of problem-setting, as an essential element for reflective practices, like design. As he stated, *"Problem setting is a process in which interactively we name the things to which we will attend and frame the context in which we will attend to them"* (Schön, 1984). Dorst and Cross (2001) studied the practices of expert designers, leading to their influential theory on the co-evolution of problem and design space. As part of this co-evolution, framing is a key practice to generatively synthesize learnings from the solution space to evolve the problem space. Dorst has expanded on this theory focusing on design problems and paradoxes (Dorst, 2006), as well as the cognitive parts of how abductive thinking of designers fuels framing practice (Dorst, 2011). Commonly argued by Dorst, the key characteristic that distinguishes expert designers from novices is the increased capacity to frame and reframe problems in constructive ways to move the design process forward. Dorst recently organized the learnings into the book *Frame Innovation* (Dorst, 2015b), in which he presents a methodology for *frame creation*. Frame creation is a nine-step process, during which many steps resemble design processes common human-centered design, however refocused to make the different steps adequate for addressing open, complex, dynamic, and networked problems. In practice, frame creation as an overarching design methodology can take years and can contain and incorporate different types of design inquiry throughout many rounds of iterations. Interviews, co-design session, and other design techniques are commonly embedded in frame creation processes. In the following, we elaborate on the steps of the frame creation process. We will especially highlight the steps in the middle of the process (5) Themes and 6) Frames), as these are the steps where we intended to spend the most time during the current study.

FRAME CREATION PROCESS

Dorst (2015a) codified frame creation as a nine-step procedure. In Table 6.1, we illustrate the nine steps and annotate them with common activities in design. The first steps out of the nine are common in human-centered design practices of generating an overview of the problem and mapping the stakeholders' needs and perspectives.

Table 6.1. The nine-step frame creation process, with our additional annotation, based on Dorst (2015a).

Step	Description	Own annotation
1) Archeology	Analyzing the history of the problem owner & the initial problem formulation	Background
2) Paradox	Analyzing the problem situation: what makes this hard?	Background
3) Context	Analyzing the inner circle of stakeholders	Stakeholders
4) Field	Exploring the broader field	Stakeholders
5) Themes	Investigating the themes that emerge in the broader field	Research, ideation
6) Frames	Identifying patterns between themes to create frames	Research, ideation
7) Futures	Exploring the possible outcomes and value propositions for the various stakeholders	Ideation, Concepting
8) Transformation	Investigating changes in stakeholders' strategies and practices required for implementation	Implementation
9) Integration	Drawing lessons from the new approach & identify new opportunities within the network	Implementation

FROM 1) ARCHEOLOGY TO 4) FIELD

The first four steps of the frame creation process are similar to a generic human-centered design process. In the steps of 1) Archeology and 2) Paradox, the background of the problem is summarized, including earlier approaches that have been attempted to solve it, and a focus on the 'paradoxes' why this problem is so hard to tackle. The following is to investigate the 3) Context and 4) Field or, in different words, the stakeholders of the problem. First, the context of the problem is investigated by summarizing direct stakeholders to the problem, and second, the broader field of the problem is explored, thinking of other stakeholders that could be part of the solution space. As a next step, the needs of the stakeholders are investigated.

5) THEMES

Dorst coined the fifth step of frame creation as 5) Themes. The core of frame creation is extracting patterns of the problem and stakeholders' needs as 'themes', and then to generate 'frames' as a solution to the themes (step 6). Dorst turned to hermeneutic phenomenology to capture 'phenomenological themes', or in other words, to capture the structure of a human experience (Van Manen, 1990). In practice, this takes place as a design team doing thematic analysis of qualitative data, which is a common process of how designers process their research and move towards ideation. In different words, thematic analysis is the modeling of a human experience based on qualitative research. In practice, thematic analysis often result in the generation of high-level concepts connected with arrows as diagrams and similar simple visualizations to represent the patterns that play a role within the given problem domain.

6) FRAMES

Once themes are fleshed out from the starting problem domain, the actual creation of frames takes place in 6) Frames. For frame creation, the following formula can be used (Dorst, 2015b, p. 78):

> *"If the problem situation is approached as if it is about: _____A_____ (stakeholder x,y,z) wanting doing, then we need _____B_____".*

A) is a reframing of the initial problem, and B) is then an initial design space to address A). For example, in the context of addressing students' mental well-being and academic stress, one frame formulation could be:

> *"If the problem situation is approached as if it is about '<u>international students having to live up to expectations</u>', then we need to '<u>rationalize metrics of success</u>' or '<u>create a safe environment for conversation</u>'".*

Frame creation process facilitators have a 'trick' when the 'If... then...' template is not leading towards 'fruitful' frames. By the deconstruction of a **metaphor**, the frame creation is inspired by an analog problem. Metaphors are often used as a design method (van Boeijen et al., 2020, p. 161), for instance, Schön (1984) also introduced the concepts of 'generative metaphors', as a way to describe the reflective design processes of architects. In the context of frame creation, metaphors are used as a way of 'dissecting' a theme by an analogous problem. In Table 6.2, we illustrate a metaphor that we used in the context of student mental well-being (not the study reported in the current chapter).

Table 6.2. For the problem formulation of 'master students trying to satisfy their knowledge hunger', using the metaphor of an 'expedition team discovers uncharted territory'.

Metaphor: Expedition team discovers knowledge	**Solution**
Interdisciplinary team	Various roles, shared responsibility, different hats to wear
'Flag'	Mascot, badge, outfit, something to leave as a mark (emotional value)
Gear: compass, map	Methodology/roadmap (where you are and where you are going)
Journal	Notebook, Book with cards to share
Need to improvise	Role-play, protocol, quick thinking, recognizing patterns

In this process, a metaphor is selected that has a connection to the theme that is being addressed but otherwise having wider connotations. For example, themes addressing 'teamwork' can use the metaphor of 'olympic sports team', and so forth. The designers dissect the used metaphor to identify the key working mechanisms in the metaphor and map the 'solutions' that these identified mechanisms are typically identified with. After a metaphor is dissected, the process returns to the frame creation, and using a working mechanism of the metaphor and the solution, frames are being created in the 'If... then...' template.

FROM 7) FUTURES TO 9) INTEGRATION

After frame creation, solutions and ideas are being generated first during step 7) Futures. This step is the classic 'conceptualization' step in any design work, when value propositions, ideas, or concepts are prototyped and tested. Step 8) Transformation and 9) Integration are concerned with embedding the solutions into the context. Since solutions for complex socio-technical problems can concern multiple stakeholders, multiple service touchpoints, and multiple products or interfaces, the importance of scaling up the solutions is not negligible. However, as they are concerned more with tactical design work than the actual frame creation, we not focus on them in further detail.

Next, we motivate our choice for social media inquiry and provide background on the ways how social media is used for inquiry.

SOCIAL MEDIA INQUIRY

Although inquiry from online sources has become common in certain scientific fields, it has sparsely been used in the field of design in general, and particularly for design inquiry. What people do and say on online platforms, such as on bulletin boards or forums decades ago, social networks today have been used as a data source by anthropologists for ethnographic research (Pink, 2016). More specifically, with the growing presence of social media in everyday lives, new methods emerged using social media data. In the following, we will present so-called digital methods as well as big data approaches for social media research.

DIGITAL METHODS: Digital methods refer to social research methods based on repurposing existing data from the internet (Rogers, 2013). Researchers using digital methods refer to data as 'digital-first', such as social media, websites, or links, explicitly not referring to digitalized data, that otherwise could exist in an analog format, like survey responses. Research protocols, tools, and methods that have arisen with digital methods research are showing an approach that is not big data research per se, as the focus is not on the sheer large amount of data collected and analyzed, but the emphasis is put on how the data acquisition is rationalized and argued, often resulting in ad-hoc designed datasets (Marres & Weltevrede, 2013). Our earlier studies using digital methods techniques underlined the importance of query design, which we have addressed earlier in Section 4.7. Researchers using digital methods have developed their research protocols, tools, and methods (e.g., Digital Methods Initiative – Tools Database, 2020), and their developed approaches have been widespread in the fields of digital humanities, political sciences and beyond. Non-expert tools based on their research protocols, such as the DMI-TCAT system (Borra & Rieder, 2014), provide a robust way of collecting tweets following certain hashtags, users, or a specific geolocation.

BIG DATA SOCIAL MEDIA RESEARCH: Social media research through big data, and more specifically the use of Twitter has long been investigated in human-computer interaction, both by interest of certain use-cases, such as emergency crisis response (Bruns & Liang, 2012), but also to understand details about social connections through social media, e.g., strength of connections (Gilbert & Karahalios, 2009). There is also a large set of studies that problematizes social media based inquiry by focusing on bias in inferring from social media (Hargittai, 2015), or ethical considerations using social media data (Tiidenberg, 2020).

More directly related to our approach in the current study, Brooker, Barnett, and Cribbin (2016) present a four-quadrant framework to categorize the types of data capture and data analysis, as a summary of the state of the art of big data research *methodologies*, as reproduced in Table 6.3.

Data capture / data analysis	Temporal analysis (event based)	Corpus analysis (topic based)
Semantically driven (query keyword)	How does a narrative about a semantic entity (i.e word, hashtags, etc.) unfold over time?	How is talk around a semantic entity organized topically (and sub-topically)?
User driven (user following)	How do users' language and tweeting practices change (or not) over time?	What topics are a specific group of users tweeting about (and how are they doing it)?

Table 6.3. Framework of social media inquiry by Brooker et al. (2016).

Brooker et al.'s framework demonstrates four distinct types of social research questions that inform both the way the data is captured and the way(s) how analysis can be conducted on the data. This framework can inform inquiry around events or topics and around following queries or users. Data analysis around an event can reveal how a phenomenon unfolds over time. Data analysis around a topic can reveal details about a discourse. Accordingly, data capture based on keyword queries allows analysis on a temporal or topic-based investigation, while data capture based on users can reveal how their communication changes over time or the topics they talk about. This framework informed our data capture to use a semantically-driven, query-based approach instead of following users. This was decided based on the context of the study, that had topics to follow around student mental well-being. We intended not to delimit the ways of data analysis and thus focused on capturing data from a long-enough period to explore the dataset either based on temporal or on corpus directions. In the next section, we present the study.

6.3 Method

To address the research objective about investigating to what extent can exploratory data inquiry be applied in the design process and what are the key components of mindset designers assume while using this approach, we set up a study. The study took place as a 5-days long design sprint pressure cooker, where we observed two design teams who were tasked to use exploratory data inquiry in a frame creation process. The participants could sign up voluntarily for the study. The study was advertised as an experimental workshop, with a learning component to learn about Frame Innovation methodology in practice and research techniques with big data.

CONTEXT OF STUDY

The context of the study was chosen to enable a design brief that is a wicked problem *per se,* a complex socio-technical domain from where a wide range of solutions can emerge, and also one that is relatable for the participants. With these considerations, we chose 'student mental-wellbeing' as the problem context. Mental well-being on campus has been in the attention of media both locally as well as around the Netherlands and the world. Academic pressure in today's world is a real problem that affects both local and foreign students for different reasons. Several stakeholders in this problem could be part of the solutions. While campuses propose old solutions in larger quantities available for the student body, such as more psychologists available, there is also space for innovation and novel thinking.

The participants' task was to look into the problem area student mental well-being through the lens of collected tweets over a two-month period (over 11 million tweets), and using frame creation and exploratory data inquiry techniques. Over a five days workshop, they learn about frame creation (day 1), working with data and exploratory data inquiry (day 2) while working on the case, and working in an iterative design research process until day 5, when presenting their results to an audience, see the schedule in Table 6.4. The results were expected to be design concepts exploring a 'fruitful' frame within the case, or one more developed design concept.

The workshop was facilitated by the author, with background in computation, design, and facilitation. Furthermore, different experts were invited to contribute at different points. The facilitation design of the workshop was developed together with an expert in frame creation facilitation. On the second day, a data scientist researcher taught a tutorial on basic data techniques. On the third day, a data steward joined for a conversation about the ethical use of data. The overall workshop approach was based on the approach of studies reported in the previous chapters.

Table 6.4. Workshop schedule of Study 6AB.

	What happened?
Day 1	› Introduction to NADI model. › Introduction of the topic about student mental wellbeing. › Group brainstorming (both group 6A & 6B) about the context and stakeholders. › Each group picked four stakeholders then did a brainstorming session to map what are the motivation behind those stakeholders. › Each group narrow down to two themes and come up with relevant metaphor. › Each group ideate on solutions using the NADI model template.
Day 2	› Introduction to big data. › Introduction to ethics in big data. › Introduction to the technical aspects of big data analysis. › Introduction to the tools to analyze big data.
Day 3	› Main dataset was provided, consisting of 11 million tweets. › Each group analyzed the data based on their theme and metaphor from day 1.
Day 4	› Each group come up with the solution based on big data analysis and fill in the NADI model template. › Each group start to make a presentation about the process and the generated concept.
Day 5	› Presentation. › Group reflection.

PARTICIPANTS

Five students (4 female, 1 male, average age = 25) participated in the study. The students were second-year or recently graduated master students from Delft University of Technology in different orientations of design (strategic design, n=1; interaction design/user research, n=2; industrial/product design, n=2). Participants had bachelor's degree in industrial design (n=4) or mechanical engineering (n=1). The participants had an average of 16 months (SD=12.90, min=4, max=30) of professional design experience (including internships). Details about the participants' self-assessed background on programming, data analysis, and technical literacy are in Table 6.5.

Table 6.5. Overview of the study participants' skill self-assessment.

Programming skills (between 1-7, 7 highest)	**Data analysis skills** (between 1-7, 7 highest)	**Technical literacy** (between 1-7, 7 highest)
1.6 (SD: 0.55)	1.4 (SD: 0.55)	4.2 (SD: 1.10)

Furthermore, the participants could select a descriptor of their skill assessment of programming and data analysis, as shown in Table 6.6.

Table 6.6. Overview of programming and data analysis skill assessments of participants.

Programming skills level overview		**Data analysis skills level overview**	
None, starting to learn	0	None, starting to learn	0
Novice: learned programming in a course, but haven't used my skills outside that since.	4	Novice: learned some Excel in a course, but haven't used my skills outside that since.	2
Beginner: I have used programming outside coursework, but not much.	1	Beginner: I use Excel occasionally.	3
Advanced beginner: e.g., I have applied programming in design projects or side projects.	0	Advanced beginner: e.g., I know Excel functions, or studied/used R, or other statistical tools.	0
Intermediate: I'm comfortable with programming, doing it weekly.	0	Intermediate: I'm comfortable with spreadsheets or R. I know statistics fairly well.	0
Advanced: I'm very comfortable with programming, doing it daily (or studied programming extensively / worked as a programmer).	0	Advanced: I'm very comfortable with data analysis, have done it daily (or studied it extensively / worked as an analyst).	0

DATASET

The prepared dataset was created through two months of Twitter data capture using DMI-TCAT (Borra & Rieder, 2014) from Twitter's Streaming API. We used a data collection following keywords of: *'mental health', 'adhd', 'anxiety', 'depression', 'mentalhealth', 'mentalhealthawarenesscare', 'mentalhealthmatters', 'mentalillness', 'stress'*. These keyword-based searches returned both #hashtag and non-hashtag versions of the same word (thus, 'stress' and '#stress' both were collected). The period of data collection was between 5 June 2019 and 5 August 2019, as well as shown on the timeline in Figure 6.1.

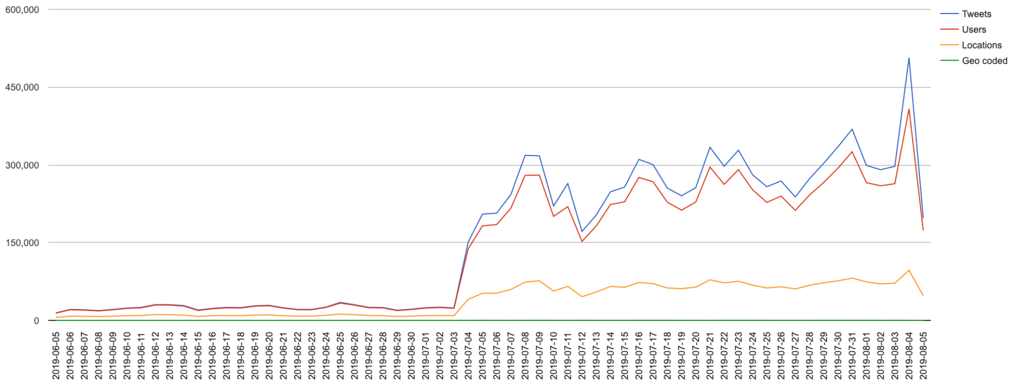


Figure 6.1. Timeline of tweets per day.

Overall, 11,852,309 tweets were collected during this period. The data collection took place prior to the study, therefore, the participants worked with historical data.

DESIGN TOOLS

The Needs and Aspirations for Design and Innovation (NADI) model (Bijl-Brouwer & Dorst, 2017) describes four levels, which are *solutions*, *scenarios*, *goals,* and *themes* connected over the same purpose in the context of design (see Table 6.7).

Table 6.7. The NADI model as described by Bijl-Brouwer & Dorst (2017).

SOLUTIONS	What do people want or need? Which products, services or interventions do people want or need?
SCENARIOS	How do people want or need to interact with the solution in the context of use?
GOALS	Why do people want to interact or behave in a certain way? What do they want to achieve within the context of the problem?
THEMES	What is the underlying structure of the experience? What are their meanings and values outside the direct context of the problem?

The model can be used to position deep insights to indicate 'how deep' those insights are. This model was used as a template for ideation and clarifying concepts. The benefit of using the NADI model is the independent analysis of themes, as the underlying needs and aspirations, which supports framing.

SOFTWARE TOOLS

On the second day of the design workshop, a few data exploration tools were introduced that were based on the curation from the previous chapters. These included Voyant-Tools for text analysis (Sinclair et al., 2018), Gephi for network graphs (Bastian et al., 2009), RAWGraphs for non-trivial charts (Mauri et al., 2017), Google Sheets and Excel for regular data jobs, as well as OpenRefine for advanced data cleaning and wrangling.

RESEARCH DATA COLLECTION

AT THE BEGINNING OF THE WORKSHOP: the participants were asked to fill a questionnaire about their backgrounds on programming, data analysis, design, as well as collecting demographics data, such as age, educational background, and expectations from the workshop. The questionnaire also featured a Likert-scale rating survey on self-assessing their confidence in skills and to what extent are they comfortable with programming, data analysis, and technology literacy in general.

DURING THE WORKSHOP: we audio-video recorded group conversations and presentations to monitor and document how the design process unfolded. Pictures were taken about the different artifacts created throughout the workshop, primarily whiteboard and paper sketches, or affinity diagrams with post-its.

AT THE END OF THE WORKSHOP: public presentations were held by the participants about the output of their work as well as reflecting on their approach (and thus, the workshop week), which was audio-recorded. The last afternoon of the workshop week ended with a focus group interview to retrospectively look at the week, and get the group collectively reflect on their experiences.

AFTER THE WORKSHOP: three participants engaged in writing a report about the workshop (to complete the requirements for educational credits for the workshop as a research project), and their report was also used as data collection to triangulate the processes that took place throughout the week.

The workshop was also followed with a questionnaire to self-assess the participants' skills development, and the Creativity Support Index (2014) instrument was used to inquire about the participants' experiences and assessment of the *Data Exploration for Design* method to be used in the context of a Frame Innovation workshop.

RESEARCH DATA ANALYSIS

The audio-video recording of the focus group interview was transcribed and qualitatively coded following an open coding protocol. The pictures and video clips throughout the workshop were used to triangulate the design process of the two teams of the participants, which was described in their report as well.

6.4 Results

In this section, we first provide a rich process description of Study 6A and 6B, in order to depict the challenges and thought-processes the participants followed throughout the workshop. After describing the two studies, we compare them and present the results of the creativity support assessment of the overall approach.

DAY 1 - LEARNING ABOUT FRAME CREATION

The goal of the first day was to let the participants practice the frame creation process, and use the first four steps of the process (namely, 1) Archeology, 2) Paradox, 3) Context and 4) Field) as an immersion into the problem brief. During these steps, the five participants worked together. Following the first two steps, they immersed into the problem domain of student mental well-being and summarized the stakeholders directly and broadly involved in the problem area, also as a way of sensitizing for the upcoming steps. After this point, the participants split into two groups and continued working in these two groups until the end of the workshop (the groups will be referred to as

6A and 6B from now on, number six referring to the current chapter number). To continue the frame creation process, both groups first selected three stakeholders to focus on and then assessed the selected stakeholders' needs. These identified needs were the basis of starting with step 5) Themes. The groups searched for themes over the needs of stakeholders and enriched the themes through concept maps from their own domain knowledge of being students. During the afternoon of the first day, the groups transitioned to step 6) Frames. They used metaphors as a tool to 'dissect' analogous problems to their themes and then used the metaphors as the basis of frame creation. Both groups were introduced NADI templates to capture frames and early concepts (towards step 7) Futures).

At the end of the first day, the groups summarized their processes with short presentations for the other team, and these presentations were also recorded.

DAY 2 - LEARNING ABOUT DATA

The goal of the second day was to develop the participants' competence and confidence in working with data. During the day, the participants received tutorials about necessary technical competences, but also about ethical usage of data and then using data for design inquiry. First, the participants were guided through a set of Jupyter notebooks that taught basic data operations of comma-separated value (.csv) files. For example, the participants followed training on how to open a file using Python, do data wrangling, and data cleaning on different rows, finding minimum and maximum values. After the basics, the notebooks featured techniques that can be done with data from Twitter, such as how to find certain text strings, or list users. The conclusion of the tutorials from the notebooks was about basic natural language processing techniques, such as how to create bag-of-words, chart word frequencies, remove stop words, and use a word tagger.

In the second part of the day, the participants had a guided conversation with a data steward about ethical uses of data, focused on social media data. As a rule-of-thumb, the primary guideline was to use digital data as they would approach research data they collected through interviews, thus anonymized. Furthermore, approach social media data with the mindset of what kind of uses they would not like their own data to be used for.

The last part of the day was to combine the learnings and bridge it back to the frame creation. The participants learned about the *Exploratory Data Inquiry* methodology in detail, as well as about the recommended software tools for data operations (OpenRefine), text analysis (Voyant-tools), network analysis (Gephi), and at last making graphs and charts (RAWGraphs).

DAY 3-4-5 - ITERATIVE DESIGN PROCESS

The third, fourth, and fifth days were spent on iterating on the frame creation process, particularly step 5-6-7. In the following, the two groups' processes are presented in detail.

STUDY 6A

On the first day of learning about the frame creation methodology, Group A explored two related themes of *'balancing'* and *'nurturing'*. Their used metaphor was a *"marathon runner, who has to understand his/her own pace to manage their performance until the finish line"*, used as their starting point on day 3. The next section elaborates on the steps taken by the group. Figure 6.2 shows how group A described their own data exploration process.

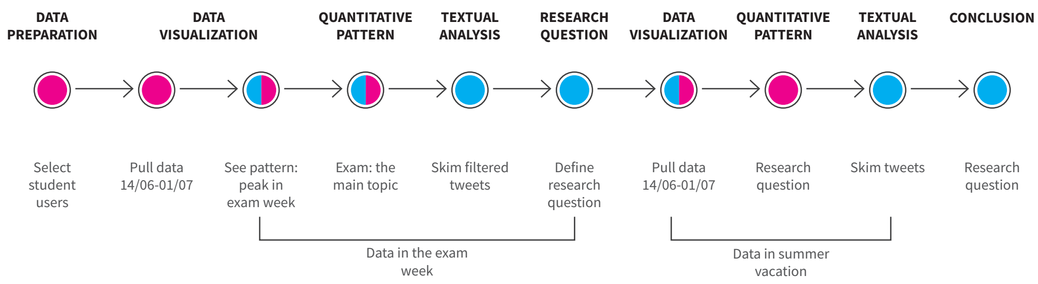


Figure 6.2. Data exploration process of Group 6A, as described by them.

STEPS

1) FILTERING THE DATASET FOR STUDENTS: The group sampled the full dataset between 14-30 June 2019, and searched for users who have the words ‘student, studying, bachelor, master, or phd’ in their user profiles. The time-frame of 14 to 30 June was chosen because it was typically considered as an exam period, and the group assumed to find related tweeting activity. From this step, the group identified 387 users as students from the main dataset.

2) LOOKING FOR PATTERNS ABOUT WHAT STUDENTS ARE TALKING ABOUT:
2A. Exporting the tweets of the students from 15 to 30 June. Since the main dataset was already related to stress and mental well-being, the result of this step was the list of stress-related tweets from the students. The group found a peak of stress-related tweets between 20 to 25 June.
2B. In parallel, the group also explored the word cloud of these tweets. Since the dataset was already about stress and depression, related terms also appeared to be the most significant.
2C. After the group saw a significant peak in the exam period, the group exported tweets between 1 to 5 August to see if there was any difference in tweeting activity between the exam period and summer vacation.

3) TRY TO UNDERSTAND WHAT THE STUDENTS ARE TALKING ABOUT: The exam period tweets were skimmed, and the group identified a core topic about exam-related stress. The group filtered the tweets to find tweets containing the word ‘exam’ or ‘deadline’ in them. The group read the filtered tweets and extracted insights into an Excel document.

4) ARTICULATING THE RESEARCH QUESTION: The group initially explored the data without having a detailed research question. At this point, the group tried to connect the insights from the previous step back to the themes identified during the first day. The overall theme was about balancing & nurturing, and the main insight was that exams are significant stressors for students.
At this point, the group noticed that they actually had an implicit

research question but were not able to articulate it. Since they started with ‘balancing’ and ‘nurturing’, the research questions were *“What to balance?”* and *“What to nurture?”*.

5) IDEATION ABOUT EXAM-RELATED STRESS: The group did a brainstorming about what might be related to exam stress, to enrich their conceptual understanding around the topic.

6) DESIGN QUESTION: By using the metaphor of a ‘marathon runner balancing her resources throughout the race’, the group then ideate on a design question. The result was: “How might we facilitate students to balance their stress points throughout their study?”

7) METAPHOR OF SUPER MARIO: Following the design question from the previous point, the group generated another metaphor based on Super Mario collecting points and going through levels throughout the game. The group used this metaphor to work out another NADI worksheet towards a concept.

8) CONCEPTUALIZATION: The group’s final concept on the fifth day was a product-service system proposition that facilitates students to create a roadmap to have an overview of their whole study while maintaining their short-term goals.

STUDY 6B

On the first day of learning about the frame creation methodology, Group B selected student peers, health associations, and student unions as stakeholders. They explored three themes of *‘bonding’, ‘encouragement’,* and ‘*fun time’*. To start-off on day 3, they chose to focus on ‘bonding’ as the next step and created a mind-map around the topic based on their own experiences what bonding means for themselves. As an output, the group formulated a list of research questions to explore further. The next section elaborates on the steps taken by the group. Figure 6.3 shows how group B described their own data exploration process.

STEPS

1) DEFINING RESEARCH QUESTIONS: By generating a mind-map together, the group concluded on one general question with two sub-questions. The general question was: *"How does friendship influence students mental health?"*, and the two sub-questions:

› In which situations does empathy relate to bonding among students?

› How does common space (such as the library or the coffee corner) influence student well being?

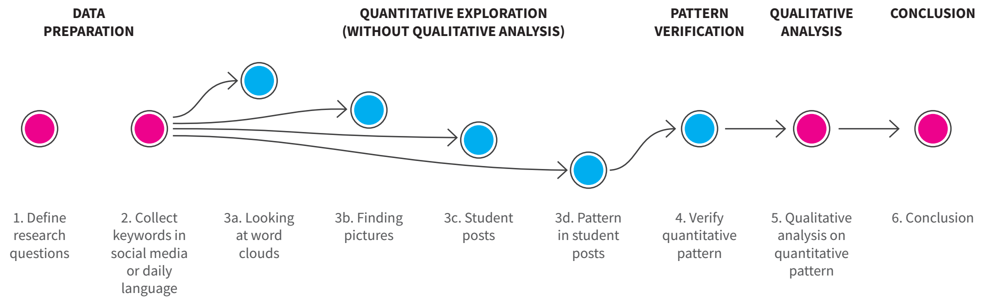


Figure 6.3. Data exploration process of Group 6B.

2) COLLECTING KEYWORDS IN SOCIAL MEDIA AND DAILY LANGUAGE: Because 'bonding' as a word was a high-level concept that might not yield relevant search results from the dataset like a colloquial word, the group began generating a list of alternative keywords by brainstorming and looking up synonyms in dictionaries. This was done in order to be able to filter the main dataset to more specific tweets.

3) EXPLORING DIFFERENT DATASETS: multiple inquiries happened in parallel.
3A. SEARCHING FOR THE KEYWORDS IN A QUERY: One of the keyword combinations was 'empathy' OR 'feel the same' OR 'pity' OR 'affinity' OR 'compassion'. The group analyzed the search results in Voyant Tools, by generating visualizations like word-cloud and word bubbles. After a while, they concluded that meaning is hard to extract from these visualizations besides an initial glimpse of what is in the corpus

of filtered tweets. By scanning through the content, the group noticed that the about 1000 tweets were re-tweets of the same content, misguiding the word cloud approach of exploring the data.

3B. LOOKING FOR INSTAGRAM PICTURES RELATED TO STUDENT'S DEPRESSION OR STRESS: The group also downloaded a dataset of 2558 recent Instagram pictures that were located in their faculty. They looked for posts that contain the keyword 'stress' and 'depression'. To their surprise, only 15 results were found from 2557 posts. Within the results, several posts contained "not depressing" or "I can't stress it more", which are opposing phrases to their inquiry. While most of the contents in the pictures in this dataset were about student design work and working people, the group reckoned that students who experience depression or stress would not tag the post with the faculty's location.

3C. FILTERING STUDENT'S POSTS DURING BEFORE AND DURING VACATION TIME: The group filtered the main dataset to posts that contain 'student OR studying' in the user's bio between 5 to 19 June 2019. This dataset was put aside as the group did not have an effective approach the analyze the data quantitatively.

3D. FILTERING STUDENT'S POSTS THAT MENTIONED INTERPERSONAL RELATIONS. The group got inspired by the other group's approach of extracting tweets by students, but including keyword 'friend OR friends OR family OR families OR team OR colleague' in the tweets' post body, for the dates between 20 June and 6 July 2019. At this time, the group noticed a peak in the number of search results rising from 4 July onwards.

4) VERIFYING THE QUANTITATIVE PATTERN: Noticing a 'quantitative pattern' from 3d, the group wanted to validate if the pattern was accidental or a remarkable change in the dataset. They conducted the same search as in 3d, but with the period following the one from the dataset in 3d. Besides, they wanted to examine whether it was by friend, family, or colleagues that contributed to the peak. They searched for the keywords separately, and it turned out that each of the keywords showed the peak. By then, the group was convinced about the fact that there was a peak of mentioning friends, family, and colleagues by students from 4 July 2019. The group interpreted 4 July 2019 as the beginning of summer vacation.

5) QUALITATIVE ANALYSIS OF THE QUANTITATIVE PATTERN: The group downloaded two datasets in the same way as 3d. As word-clouds for reading tweets were not a successful approach in 3a, they chose to read the tweets one-by-one and extract insights into an Excel document (this step was inspired by the other group). The group split up to read the two datasets. They scanned through the content and marked the insightful ones. After they finished reading, the group synthesized their findings. They summarized several themes (as in qualitative analysis methods). The central theme was *"mental issue is stopping students from socializing"*.

6) METAPHOR: With the insights from step 4-5, the group defined the ideal solution as *'a subtle socialization'*. They explored this theme through the metaphor of blowing bubbles. They analyzed the properties of the metaphor and used the analysis as the basis for ideation.

7) CONCEPTUALIZATION: Following the metaphor from 6), the group ideated a concept using the NADI worksheet. The generated concept was an app to enable students to find like-minded students after a stressful period.

The groups worked in the same room with several interim occasions to update their progress to each other, and in this way also influencing the actions of each other. Notable differences in approach were the timing of articulating more specific research questions. Group A had no implicit research question when approaching data exploration, and they made the leading research question explicit only at a later point. Meanwhile, Group B started with a research question from the beginning; however, they diverged from the research question shortly. Group A filtered down the main dataset early in the data exploration, which made their sequential steps more straightforward. Meanwhile, Group B filtered the main dataset at a later point, after struggling to extract value from a broader dataset.

In the following, we present the outcomes of the creativity support evaluation of using Exploratory Design Inquiry with Frame Innovation as a design inquiry approach.

CREATIVITY SUPPORT EVALUATION

At the end of the workshop, a creativity support evaluation survey was conducted to capture quantitative assessment about the overall approach of using exploratory data inquiry in a frame creation process. Next, we first elaborate on the Creativity Support Index assessment tool we used, the results, and then we interpret the results.

Following the example by Cherry and Latulipe (2014), we report the results with respect to average factor counts, factor score, and weighted factor score in Table 6.8. The *Average factor counts* indicate the number of times participants chose a given factor important (between 0 and 5) in the specific context of using data exploration as a design inquiry approach. The highest-ranking of *Exploration* is expected, given the context of fostering data exploration in design inquiry. The factors of *Expressiveness, Collaboration* and *Results Worth Effort* are found moderately important, and *Enjoyment* was scored considerably lower than the rest. The *Average factor score* indicates how well the participants scored the design method (between 0 and 20) for certain factors. Interestingly, the participants marked *Enjoyment* the highest. The *average weighted factor scores* are the most sensitive to factors that are marked important, and *Exploration* was scored the highest factor. Besides *Exploration, Expressiveness,* and *Results Worth Effort* are the two factors that provide an extra dimension to what has been expected and valued for the design method.

Scale	**Avg. factor counts (SD)** (between 0-5, highest 5)	**Avg. factor score (SD)** (between 0-20, highest 20)	**Avg. weighted factor score (SD)** (between 0-100, highest 100)
Results Worth Effort	2.60 (1.82)	15.80 (1.20)	**42.80** (30.52)
Exploration	**3.80** (0.84)	13.60 (1.87)	**52.40** (15.98)
Collaboration	2.60 (1.82)	14.60 (1.64)	37.20 (26.47)
Immersion	2.20 (1.10)	10.80 (2.17)	20.00 (4.00)
Expressiveness	2.80 (0.84)	15.40 (1.16)	**43.60** (15.96)
Enjoyment	0.80 (1.30)	**17.40** (0.82)	14.60 (23.51)

Table 6.8. Creativity Support Index of Study 6AB.

INTERPRETATION

ENJOYMENT: The participants marked Enjoyment low on factor count, however, it was scored the highest in assessing the workshops' performance on this factor. We triangulated these counts with the discussions and observations throughout the workshop and with the focus group reflection interview at the end of the workshop, and found that these numbers can be explained to reflect the overall good mood and learning environment of the workshop. We may conclude that a 5-days pressure cooker environment, where teamwork is essential to move the design project forward, can make an enjoyable way of conducting the Data Exploration for Design method for Frame Innovation.

EXPLORATION: The participants marked exploration the highest as factor count, and the weighted factor score also resulted in the highest value. This result is expected, as Exploratory Data Inquiry is based on supporting data exploration as a way of design inquiry.

RESULTS WORTH EFFORT AND EXPRESSIVENESS: The participants marked both Results Worth Effort and Expressiveness as important factors of the Data Exploration for Design method for Frame Innovation. These numbers can be interpreted with the participants' experience with different tools, especially DMI-TCAT and Voyant-Tools for text analysis. In general, the participants found the tools harder to use. Results Worth the Effort can be interpreted as 'even though the work is complicated to do with these tools, the results are worth it', and thus indicating the type of insights the tools enable are promising to pursue. Similarly, Expressiveness can be interpreted through the tools as well. Although the tools are harder to use, the type of visualizations that can be generated not only provide new insights in an expressive manner but might also reveal hidden connections in the data.

6.5 Discussion

The focus group interview at the end of the workshop allowed the participants to reflect on the approach of using exploratory data inquiry – more generally, data – as a mode of inquiry in frame creation, more generally in the design process. The following quote summarizes the complexity of such an approach: *"…the hard part here is being able to create the connections between the qualitative and the [computational]. […] The tools were nice, because they open your playground, but the hard thing is not to learn the tools, but to learn where to look at, is it valid enough, or if your assumption is super risky"* -PH, and then continued on iterating with tools: *"the good thing with computation is that you can iterate faster. You could extract different datasets faster, so you can try to look for different words"* -PH. As the participant reflects, the interplay of combining a qualitative and computational approach means new challenges to take into account. To be able to leverage such an approach, there is a mindset that leads the thinking process when designers embed such a computational approach into the design process. In the following, we unpack these notions in further detail, as they have particular relevance for the development of the conceptual framework.

COMBINING COMPUTATIONAL AND QUALITATIVE THINKING

First, when the participants filtered down the dataset to narrower time-frames (i.e., two weeks periods), they initially explored tweets through quantitative text analysis. Using word clouds and n-grams, they tried to develop a stronger sense of what people are talking about. In the field of digital humanities, this approach is referred to as **'distant reading'**, as opposed to **'close reading'** (Moretti, 2013). However, they found looking at tweets from such a quantitative approach only gives limited insights into the human experiences the tweets contained. A participant phrase these limits as the following: *"There are different levels of things you can get out from a text. A basic one is a word cloud, that you are just looking at the words, but then if you go deeper [in analysis] why does this word in a context was used, and you start thinking 'why are they actually saying that as a group?'"* -PH. So, first approaching the data from quantitative means (i.e., word clouds, but also n-grams) can be used as a starting point, but then for actual meaning, a closer reading of the data points is necessary. The participants were all trained in a human-centered design tradition where the design process integrates multiple ways to empathize with the people (Kouprie & Sleeswijk Visser, 2009). From that perspective, it can be argued that they lacked a way how to practice empathy when only looking at the tweets in such aggregated ways.

To resolve the shortcomings of distant reading – and the produced artifacts, such as word clouds – the participants started to read the subset of tweets and **qualitatively coded** them. This process helped in developing empathy about the users' lives and helped in the thematic construction of the human experiences in their focal points of their problem inquiry. As one participant put it: *"We all have to know what the data is, it is very important. Because, for any type of data analysis, you need to link it to the real world, so if you don't read the tweets, they are just numbers and charts. It doesn't mean anything to you"* -PQ. And then the participant added: *"If I don't read the tweets, you don't see the big picture of what is there. If I just search for 'family' and 'friends' from the dataset, I can only know how they talk about family and friends, but I don't know how they talk about other things"* -PQ. However, the participants also pointed it out that when they qualitatively coded the tweets, they could find many inspirational anecdotes, and it was easy to forget to look at these from the larger picture of the human perspective. They compared such an approach to analyzing data from qualitative inquiry: *"after we have all the interview transcripts and we analyze and over-analyze, then we go already very deep there and we forget where is the surface"* -PN. In other words, while abstracting out human experiences from qualitative research, it is a similar challenge not to loose sense of the researched human experience when approaching that from exploratory data inquiry.

The participants interchangeably used these two modes of operation; when changing the subset of data under analysis, such as different queries, or changing filtering dates, distant reading is a 'quick' approach. However, they required qualitative reading to understand the data. Furthermore the qualitative understanding could provide the necessary changes in the inquiry that moved the process forward, in a way creating an iterative structure of moving from distant to qualitative and repeat. As summarized by one of the participants: *"I see as we can start with data, and then validate our assumptions. So, learnings from the data mining, we can make some concepts, and inside* *the concepts there are some assumptions. Maybe inside the concepts, the next step can be testing those assumptions, so qualitative user testing, interviews, etc. That will be data-first and then doing qualitative research. Or, we can start from qualitative research, and then we look at the data to see the scale of the bigger situation. Data can be in the first, or data can be in the last part."* -PN. Following the logic of this participant, inferring from the data is combined with qualitative research, but data can come either as first or last; as such, either as a source of inspiration or as means to validate. Such an intertwined approach of using exploratory data inquiry at different phases of the research was put in different words by another participant: *"[Exploratory Data Inquiry is] really useful to use data to create an overview of the user group I'm doing research with. So it can form a first impression of them, which leads to a lot of different directions and I can choose which ones I can go deeper in."* -PF. Thus, exploratory data inquiry may require complementary design techniques for empathizing, but the nature of design practice enables that.

The study participants emphasized that by combining data thinking (i.e., computational thinking) and design thinking (i.e., qualitative thinking), a *'hybrid mindset'* emerged. Such a mindset builds on designers looking at data either quantitatively or qualitatively, and choose appropriate methods of exploration and analysis according to the needs of the design process.

TOWARDS DESIGNER DATA LITERACY

The participants gained hands-on skills working with data throughout the workshop. This learning effect, however, not only took place as a way of skill acquisition, but also as a shift in mindset, indicating the development of more holistic data literacy. Data literacy is a type of literacy that attempts to center data in a set of skills, competences, and practices. As data is framed in multiple ways as an engineering problem or as a more abstract socio-technical phenomenon (see more in Section 2.3), different concurrent data literacy perspectives exist, and no previous work has framed it from the perspective of designers. Wolff and colleagues (2016) discuss data literacy for citizens and learners as a way of teaching and empowering non-experts of data. Data infrastructure literacy focuses on competences to account for, intervene around, and participate in wider socio-technical infrastructures around data (Gray et al., 2018). Data information literacy is framed by librarians, a profession that potentially has the longest traditions of data literacy, referring to twelve competencies and skills related to the practices around data management, curation, and reuse (Carlson & Johnston, 2015, pp. 44–45). Shared in these different data literacy interpretations that they do not only codify skills and competencies, but also represent a mindset or thinking patterns, similarly to design methods (see Section 5.2). Knowing the ways how to operate data, what tools exist, and what are their purposes are fundamental elements of data literacy, but nuances how to think with data to make it relevant for design is what makes it specific for designers. While the focus of the current study was not to describe data literacy of designers, investigating the mindset of how designers use exploratory data inquiry in the design process can contribute towards a designer data literacy.

The current study highlights that when data is approached for *inquiry* in the design process, there is a mindset shift that can be described as 'developing the data lens', or in other words, to develop an eye on how to look at the world to see potential data sources. However, such a lens provides value only if the designer can make rational and strategic choices based on what kind of inferences can be achieved from certain types of data and methods of analysis, furthermore to have a sense of problematic assumptions, ethical usage, and at last creating value for the design process.

BIAS AND LIMITS: The participants compared exploratory data inquiry to other types of inquiries, such as interviews or field studies, as alternative ways of qualitative data collection. Although bias and limits exist with all types of data sources and modes of data collection, with interviews or other qualitative techniques, the biases are less prominent or problematic: *"...the data itself limits [our inquiry] to internet users. The context, the users are already narrowed down, so it can only be a part of research and design. I mean, the scope is really small"* -PF. Interestingly, the participants were less concerned about the limitations of qualitative studies and the biases and ethical problems related to them. Limited concern can be accounted partly because working with publicly accessible online data, problems surface by large-scale data aggregation, and the high-level patterns one can find from big volumes of data. As a solution, the participants emphasized to complement the computational approach with keeping a human-level lens, which is practically by using qualitative inquiry as a way of combining approaches: *"I can find interesting data from Twitter, but I still need other efforts to define my user group, to define others who don't use Twitter. What will be their behavior? What will be their patterns about the problem?"* -PF**.** Later, the participant added, *"for a next project, I should start thinking about what kind of people I will focus on and then choose what kind of data I can use"* -PF. Focusing on the target group also flags the importance of keeping datasets relevant, which we further discuss next.

KEEPING DATASETS RELEVANT: The participants highlighted that just using any available dataset is problematic, and urged the use of datasets related to the problem: *"you need to find datasets that are related to the problem that you are trying to solve"* -PH. Later, the participant added: *"dataset can be comments on forums, for example about games, or can be reviews from Airbnb… it can be a lot of different stuff."* -PH**.** Following the logic in this quote highlights the development of the *'data lens'*, but also shows the problem, that with a co-evolving problem and design space, sometimes new types of datasets need to be used as potential data sources. Such flexibility can only be achieved in practice if the designers themselves are able to identify valuable data sources, and it also helps if they can acquire, clean, and process potential data sources to some extent.

To summarize, using exploratory data inquiry in the design process imposes new considerations on how to approach and think about inquiry, which is related to the nature of using data. By *'developing a data eye'*, designers can look at the world as a potential data source, given they have a holistic understanding also about how to keep data exploration relevant to the design process, and that they have an awareness of the potential biases, limits and ethical considerations about data.

LIMITATIONS

The current study provides an in-depth view of five participants' design process during a one-week long workshop, which imposes several limitations on the study, both in the sample size of participants as well as the observed period. All participants were fifth-year master-level design students, with limited experience in professional design practice. The one-week long design workshop resembled a pressure cooker design environment, however, this is not representative to design projects in practice that cover more extended periods. Furthermore, the participants used Twitter data curated by us, which approach would unlikely to take place in real life.

The limitations above all indicate future opportunities to further expand on the work. First, further investigating what kind of mindset expert designers assume would be beneficial to validate the findings. Second, analyzing cases over longer time-frames would reveal further how exploratory data inquiry influences strategic design choices, as the current study only revealed a rather 'tactical design' level.

6.6 Conclusions

The research questions of the current study were concerned about to what extent exploratory data inquiry can be applied in the design process, and to identify the key components of the designers' mindset when using data exploration as a primary mode of design inquiry. From the current study, it can be concluded that *Exploratory Data Inquiry* can intertwine into the design process similar to other design techniques. As designers pick tools, techniques, methods as the unfolding co-evolution of problem and design space leads the design process, *Exploratory Data Inquiry* can be used by designers to inquire the digital world by formulating what data to acquire, acquire the data, and explore and analyze the data as an intertwined part of the design process. The second research question of the study was concerned about the characteristics of the mindset designers assume, for which we found two main characteristics. One characteristic is the thinking process that combines data thinking with design thinking, often based on the same dataset. The second characteristic is developing an eye for looking at the world as a potential data source, but with the sensibilities of a holistic literacy around data to know what can be inferred from data and how. In practice, this means knowledge of limitations, biases, and ethical considerations such an approach would carry.

LESSONS LEARNED FOR DEVELOPING METHODOLOGICAL CONTRIBUTIONS

This study shows how *Exploratory Data Inquiry* is adapted in design practice. As one of the goals of the current study was to 'stress test' the design approach, the contributions to the conceptual framework reveal valuable input what to consider in real-life usage of design inquiry through data.

TOOLS: Compared to previous studies, in the current study, the dataset was a magnitude larger, which imposed new challenges. We learned from this study that working with 11 million tweets requires computational power, and increased awareness of data practices to be able to extract meaningful subsets of data not by chance but by strategical choices.

MINDSET: During this study, we also learned about the alternation between computational/data thinking and design/qualitative thinking (see Figure 6.4). This took place combining quantitative text analysis and qualitative coding of datapoints as an intertwined way of looking at the same data.

PROCESS: The study showed that the general mindset and approach of Exploratory Data Inquiry could be well-integrated into a larger design framework, such as Frame Innovation. Related to the alternation of computational/data thinking and design/qualitative thinking, from a process perspective it can be seen that the design/qualitative thinking might take place 'outside' of the Exploratory Data Inquiry process, where the design team is aligning on the goals of inquiry, trying to relate the inferences with other inquiries.

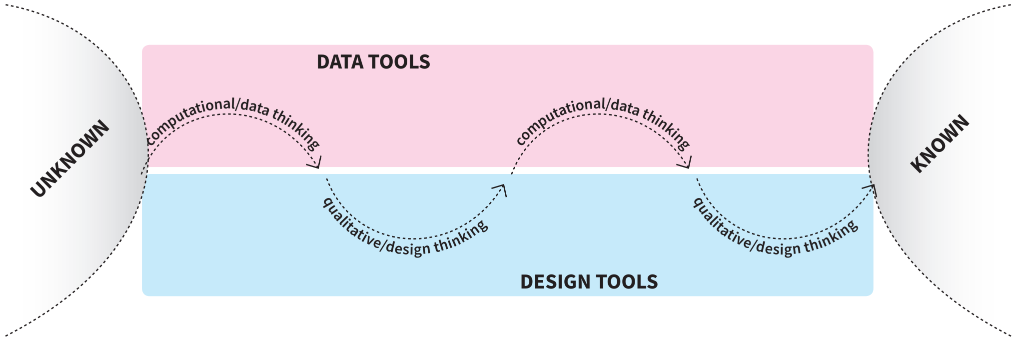


Figure 6.4. Computational/data thinking and the design/qualitative thinking intertwines through data and design tools, respectively.

The previous chapters started from a 'naive' investigation of how data science practices could be appropriated in the design practice (Chapter 3), and then continued with analyzing data science practices through a creative process lens (Chapter 4). Analyzing data science practices through a creative lens resulted in a methodology for Exploratory Data Inquiry, which we used for developing the Data Exploration for Design method in Chapter 5. The current chapter contextualized the use of Data Exploration for Design method in a more realistic design case, during which we studied the mindset designers assume when approaching design inquiry through data. In the next chapter, we will first summarize the different studies and reflect on them through a set of observations as a way of revisiting and answering the research questions of the current dissertation.

Chapter 7

Synthesis and Discussion

In this final chapter, we reflect upon the Design Inquiry Through Data framework by synthesizing the empirical investigations of Chapter 3, 4, 5, and 6, based on their process, mindset, and tools. We conclude the Design Inquiry Through Data framework as a composition of opportunistic process of data exploration, a hybrid mindset of combining computational/data thinking and qualitative/design thinking, and the use of visualizations as prototypes and boundary objects in design inquiry. In the second half of the chapter, we discuss the implications of our research findings, ethical considerations, and recommendations for future research.

7.1 Introduction

As stated in Section 1.2 earlier, the overall aim of this dissertation has been to **develop theoretical and practical knowledge on the intersection of design and data science to enable designers to use data-rich practices for design inquiry**. In the previous sections, we formulated a conceptual framework and conducted a series of studies as part of a RTD program. In Section 2.4, the three elements of process, mindset, and tools were initially selected to decompose the transitioning from the unknown situation to the known situation, which refers to a typical design scenario. In such a scenario, the designer follows a process that sets the transitioning. The process is rationalized by a mindset that the designer assumes as she makes sense of the design situation momentarily. The designer's thinking process is leveraged by the tools involved in order to interact with the data. Next, we summarize the lessons learned from the studies with respect to process, mindset, and tools, which we synthesize in Section 7.2, 7.3, and 7.4, respectively.

Throughout the RTD program overarching the different studies, our understanding of the three elements of the data-rich design practice of *Design Inquiry Through Data* evolved with every step. The learnings of each study have continuously been integrated into the following one. Table 7.1 shows that with respect to *process*, our understanding started from the emphasis of designers being involved in the overall data spectrum, from data collection to inferring insights. In the follow-up, we further clarified the holistic involvement of designers by introducing the three stages of 'problem framing', 'exploring' and 'inferring'. These three stages are iterative and consist of multiple data and design activities, not necessarily conducted in a linear manner. With respect to *mindset*, Table 7.1 highlights how designers use their familiar sensemaking patterns, such as abduction to use data for design inquiry. The studies showed that designers' creativity and mindset manifest in empathetic and creative definition of what data to collect, leading towards a thinking process that combines computational/data thinking with qualitative/design thinking. Table 7.1 illustrates that with respect to *tools*, our first study confirmed that non-expert data tools are suitable for design inquiry. In further studies, we found that using non-expert data tools can become computationally more challenging with larger datasets, by the sheer amount of data. The role of design and data tools has also been found different, with the former being more valuable to support the learning curve of the data-rich design practices.

Table 7.1. Overview of empirical studies and the lessons learned from them.

Study	RTD output	Process	Mindset	Tools
3A (Master thesis records 1) 3B (Service Design Tourism)	Initial understanding of the scope and characteristics of design inquiry through data	Navigating the whole process from data collection to inferring insights.	Abductive sense-making for data exploration, creative appropriation of tools. Qualitative methods used in concert with data practices.	Non-expert data tools are suitable for design inquiry.
4ABC (Service Design Mobility)	*Exploratory Data Inquiry* methodology (process)	Data and design process can be combined along 'problem framing', 'exploring', 'inferring'.	Defining what data to collect is non-trivial and requires designers empathy. Framing and reframing with data exploration.	Caveats of using social media data. Technological limitations and timeframes to take into account.
5 (Master thesis records 2)	*Data Exploration for Design* method and design tools based on the *Exploratory Data Inquiry* framework	The three stages of 'problem framing', 'exploring', 'inferring' are iterative, and design / data tools introduce shifts.	Holistic involvement and spending time with data is very valuable. Being part in data collection reduces bias.	Distinguishing between design tools and non-expert data tools. Design tools have a role in easing the learning curve.
6AB (Frame Innovation + data exploration)	Evaluating *Exploratory Data Inquiry* in a more realistic scenario and extracting the key characteristics of mindset.	Exploratory Data Inquiry integrates well into Frame Innovation.	Computational/data thinking and qualitative/design thinking combined around the whole process.	Dataset with actual big data introduces computational challenges for end-user tools.

From the five different studies, we gained additional insights into the three elements, which we will synthesize and reflect upon in the three next sections. First, in Section 7.2 we will reflect on the *process* and how it is governed by opportunistic logic for data exploration. Second, in Section 7.3 we will reflect on *mindset* and how the intertwining of computational/data thinking and design/qualitative thinking leads to a new hybrid mindset. Third, in Section 7.4, we will reflect on *tools* and how they are equipped with designerly characteristics of prototypes and boundary objects when used in data-rich design practices. Afterwards, we will summarize our contributions through answering the research questions and the *Design Inquiry Through Data* framework and then present the implications of the research, the ethical considerations ,and potential directions for future work.

7.2 Process – Opportunistic data exploration

In Section 2.1, we characterized designing as a process that is exploratory and opportunistic to describe the underlying logic that designers follow for inquiry. Throughout the studies, we observed the same pattern that designers try to incorporate exploratory data inquiry in an opportunistic fashion. These exploratory and opportunistic characteristics of designing have been identified as abductive sensemaking (Dorst, 2011), enabling the analysis of the studies from this perspective. In the following, we show how Dorst's (2011) framework of deduction-induction-abduction – the typical reasoning patterns designers follow – could be used to interpret the opportunistic process of data exploration. Dorst provides a framework that can highlight how designers use data, especially in comparison with more 'traditional' uses of data for analysis or prediction. Figure 7.1 shows the three-element formula Dorst (2011) introduced to describe reasoning patterns: "What? – thing", "How? – working principle", and "Aspired value".

WHAT? *(thing)*	+	**HOW?** *(working principle)*	*leads to*	**VALUE** *(aspired)*

Figure 7.1. Design reasoning framework by Dorst (2011).

In Figure 7.2 we match these elements around for opportunistic exploration of data, showing the components of a data inquiry that are available to the designers: data (a thing); a data technique (a working principle); or a question (an aspired value). In other words, what are opportunistic scenarios that can be a starting point for data exploration in a design inquiry.

DATASET *(thing)*	+	**DATA TECHNIQUE** *(working principle)*	*leads to*	**VALUE** *(aspired)*

Figure 7.2. Dorst's (2011) design reasoning framework annotated for data exploration.

1) THERE IS AN AVAILABLE DATASET

In this scenario, the designer has access to a dataset related to the problem and could use it to better understand an inquiry, as shown in Figure 7.3. For example, the product designer of a software product can have access to usage log-files of a product, and choose different working principles, such as text mining, or visualizing the product usage for analytics.

DATASET *(thing)*	+	**???** *(working principle)*	*leads to*	**???** *(aspired)*

Figure 7.3. Dataset is available for opportunistic data exploration.

For example, the dataset of master thesis records (e.g., Study 3A and Study 5) was explored in multiple ways by the participants. One specific exploration of data was as a network graph (see Figure 7.4).

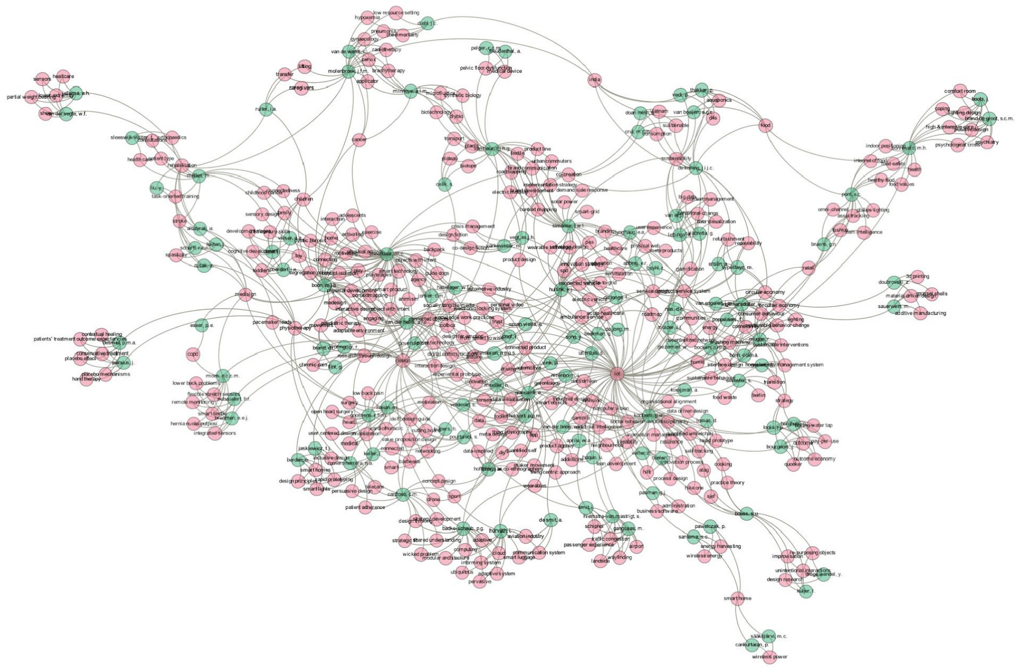


Figure 7.4. An example network graph, showing the map of keywords (pink) and faculty mentors related (green) to 'iot'. (Source: participants' slide)

Some of the design workshops had the explicit learning goal to try out new data techniques (i.e. working principles), and the participants could openly choose any data technique to experiment with, as they saw fitting their inquiry. With the same dataset, therefore, parallel approaches took place, for example by charting hierarchical connections differently (see Figure 7.5).

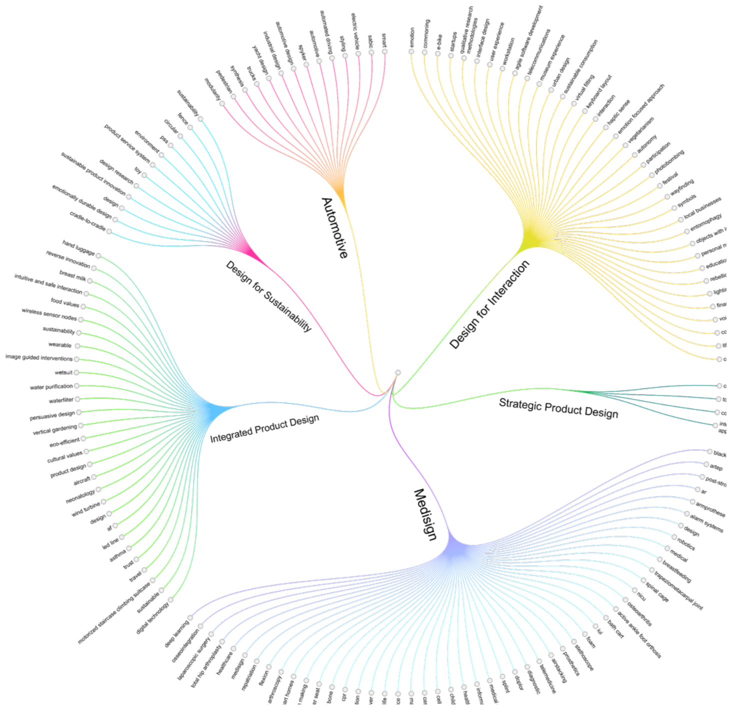


Figure 7.5. An example of a so-called 'circular dendrogram' visualization of the master thesis records dataset. Shows hierarchical connections between master programs (e.g., Strategic Product Design), and keywords of graduation thesis on the outer labels. (Source: participants' slide)

2) THERE IS A SPECIFIC DATA EXPLORATION TECHNIQUE TO BE USED

Figure 7.6 shows the scenario when a tool is already specified to explore or analyze a dataset. For example, when a designer is learning about network visualization and wants to create a network graph to visualize connections between entities, the network graph can be drawn from connections of graduation project mentors and keywords (e.g., Study 3A and 5).

??? *(thing)*	**+**	**DATA TECHNIQUE** *(working principle)*	*leads to*	**???** *(aspired)*

Figure 7.6. Data technique is available for opportunistic data exploration.

The extracted value from generating such a visualization is upon the designers' sensemaking of it; the value can be the graph itself as a representation of the data, but also the extracted insights for increasing understanding of the problem space. To illustrate, during Study 4ABC, the groups learned about the potentials of using quantitative text analysis on social media datasets. One group opportunistically went further to try quantitative text analysis also on their raw interview transcripts corpus, with the hopes that they will gain a faster overview of the interview study prior to thoroughly analyzing it.

3) THERE IS A SPECIFIC ASPIRED VALUE SOUGHT AFTER

In the scenario shown in Figure 7.7, there is a specific value sought after. Such a specific value can be an answer to a research question, a specific data point, or an insight that builds the understanding of the problem space. In this case, towards aspired value, neither the used data nor the working principle is defined.

???	+	**???**	*leads to*	**VALUE**
(thing)		*(working principle)*		*(aspired)*

Figure 7.7. Aspired value is fixed for opportunistic data exploration.

In this scenario, designers' abductive thinking can thrive on generating ideas on how to collect data for the specific value, such as finding a dataset or collecting data with sensors. Dorst (2011) underlines that ideating and testing in parallel of a "what" and a "how" is a complex creative problem. In our data-rich discussion, it means ideating on the data collection as well as the data technique to use for exploring the data. Consequently, knowing a variety of working principles (i.e., a variety of data techniques to explore and analyze data), and knowing what data makes sense to collect to be able to process it with the different data techniques become increasingly important, highlighting a need for the holistic development of data competences and a path for mastering a data-rich design practice.

The open form of reasoning process explained above indicates how in practice design inquiry and data exploration can happen intertwined and in a methodical way. Designers can opportunistically take advantage of exploring a dataset or use a data technique to extract otherwise 'hidden' relations from the data. Next, we examine the intertwined mindset of using computational/data thinking and qualitative/design thinking in the same design inquiry through the data process.

7.3 Mindset - Hybrid mindset

With the evolution of the studies, we developed a more precise perspective that computational/data thinking and qualitative/design thinking intertwine (see Section 6.6). For example, during the studies using social media data (Study 4ABC, Study 6AB), quantitative analysis of the datasets was combined with qualitative coding or simply 'reading' the tweets in the datasets. Such an approach challenges the 'traditional' understanding of dividing data and data practices between qualitative and quantitative data. The studies showed that when designers work with complex and heterogeneous datasets, they combine qualitative and quantitative understandings of data. Designers combining qualitative and quantitative approaches of data leads to ambiguity in how data to be used for design inquiry. However, such ambiguity around data can be advantageous instead of limiting. First, we will examine the studies by the lenses of mixed methods research that builds on combining quantitative and qualitative data, and then by a more emergent framing of big data and thick data. Afterwards, we will synthesize these and relate to our observations.

RELATIONSHIP TO MIXED METHODS

Mixed methods research designs are based on the combination of qualitative and quantitative methods in different configurations. Therefore, mixed methods could potentially describe the observed combination of computational/data thinking and qualitative/design thinking. Creswell and Clark (2017) identify three main mixed methods research designs where the qualitative and quantitative methods follow each other in an 1) *explanatory* capacity; in an 2) *exploratory* capacity; or in a 3) *converging* capacity. In an explanatory capacity, quantitative and qualitative methods are used in a sequence, when the latter provides explanations on the findings of the former. In an exploratory capacity, quantitative and qualitative methods are used in a sequence, when the former is used to inform the latter method, such as an initial interview study informs the development of a larger-scale survey based on the identified patterns from the former. In a *converging* capacity, the goal of the mixed methods research is to merge and compare the results of qualitative and quantitative methods towards a shared interpretation. Such convergent design might benefit from the triangulation of different qualitative and quantitative data sources, but also enables inquiries that are more applied and pragmatic for real-world cases. In principle, mixed methods research is always based on separate quantitative and qualitative data collection and analysis, and that was not the case during the studies. The studies that used social media datasets (Study 4ABC, Study 6AB) used the same dataset for both quantitative and qualitative data practices, and beyond the dataset, the practices also seamlessly moved the inquiry forward in an integrated manner, not by distinguishable sequence.

BIG DATA AND THICK DATA

Mixed methods are research strategies of scientific research, but they have limited applicability to the applied research conducted within design practice. To analyze the studies through a more applied lens, we turn to Wang's (2013) notion of "Big Data and Thick Data", an integrative approach of combining big data analytics with qualitative data collection to inform inquiries or interpret results better. Wang's perspective as an ethnographer is to show the value to the rich data collection and insights qualitative methods can bring to interpreting the otherwise abstract analysis based on numbers. According to an interview with Wang (Neill, 2020), 'thick data' is essentially a rebranding of qualitative data. As we discussed earlier in Section 2.2, the definitions of data are challenged, particularly in the big data era. Earlier, we referred to 'data' as complex and heterogeneous datasets (Mayer-Schönberger & Cukier, 2013; Kitchin, 2014a); we find Wang's interpretation suiting this emerging understanding of data over the more traditional quantitative and qualitative divide of research inquiry. During the studies of Study 6AB, when the design teams were limited by using one approach, such as quantitative text analysis, they could move forward choosing a qualitative approach, such as reading the tweets in their dataset. The simple close reading and qualitative coding of the tweets resulted in insights about what was happening in time, and that understanding made the teams return to calibrating what dataset they are filtering on.

This seamless integration of computational/data thinking and qualitative/design thinking points beyond alternating between two modes of mind. More precisely, the integration shows the emergence of a hybrid approach that we will call the 'hybrid mindset'. We credit the seamless integration of qualitative and quantitative approaches to the dual definitions of data existing in design. On the one hand, design traditions come from science and engineering, in which fields 'data' is traditionally defined as measurements from an instrument. This logic is dominant in using sensor data to quantify a certain aspect of the physical world. On the other hand, designers also build on traditions coming from social sciences like anthropology, in which fields 'data' is traditionally defined as facts about a phenomenon. For example, an ethnographer will call field notes, pictures, or interview snippets as 'data'. This logic is dominant in techniques, such as contextmapping (Sleeswijk Visser et al., 2005) to capture deep insights about users.

Both mixed methods and big data and thick data indicate a convergence in the big data era (as has been detailed in Section 2.2). The convergence of two data paradigms can benefit designers to use qualitative and quantitative data in for design inquiry flexibly. For example, taking the case of social media data, these two different understanding of data has resulted in two different approaches. Computational social scientists (Lazer et al., 2009) use massive amounts of data to model and visualize massive networks of human interactions via social media. Meanwhile, digital ethnographers (Pink, 2016) mine social media, such as photos, tweets, or blog posts to describe communities that flourish in online spaces (Tiidenberg & Gómez Cruz, 2015). During the studies, both types of understanding of data have appeared and were used as part of design inquiry. In other words, the two definitions of data were used interchangeably, using qualitative tools on quantitative data, or quantitative tools on qualitative data.

In the next section, we interpret *tools,* and more specifically, how visualizations take the role of prototypes and boundary objects in design inquiry through data.

7.4 Tools – visualizations as prototypes and boundary objects

With the evolution of the studies, we gained an increased perspective that the appropriation of computational tools for designerly working is a key characteristic of design inquiry through data. Computational tools for the non-designerly aspects of the work, such as cleaning the dataset, filtering the data, or transforming the dataset, have remained at the level of helping designers to operate and leverage data. However, we observed the appropriation of computational tools to visualize the data as designerly tools. During the studies, we intentionally refrained from emphasizing visualizations, not wanting to steer the studies towards information visualization and infographics problems. Nevertheless, visualizations turned out to be impossible to be ignored in the context of working with data. Visualizations are essential to turn raw data into 'consumable' representations for human cognition (Card et al., 1999) by forming abstractions over raw data, highlight and contextualizing specific aspects of the raw data. In this representation role, visualizations are often associated with designers and the substantial expectations of their designed properties to be aesthetic, effective, and functional. However, throughout the studies, such aesthetic properties of visualizations were overlooked, as visualizations were approached as 'means-to-an-end' of an inquiry workflow, not as a designed output *per se*. For example, throughout the studies, design teams generated word-clouds by common online word-cloud generators to gain a distant reading perspective on their datasets. These word-clouds were never used as part of a report or to communicate outside the design team. Nevertheless, the word-clouds captured and stored inquiry directions, as well as 'greasing' the designers' sensemaking to provide a glimpse into the dataset. The use of visualizations as interim tools for supporting design cognition equips the visualizations with properties of prototypes and 'boundary objects' in the design process. The prototypes and boundary objects lens enables us to analyze and evaluate visualizations through existing design theory frameworks. Furthermore, the 'interim visualizations' like the exemplary word-clouds are inherently representations of reality and can be ambiguous, but such property can be beneficiary and productive, as we will analyze later.

Lim, Stolterman, and Tenenberg (2008) theorized on the role of prototypes as 'filters' and as 'manifestations' in interaction design. If we position visualizations into design inquiry's transitioning from unknown towards the known, visualizations can be interpreted as a way of highlighting one aspect of the problem space, with the dimensions, properties shown or hidden. For example, the dataset from Study 3A and Study 5 enables to create a network graph that shows the connections between keywords (such as 'sustainability') and master thesis supervisors (such as 'Dr. Smith') hides properties such as who was the author of a master thesis that established the connection between keywords and supervisor, see Figure 7.9. If we want to know how popular 'sustainability' is as a thesis keyword, a network graph can show it in an ambiguous way, as shown in the next section, however the chart in Figure 7.8 can show this more effectively and with less ambiguity.

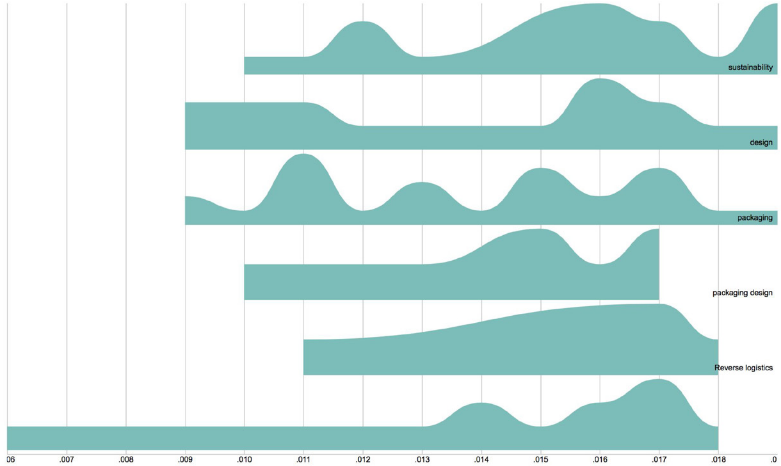


Figure 7.8. Number of keywords per year visualized, indicating the popularity of a certain topic. (Source: participants' slide).

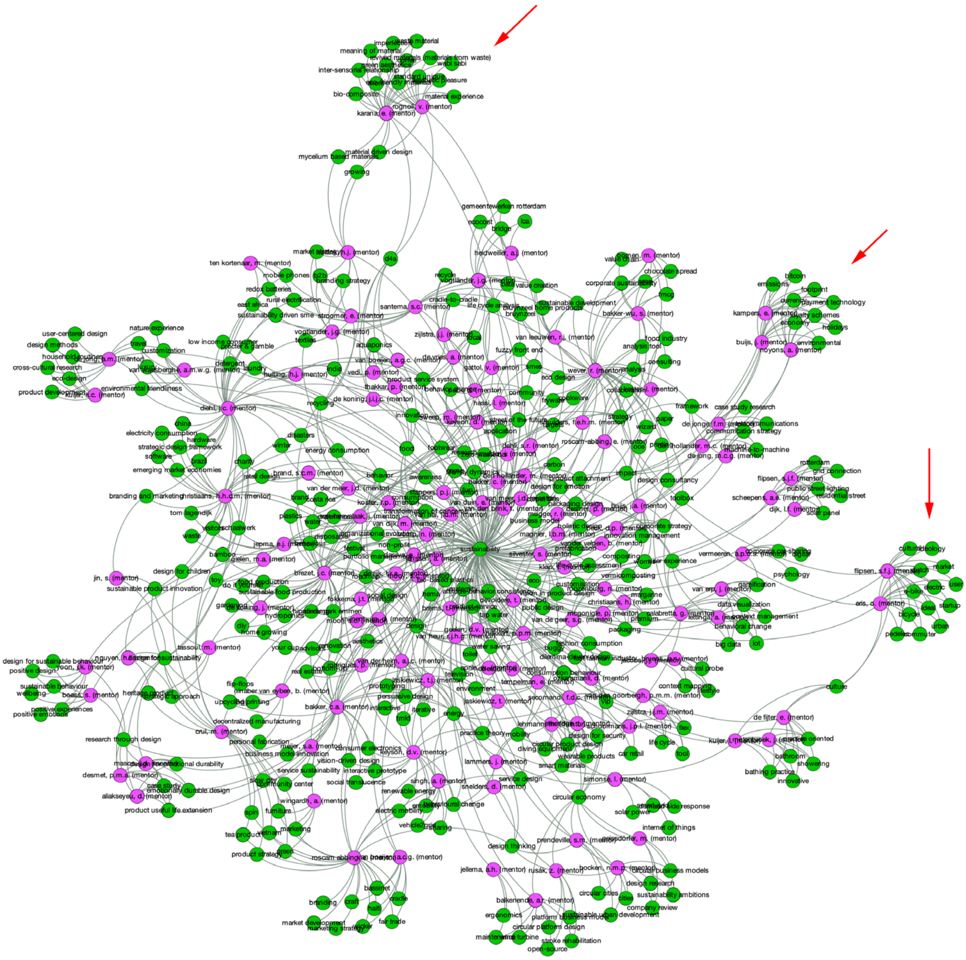


Figure 7.9. Example of a network visualization from master thesis dataset on ‘sustainability’. The green dots represent keywords, and the purple dots represent faculty mentors. The clusters on the peripheral parts of the visualization, marked with arrows, indicate noticeable clusters focused on certain sub-topics within ‘sustainability’. (Source: Peter Kun)

They define **filtering** as: *“[...] by selecting aspects of a design idea, the designer focuses on particular regions within an imagined or possible design space. The designer screens out unnecessary aspects of the design that a particular prototype does not need to explore. Designers may purposefully do this so that they can extract knowledge about specific aspects of the design more precisely and effectively. The decision of what to filter out is* *always based on the purpose of prototyping"* (Lim et al., 2008, p. 7:3). If we apply the prototype-as-a-filter lens on the examples from above, this lens highlights how the visualizations are made for inquiring towards specific queries, for narrowing down the 'unknown' aspect from the conceptual framework, as showing every property of a dataset hardly leads to scoped answers. However, it is still an exploratory search process to identify the interesting aspects of data or to understand the properties of the data and their relations and to make decisions about what to visualize and what to leave out.

They define **manifestations** as: *"A designer can determine the manifestation dimensions of a prototype by considering the economic principle of prototyping, which we define as follows: the best prototype is one that, in the simplest and most efficient way, makes the possibilities and limitations of a design idea visible and measurable. If we keep the economic principle of prototyping in mind, determining the values of the manifestation dimensions—that is, the materials, resolution, and scope of the prototype—can be approached in a rational and systematic way"* (Lim et al., 2008, p. 7:3). If we apply the prototyping-as-a-manifestation lens on the examples from above, it is crucial to know how much effort a certain visualization takes to be created. Network graphs are non-trivial to be drawn using spreadsheet software (i.e., Excel). However, if a designer learns the workflow of how to generate a network graph and use a specific non-expert tool (i.e., Gephi), they can be made quickly. Another example is the word-clouds from above; by using random word-cloud generators found online, the economics of prototyping was low. The quickness and ease of effort to make word-clouds enabled quick iterations on many different slices of the dataset, to investigate whether something remarkable pops up.

There are two take-aways from these two aspects. Visualizations as filters necessitate the learning of what kind of visualization affords certain types of inferences or communicates certain types of insights. Filtering out aspects of a dataset leads to biases, and these need to be taken into account. Visualizations as manifestations necessitate learning the 'tools of the trade', to become efficient in using a wide-palette of tools for different sorts of data, so they are economically viable manifestations throughout the design process. In other words, if the designer is able to create a visualization without needing to wait for an expert, that will enable the designer to manifest thoughts and aspects of the problem space to move fast.

VISUALIZATIONS AND AMBIGUITY

When visualizations are considered from a prototype-as-a-filter perspective, a crucial aspect is what kind of properties are left out from the visualization, or what properties are ambiguous. Although ambiguity can be a resource in the design process (Gaver et al., 2003), it remains unclear how it relates to visualizations. Venturini, Jacomy, and Jensen (2019) discuss network graph visualizations and their reading as an ambiguous way of inferring, yet still preferred over the more precise, calculated approaches that exist, such as network centrality measures. They highlight a scalability problem of network visualizations, as a problem of 'exactitude', referring to exact or ambiguous inferences that can be read from the visualization. They see ambiguity as a strength, as calculated approaches discard complexity a network graph immediately shows, while also reducing the network only to quantifiable properties. Visualizations then have usefulness to explore uncertain phenomena, which relates to our investigation of using data in design exploration. A small- or medium-sized network is relatively easy to visualize, but there is a point when it becomes too complex to represent all datapoints productively. As a solution, they propose two approaches to reading networks. Their first approach is a diagrammatic reading of small networks, which effectively means networks to be considered as diagrams, preventing visual clutter. In these cases, there is an increased reading comfort in showing the data as a network. Their second approach is a topological reading when the network's topology shows patterns that can be detected by spatial arrangements. This latter is similar to the visual clustering of affinity diagrams in design. For example, during the studies, the more successful network visualizations were those that attempted to visualize a small- or medium-sized network, see Figure 7.9. Visualizing massive networks was an interesting exercise based on the library dataset (Study 3A, Study 5), but the participants had trouble to infer meaningful insights. Meaningful was hard to interpret not only because of the sheer size of the network but also by the inexact clusters shown on the network graph.

The perspective of seeing the value in the ambiguity of visualizations is unusual from an engineering and scientific tradition of visual analytics, where data – and therefore visualization – is expected to show objective truth. During the studies, visualizations were used primarily as a means-to-an-end to 'grease the cognition' of the designers during data exploration. To conclude, visualizations are used as prototypes for design inquiry, and in that case, the ambiguity is a resource for inspiration and for feeding the abductive sensemaking process (Dorst, 2011; Kolko, 2009). Equipping visualizations with characteristics of prototypes and boundary objects is a more nuanced understanding of how the appropriation of data tools happen in practice.

In the next section, we return to answer the research questions of the dissertation and conclude on the final *Design Inquiry Through Data* framework by characterizing a design approach cumulating the insights gained in the previous sections.

7.5 Contributions

In Section 1.2, we defined the main research question to guide the dissertation towards its aims:

> **Main RQ:** *"How can designers integrate data practices into design inquiry?"*

Below we answer the five specific research questions that were addressed throughout the studies. As the questions were connected to specific chapters, we refer to the connected chapters as well to clarify how the answers were reached. Afterwards, we conclude the final *Design Inquiry Through Data* framework.

> **RQ1:** *How can design and data science be aligned as mode of inquiry?*

Towards addressing the dissertation's overarching aim and main research question, Chapter 2 summarized the literature on data and design processes, framing a future data-rich design practice – design inquiry through data – where data science practices are used for design inquiry. Combining design and data science at the fundamental level of inquiry shows shared characteristics that enable alignment. The nature of design inquiry can be characterized as open-ended exploratory and opportunistic. These characteristics share nature with data exploration being opportunistic and reactionary to the continuous learning throughout the data exploration process. These aligned characteristics informed the setup of a conceptual framework for design inquiry through data.

> **RQ2:** *How do designers appropriate non-expert data science practices for design inquiry?*

In Chapter 3, a 'naive' exploratory study was presented, in which we observed the appropriation of data practices by designers, investigating how data practices can intertwine with designing. The findings show that design inquiry through data can be conducted through the appropriation of non-expert data tools. More specifically, designers creatively appropriate data practices using abductive sensemaking, and this is enabled by intertwining the design and data steps by *navigating the whole process*. In other words, at the beginning of the process, designers partake in defining the data collection by formulating the 'unknown' as a question to lead the inquiry. Furthermore, designers explore the data and infer insights from the data to fuel other design activities, such as qualitative inquiries used in concert.

> **RQ3:** *How can data science practices be characterized through a creative process lens?*

In Chapter 4, we deepened our conceptual framework to consider data science practices as a *creative process*, and in that way, observe how data practices are used creatively for designing. The study indicates that framing data practices as a creative process can explain an opportunistic practice, that still follows traditional stages of 'problem framing', 'exploring', 'inferring', but engrain data practices all over.

In these stages the creativity and empathy of designers manifest through *query design* to ensure that the data collection is representing and helping to investigate human experiences. We concluded the *Exploratory Data Inquiry* methodology that combines data practices and design practices.

> **RQ4:** *How can a design method support design inquiry through data?*

In Chapter 5, based on the *Exploratory Data Inquiry* methodology, we developed the *Data Exploration for Design* method, which we evaluated in an empirical study. The results of the study show that distinguishing between data tools and design tools enable focusing on the 'cognitive' aspects of data-rich design practice by design tools. Design tools also have the second function to support the learning curve of the approach, and to scaffold how to integrate the method into the mindset of designers. The outcomes of the creativity support assessment of *Data Exploration for Design* method shows that participants valued the exploration aspects of the method, and how the use of the method produced results worth the effort. The mindset of the method needs to lead to inferring correctly from the dataset, and that includes acknowledging biases in the data, as well as building sufficient domain and contextual understanding about the problem, in effect leading to spending time with the data.

> **RQ5:** *How do designers adopt a data-rich design methodology in design practice?*

In Chapter 6, we conducted a study to observe how the *Exploratory Data Inquiry* methodology is adopted for Frame Innovation as the overarching design methodology and studied this in a more realistic design situation. The study indicates that designers intertwine computational/data thinking with design/qualitative thinking, and the combination of these are applied on the same dataset, for example, distant reading the data by quantitative tools and closely reading the data by qualitative coding. The study confirmed that *Exploratory Data Inquiry* methodology could be well-integrated into a broader design framework, such as Frame Innovation. In such a context, the design/qualitative reasoning might take place 'outside' of the process of design inquiry through data. Another important finding is an increased understanding of the data literacy of designers focused on bias and limits of inferences from data, and the sense of capturing and working with relevant datasets, and not necessarily datasets that are already available.

From Sections 7.1 to 7.4. we synthesized the key findings related to research questions RQ1 to RQ5. We now revisit the main research question and discuss how these findings relate to the *Design Inquiry Through Data* framework.

ALIGNING THEORY AND PRACTICE

Besides the theoretical and practical knowledge contributions addressed by the research questions, the other main driver of this dissertation has been the development of methodological contributions to support future data-rich design practices. In Section 2.4, we introduced a conceptual framework that has made the development of design workshops, a methodology, a method, and tools operational. In the current section, we cumulate the lessons learned from Section 7.1 to 7.4 to inform the final *Design Inquiry Through Data* framework. As stated in Section 2.4, *Design Inquiry Through Data* refers to the design inquiry through data process in which a designer follows an opportunistic data exploration process for inquiry. We have reached an improved understanding that provides further details. Figure 7.10 shows more specifics in this process, where the designer combines computational/data thinking with qualitative/design thinking throughout all the different steps. The process can be separated into three conceptual stages, containing 'problem framing', 'exploring', and 'inferring'. While these stages appear linear, in practice, they are iterative and involve multiple sub-steps, as shown in Figure 7.10. During *Design Inquiry Through Data*, designers use design tools as process guides to support the learning curve and existing computational tools to leverage and interact with data. Prominent tools in this process are visualizations, which resemble traditional prototypes and boundary objects.

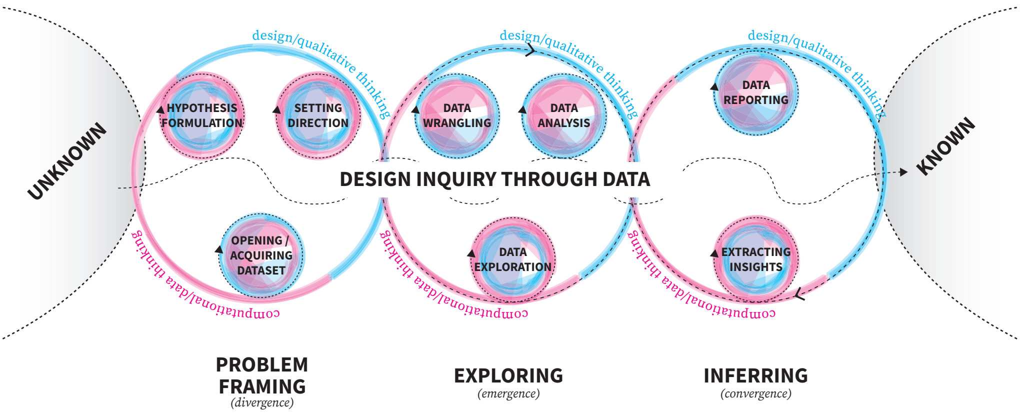


Figure 7.10. The final *Design Inquiry Through Data* framework shows the synergy between data practices and design practices, as intertwined steps of computational/data thinking (color pink) and design/qualitative thinking (color blue). In the interim steps, the color mix represents which thinking and tools are dominant (outer color is dominant)

7.6 Implications

The main intention of this dissertation was to enable more designers to be able to work with data, and open up the thinking of designers on the different, novel ways how data could contribute to design practice. Stolterman (2008) makes a detailed argument that design methods, tools, and approaches must address the nature of the design process, and need to be *designerly* in order to have uptake by design practitioners. This dissertation shows one way to handle the paradox of providing designers not only with a step-by-step guide, a very descriptive design method, but also different framings on how design inquiry through data could be utilized in a situated design practice. As have been highlighted in the dissertation earlier, when data and design intertwine at a fundamental level of inquiry, designerly uses of data emerge. We approached the research from this perspective and thus strived to develop the *Design Inquiry Through Data* framework, the *Exploratory Data Inquiry* methodology, and the *Data Exploration for Design* method, with the emphasis being on the former two. In keeping with Stolterman (2008), we focused on 'prepared-for-action', as opposed to 'guided-in-action'. Prepared-for-action is the idea that in order to have designers able to deal with the particular richnesses of complex design problems, it is impossible to prepare guidance for all sorts of design situations. The best that design education can do is to train reflective practitioners (Schön, 1984), who will be competent to adapt their design repertoire to address all sorts of problems. Following the importance of design education to influence design practice, next, we focus further on unpacking implications for design education.

The different empirical studies presented took place as design workshops with a learning agenda to enable design students to be able to work with data. From our position, the valuable way of teaching data competences to designers is focused on hands-on doing and shaping the mindset. The studies illustrate that the most valuable strategic use of data comes when designers are part of defining what data to acquire. In order to be effectively involved in defining data acquisition, designers need to be cognizant of what data can be acquired, how can it be made sense of, what can be inferred, and above all, what can be ethical to use. The dissertation contributes to design education in this regard by providing rich descriptions of studies that tell how design students learn data practices for design inquiry. We need to emphasize, that we offer an alternative for design education about data with less focus on the visualization of data, but more on thinking with data and collecting the right data for inquiry.

7.7 Ethical considerations

Although addressing designer ethics was not among the goals of this dissertation, but ethical considerations need to be addressed when working with digital data. During the different cases, participants developed their data literacy of how to use digital data, but also the limits and problematic boundaries, such as bias, or inferring based on limited data collection. In other words, there are clear benefits of empowering designers in being able to work with data – designers become more aware of data and consequences of its use. If we support designers in being more data literate, we aspire for human-centered and ethical use of people's data. It is nevertheless controversial, that some findings of this dissertation urges designers to be involved in data capture in order to get access to 'good quality' data for innovation, and thus foster the capturing of more data. When designers learn tools to collect and aggregate digital data, and often at scale, designers must be cognizant of the responsibility that comes with aggregated data. During the studies, it has been especially controversial to show the relative simplicity of such computational tasks and that it is within only a few days of programming for a master-level design student. As the common proverb goes, *"With great powers come great responsibilities"*, which is applicable in this context; with the same tools, automated marketing bots can be made to spam people, or to aggregate personal data, which can be easily abused.

Since the beginning of the dissertation's research project (2016), online data and privacy have been a turbulent space. Among the user-empowering directions, regulations such as the European Union's General Data Protection Regulation (GDPR) (*EU data protection rules*, 2020) has come to ruling, affecting public data handling on the internet to a large extent. On the other hand, news on abusing privacy and algorithms are too common from the tech industry. These circumstances make it hard to predict future opportunities and limitations on using design inquiry through data. However, as digital humanities inspired the approach, the scholars of internet research and new media can be a source of inspiration for already established ethical protocols about using digital data for research (Tiidenberg, 2020). Nevertheless, the conduction of social science research in the digital space is different from design practice, such as less scrutiny than academic peer review. Thus, ethical considerations of designers using exploratory data inquiry need to be taken into account. A reasonable and straightforward principle to follow is to self-reflect: *'Would such a research approach or handling of my data make me uncomfortable?'* If the answer is yes, then it is probably a problematic approach. Furthermore, making anonymization of data more accessible, such as with code snippet examples, could be towards mitigating risks.

7.8 Recommendations for future research

This dissertation has been approached from a strong position that future designers will face an increasing amount of data in their practice, and enabling them to work with data would contribute to 'future-proofing' their career. To explore data-rich design practices, we conducted an exploratory research project, which signposts new questions to follow in the future.

EXPERIMENTAL SETUPS

Future studies on the *Exploratory Data Inquiry* methodology could address different research variables: experimental setup of groups with and without *Exploratory Data Inquiry*, different timescales, and different participant demographics.

EXPERIMENTAL DESIGNS FOR VALIDATION: Validating the added value of design methodologies, such as Exploratory Data Inquiry, could be achieved with experimental setups comparing groups with and without using Exploratory Data Inquiry. As currently the shortest studies have been conducted as one-day workshops, this format could be expanded with monitoring groups without Exploratory Data Inquiry.

DIFFERENT TIMESCALES: An alternative exploration could be to study the use of Exploratory Data Inquiry over a longer time-frame, such as a complete semester project. A longer time-frame would enable to the mastery of the learning curve, observing how gained proficiency influences the approach and to observe how a more extensive design process would unfold when approached through data.

DIFFERENT PARTICIPANT DEMOGRAPHICS: Design practitioners would be appropriate demographics to study next, based on the established Exploratory Data Inquiry methodology from the current work. While design practitioners can already have developed approaches and favorite tools and methods to be used, their perspective in using Exploratory Data Inquiry could reveal new aspects that might have been approached naively before. Besides design practitioners, it is also interesting to consider alternative demographics, such as data scientists or non-expert data communities, to investigate how their work can be seen with the lenses of a creative process, as used in Chapter 4. Whereas creativity is deeply ingrained in problem-solving, it is likely that how it manifests in the context of data practices will inspire new tools beyond design.

NEW INTERFACES FOR COMBINING BIG DATA AND QUALITATIVE DATA INQUIRY

While the notion of combining big data and 'thick data' (i.e., qualitative data) has been around for years, it has been unclear how can they be combined on a tools level. The outcomes of the dissertation show the possibility for new conceptualizations by combining these two aspects. For example, abductive sensemaking (and framing) have been an important notion for sensemaking of qualitative data, yet existing tools only support aggregated operations on datasets, such as visualizing time series. Integrating interactions with 'big data', such as time series or network graphs interfaces, could be to be combined with close/granular reading of datapoints on the same interface. Additionally, another main finding has been the importance of designers being involved in defining the data collection. For that, data querying and collection (e.g., scraping) in the same interface would provide tools that support an ideal scenario of exploratory data inquiry.

USING EMERGING DATA TECHNOLOGY FOR DESIGN METHODS

Beyond strengthening the conclusions around *Exploratory Data Inquiry* methodology, the dissertation indicates broader research opportunities as well. First, the use of emerging technology, such as artificial intelligence algorithms for computer vision or natural language processing, could be utilized as techniques to enrich *how* we design. To clarify, we mean to use emerging technology not for form-giving as part of the designed solution, but to augment the cognition, intelligence, and creativity of designers while designing. Some elements of this dissertation have already illustrated that techniques such as quantitative text analysis or network visualizations are within reach for designers, and new techniques could be made more accessible as well. As different off-the-shelf algorithms and data practices becoming only more-and-more approachable, it is just a matter of effort to scaffold and tailor them to be useful for designers. For that quest, this dissertation can be a reference point to navigate how designers think and act when data is integrated into their practice.

References

Alspaugh, S., Zokaei, N., Liu, A., Jin, C., & Hearst, M. A. (2019). Futzing and Moseying: Interviews with Professional Data Analysts on Exploration Practices. *IEEE Transactions on Visualization and Computer Graphics, 25*(1), 22–31. https://doi.org/10.1109/TVCG.2018.2865040

Anand, S. S., Bell, D. A., & Hughes, J. G. (1995). The role of domain knowledge in data mining. *Proceedings of the Fourth International Conference on Information and Knowledge Management*, 37–43. https://doi.org/10.1145/221270.221321

Bastian, M., Heymann, S., Jacomy, M., & others. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. *Proceedings of the Third International ICWSM Conference, 8*(2009), 361–362.

Baumer, B. (2015). A Data Science Course for Undergraduates: Thinking With Data. *The American Statistician, 69*(4), 334–342. https://doi.org/10.1080/00031305.2015.1081105

Bijl-Brouwer, M. van der, & Dorst, K. (2017). Advancing the strategic impact of human-centred design. *Design Studies, 53,* 1–23. https://doi.org/10.1016/j.destud.2017.06.003

Bijl-Brouwer, M. van der. (2019). Problem Framing Expertise in Public and Social Innovation. *She Ji: The Journal of Design, Economics, and Innovation, 5*(1), 29–43. https://doi.org/10.1016/j.sheji.2019.01.003

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. *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, 301–311. https://doi.org/10.1145/2901790.2901855

Bogers, S., van Kollenburg, J., Deckers, E., Frens, J., & Hummels, C. (2018). A situated exploration of designing for personal health ecosystems through data-enabled design. *Proceedings of the 2018 Designing Interactive Systems Conference,* 109–120. https://doi.org/10.1145/3196709.3196769

Borra, E., & Rieder, B. (2014). Programmed Method: Developing a

Toolset for Capturing and Analyzing tweets. *Aslib Journal of Information Management, 66*(3), 262–278. https://doi.org/10.1108/AJIM-09-2013-0094

boyd, d., & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society, 15*(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878

Brooker, P., Barnett, J., & Cribbin, T. (2016). Doing social media analytics. *Big Data & Society, 3*(2), 2053951716658060. https://doi.org/10.1177/2053951716658060

Bruns, A., & Liang, Y. (2012). Tools and methods for capturing Twitter data during natural disasters. *First Monday, 17*(4). https://doi.org/10.5210/fm.v17i4.3937

Buchanan, R. (1992). Wicked Problems in Design Thinking. *Design Issues, 8*(2), 5–21. JSTOR. https://doi.org/10.2307/1511637

Buchanan, R. (2001). Design Research and the New Learning. *Design Issues, 17*(4), 3–23. https://doi.org/10.1162/07479360152681056

Cao, L. (2017). Data Science: A Comprehensive Overview. *ACM Computing Surveys (CSUR), 50*(3), 43. https://doi.org/10.1145/3076253

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann.

Cardoso, C., Badke-Schaub, P., & Eris, O. (2016). Inflection moments in design discourse: How questions drive problem framing during idea generation. *Design Studies, 46*, 59–78. https://doi.org/10.1016/j.destud.2016.07.002

Carlson, J., & Johnston, L. (2015). *Data Information Literacy: Librarians, Data, and the Education of a New Generation of Researchers*. Purdue University Press. https://docs.lib.purdue.edu/purduepress_ebooks/43

Cherry, E., & Latulipe, C. (2014). Quantifying the Creativity Support of Digital Tools Through the Creativity Support Index. *ACM Trans. Comput.-Hum. Interact., 21*(4), 21:1–21:25. https://doi.org/10.1145/2617588

Coulson, S., Woods, M., Scott, M., Hemment, D., & Balestrini, M. (2018). Stop the Noise! Enhancing Meaningfulness in Participatory

Sensing with Community Level Indicators. Proceedings of the 2018 Designing Interactive Systems Conference, 1183–1192. https://doi.org/10.1145/3196709.3196762

Crabtree, A., & Mortier, R. (2015). Human Data Interaction: Historical Lessons from Social Studies and CSCW. In *ECSCW 2015: Proceedings of the 14th European Conference on Computer Supported Cooperative Work, 19-23 September 2015, Oslo, Norway* (pp. 3–21). Springer International Publishing. https://doi.org/10.1007/978-3-319-20499-4_1

Creswell, J. W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications.

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and Conducting Mixed Methods Research*. SAGE Publications.

Cross, N. (1982). Designerly ways of knowing. *Design Studies, 3*(4), 221–227. https://doi.org/10.1016/0142-694X(82)90040-0

Cross, N. (2004). Expertise in design: An overview. *Design Studies, 25*(5), 427–441. https://doi.org/10.1016/j.destud.2004.06.002

Daalhuizen, J. J. (2014). *Method Usage in Design: How methods function as mental tools for designers*. Delft University of Technology.

Dalsgaard, P. (2014). Pragmatism and Design Thinking. *International Journal of Design, 8*(1), 143–155.

Dalsgaard, P. (2017). Instruments of Inquiry: Understanding the Nature and Role of Tools in Design. *International Journal of Design, 11*(1), 21–33.

Deng, Y., Antle, A. N., & Neustaedter, C. (2014). Tango cards: A card-based design tool for informing the design of tangible learning games. *Proceedings of the 2014 Conference on Designing Interactive Systems*, 695–704. https://doi.org/10.1145/2598510.2598601

Deutsch, R. (Ed.). (2015). *Data-Driven Design and Construction*. John Wiley & Sons, Inc. https://doi.org/10.1002/9781119149682

Digital Methods Initiative – Tools Database. (2020). Retrieved April 18, 2020, from https://wiki.digitalmethods.net/Dmi/ToolDatabase

Dixon, B. (2019). Experiments in Experience: Towards an Alignment of Research through Design and John Dewey's Pragmatism. *Design Issues*, 35(2), 5–16. https://doi.org/10.1162/desi_a_00531

D'Ignazio, C. (2017). Creative data literacy. *Information Design Journal, 23*(1), 6–18. https://doi.org/10.1075/idj.23.1.03dig

D'Ignazio, C., & Bhargava, R. (2016). DataBasic: Design principles, tools and activities for data literacy learners. *The Journal of Community Informatics, 12*(3).

Dorst, K. (2006). Design Problems and Design Paradoxes. *Design Issues, 22*(3), 4–17. https://doi.org/10.1162/desi.2006.22.3.4

Dorst, K. (2011). The Core of 'Design Thinking' and its Application. *Design Studies, 32*(6), 521–532. https://doi.org/10.1016/j.destud.2011.07.006

Dorst, K. (2015a). Frame Creation and Design in the Expanded Field. *She Ji: The Journal of Design, Economics, and Innovation, 1*(1), 22–33. https://doi.org/10.1016/j.sheji.2015.07.003

Dorst, K. (2015b). *Frame innovation: Create new thinking by design*. MIT press.

Dorst, K., & Cross, N. (2001). Creativity in the Design Process: Co-evolution of Problem-Solution. *Design Studies, 22*(5), 425–437. https://doi.org/10.1016/S0142-694X(01)00009-6

Dourish, P. (2006). Implications for design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 541–550. https://doi.org/10.1145/1124772.1124855

Dove, G. (2015). *CoDesign with data*. City University London. https://openaccess.city.ac.uk/id/eprint/14902/

Dove, G., Halskov, K., Forlizzi, J., & Zimmerman, J. (2017). UX Design Innovation: Challenges for Working with Machine Learning As a Design Material. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 278–288. https://doi.org/10.1145/3025453.3025739

Dove, G., & Jones, S. (2014). Using data to stimulate creative thinking in the design of new products and services. *Proceedings of the 2014 Conference on Designing Interactive Systems*, 443–452. https://doi.org/10.1145/2598510.2598564

EU data protection rules. (2020). European Commission - European Commission. Retrieved April 19, 2020, from https://ec.europa.eu/info/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules/eu-data-protection-rules_en

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *AI Magazine, 17*(3), 37–37. https://doi.org/10.1609/aimag.v17i3.1230

Feinberg, M. (2017). A Design Perspective on Data. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2952–2963. https://doi.org/10.1145/3025453.3025837

Gaver, B., Beaver, J., & Benford, S. (2003). Ambiguity as a Resource for Design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 233–240. https://doi.org/10.1145/642611.642653

Gero, J. S. (1996). Creativity, emergence and evolution in design. *Knowledge-Based Systems*, *9*(7), 435–448. https://doi.org/10.1016/S0950-7051(96)01054-4

Giaccardi, E. (2019). Integrating Human and Nonhuman Perspectives in Data-Enabled RTD. In *Resourceful Ageing*.

Giaccardi, E., Cila, N., Speed, C., & Caldwell, M. (2016). Thing Ethnography: Doing Design Research with Non-Humans. *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, 377–387. https://doi.org/10.1145/2901790.2901905

Gilbert, E., & Karahalios, K. (2009). Predicting Tie Strength with Social Media. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 211–220. https://doi.org/10.1145/1518701.1518736

Gray, C. M. (2016). 'It's More of a Mindset Than a Method': UX Practitioners' Conception of Design Methods. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 4044–4055. https://doi.org/10.1145/2858036.2858410

Gray, J., Chambers, L., & Bounegru, L. (2012). *The Data Journalism Handbook: How Journalists Can Use Data to Improve the News*. O'Reilly Media.

Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. *Big Data & Society*, *5*(2), 2053951718786316. https://doi.org/10.1177/2053951718786316

Guindon, R. (1990). Designing the design process: Exploiting opportunistic thoughts. *Human-Computer Interaction*, *5*(2), 305–344. https://doi.org/10.1080/07370024.1990.9667157

Hargittai, E. (2015). Is Bigger Always Better? Potential Biases of Big Data Derived from Social Network Sites. *The ANNALS of the American Academy of Political and Social Science*, *659*(1), 63–76. https://doi.org/10.1177/0002716215570866

He, S., & Adar, E. (2017). VizItCards: A Card-Based Toolkit for

Infovis Design Education. *IEEE Transactions on Visualization and Computer Graphics, 23*(1), 561–570. https://doi.org/10.1109/TVCG.2016.2599338

Hill, B. M., Dailey, D., Guy, R. T., Lewis, B., Matsuzaki, M., & Morgan, J. T. (2017). Democratizing Data Science: The Community Data Science Workshops and Classes. In S. A. Matei, N. Jullien, & S. P. Goggins (Eds.), *Big Data Factories: Collaborative Approaches* (pp. 115–135). Springer International Publishing. https://doi.org/10.1007/978-3-319-59186-5_9

Hornecker, E. (2010). Creative Idea Exploration Within the Structure of a Guiding Framework: The Card Brainstorming Game. *Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction,* 101–108. https://doi.org/10.1145/1709886.1709905

Jansen, Y., Dragicevic, P., Isenberg, P., Alexander, J., Karnik, A., Kildal, J., Subramanian, S., & Hornbæk, K. (2015). Opportunities and Challenges for Data Physicalization. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems,* 3227–3236. https://doi.org/10.1145/2702123.2702180

Jones, J. C. (1970). *Design Methods: Seeds of Human Futures*. Wiley-Interscience.

Jones, J. C. (1992). *Design Methods*. John Wiley & Sons.

Kandel, S., Heer, J., Plaisant, C., Kennedy, J., Van Ham, F., Riche, N. H., Weaver, C., Lee, B., Brodbeck, D., & Buono, P. (2011). Research Directions in Data Wrangling: Visualizations and Transformations for Usable and Credible Data. *Information Visualization, 10*(4), 271–288. https://doi.org/10.1177/1473871611415994

Kaner, S. (2014). *Facilitator's guide to participatory decision-making*. John Wiley & Sons.

King, R., Churchill, E. F., & Tan, C. (2017). *Designing with Data: Improving the User Experience with A/B Testing* (1st ed.). O'Reilly Media, Inc.

Kitchin, R. (2014a). *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. SAGE Publications. https://doi.org/10.4135/9781473909472

Kitchin, R. (2014b). Big Data, new epistemologies and paradigm shifts. *Big Data & Society, 1*(1). https://doi.org/10.1177/2053951714528481

Kling, R., & Scacchi, W. (1982). *The Web of Computing: Computer Technology as Social Organization* (M. C. Yovits, Ed.; Vol. 21, pp. 1–90). Elsevier. https://doi.org/10.1016/S0065-2458(08)60567-7

Kolko, J. (2009). Abductive Thinking and Sensemaking: The Drivers of Design Synthesis. *Design Issues, 26*(1), 15–28. https://doi.org/10.1162/desi.2010.26.1.15

Kouprie, M., & Sleeswijk Visser, F. (2009). A framework for empathy in design: Stepping into and out of the user's life. *Journal of Engineering Design, 20*(5), 437–448. https://doi.org/10.1080/09544820902875033

Kozinets, R. V. (2002). The Field Behind the Screen: Using Netnography for Marketing Research in Online Communities. *Journal of Marketing Research, 39*(1), 61–72. https://doi.org/10.1509/jmkr.39.1.61.18935

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., & Gutmann, M. (2009). Computational social science. *Science, 323*(5915), 721–723.

Lewin, K. (1946). Action Research and Minority Problems. *Journal of Social Issues, 2*(4), 34–46. https://doi.org/10/bkf3rx

Li, I., Dey, A., & Forlizzi, J. (2010). A Stage-Based Model of Personal Informatics Systems. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 557–566. https://doi.org/10.1145/1753326.1753409

Lim, Y.-K., Stolterman, E., & Tenenberg, J. (2008). The Anatomy of Prototypes: Prototypes as Filters, Prototypes as Manifestations of Design Ideas. *ACM Trans. Comput.-Hum. Interact., 15*(2). https://doi.org/10.1145/1375761.1375762

Litt, E., & Hargittai, E. (2016). 'Just Cast the Net, and Hopefully the Right Fish Swim into It': Audience Management on Social Network Sites. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing,* 1488–1500. https://doi.org/10.1145/2818048.2819933

Liu, Y.-C., Chakrabarti, A., & Bligh, T. (2003). Towards an 'ideal' approach for concept generation. *Design Studies, 24*(4), 341–355. https://doi.org/10.1016/S0142-694X(03)00003-6

Louridas, P. (1999). Design as Bricolage: Anthropology Meets Design Thinking. *Design Studies, 20*(6), 517–535. https://doi.org/10.1016/S0142-694X(98)00044-1

Lycett, M. (2013). 'Datafication': Making Sense of (Big) Data in a Complex World. *European Journal of Information Systems, 22*(4), 381–386. https://doi.org/10.1057/ejis.2013.10

Maher, M. L., Poon, J., & Boulanger, S. (1996). Formalising Design Exploration as Co-Evolution. In J. S. Gero & F. Sudweeks (Eds.), *Advances in Formal Design Methods for CAD: Proceedings of the IFIP WG5.2 Workshop on Formal Design Methods for Computer-Aided Design, June 1995* (pp. 3–30). Springer US. https://doi.org/10.1007/978-0-387-34925-1_1

Marchionini, G. (2006). Exploratory Search: From Finding to Understanding. *Communications of the ACM*, *49*(4), 41–46. https://doi.org/10.1145/1121949.1121979

Marres, N., & Weltevrede, E. (2013). Scraping the Social? *Journal of Cultural Economy*, *6*(3), 313–335. https://doi.org/10.1080/17530350.2013.772070

Mattelmäki, T. (2006). *Design probes*. Aalto University.

Mauri, M., Elli, T., Caviglia, G., Uboldi, G., & Azzi, M. (2017). RAWGraphs: A Visualisation Platform to Create Open Outputs. *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter*, 1–5. https://doi.org/10.1145/3125571.3125585

Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data: A Revolution that Will Transform how We Live, Work, and Think*. Houghton Mifflin Harcourt.

Mitchell, R. (2018). *Web Scraping with Python: Collecting More Data from the Modern Web*. O'Reilly Media, Inc.

Moretti, F. (2013). *Distant Reading*. Verso Books.

Mortier, R., Haddadi, H., Henderson, T., McAuley, D., & Crowcroft, J. (2014). Human-Data Interaction: The Human Face of the Data-Driven Society. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2508051

Neill, C. (2020). *People are Your Data*. Dscout. Retrieved April 19, 2020, from https://dscout.com/people-nerds/people-are-your-data-tricia-wang

Nelson, H. G., & Stolterman, E. (2012). *The Design Way: Intentional Change in an Unpredictable World*.

Norman, D. A., & Stappers, P. J. (2015). DesignX: Complex

Sociotechnical Systems. *She Ji: The Journal of Design, Economics, and Innovation*, *1*(2), 83–106. https://doi.org/10.1016/j.sheji.2016.01.002

O'Neil, C., & Schutt, R. (2013). *Doing Data Science: Straight Talk from the Frontline*. O'Reilly Media.

OpenRefine. (2020). *OpenRefine—A Free, Open Source, Powerful Tool for Working with Messy Data*. Retrieved April 18, 2020, from https://openrefine.org

Passi, S., & Jackson, S. (2017). Data Vision: Learning to See Through Algorithmic Abstraction. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 2436–2447. https://doi.org/10.1145/2998181.2998331

Pink, S. (2016). Digital ethnography. *Innovative Methods in Media and Communication Research*, 161–165.

Redström, J. (2017). *Making Design Theory*. MIT Press.

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*(2), 155–169. https://doi.org/10.1007/BF01405730

Rogers, R. (2013). *Digital Methods*. The MIT Press.

Rogers, R. (2017). Foundations of Digital Methods—Query Design. In *The Datafied Society. Studying Culture through Data*. Amsterdam University Press.

Roy, R., & Warren, J. P. (2019). Card-based design tools: A review and analysis of 155 card decks for designers and designing. *Design Studies*, *63*, 125–154. https://doi.org/10.1016/j.destud.2019.04.002

Sanders, E. B.-N., Brandt, E., & Binder, T. (2010). A framework for organizing the tools and techniques of participatory design. *Proceedings of the 11th Biennial Participatory Design Conference*, 195–198. https://doi.org/10.1145/1900441.1900476

Sanders, E. B.-N., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. *CoDesign*, *4*(1), 5–18. https://doi.org/10.1080/15710880701875068

Sanders, E. B.-N., & Stappers, P. J. (2014). Probes, toolkits and prototypes: Three approaches to making in codesigning. *CoDesign*, *10*(1), 5–14. https://doi.org/10.1080/15710882.2014.888183

Sawyer, R. K. (2011). *Explaining creativity: The science of human innovation*. Oxford University Press.

Schön, D. A. (1984). *The reflective practitioner: How professionals think in action*. Basic books.
Schön, D. A. (1988). Designing: Rules, types and worlds. *Design Studies*, 9(3), 181–190. https://doi.org/10.1016/0142-694X(88)90047-6
Schønheyder, J. F., & Nordby, K. (2018). The use and evolution of design methods in professional design practice. *Design Studies*, *58*, 36–62. https://doi.org/10.1016/j.destud.2018.04.001
School of Data. (2016). School of Data - Online Courses. Retrieved April 18, 2020, from https://schoolofdata.org/courses/
Sinclair, S., & Rockwell, G. (2012). Teaching Computer-Assisted Text Analysis: Approaches to Learning New Methodologies. *Digital Humanities Pedagogy: Practices, Principles, and Politics. Ed. Brett D. Hirsch. Cambridge: Open Book*, 241–63.
Sinclair, S., Rockwell, G., & others. (2018). *Voyant tools*. Retrieved April 18, 2020, from http://voyant-tools.org/
Sleeswijk Visser, F., Stappers, P. J., Lugt, R. van der, & Sanders, E. B.-N. (2005). Contextmapping: Experiences from practice. *CoDesign*, 1(2), 119–149. https://doi.org/10.1080/15710880500135987
Speed, C., & Oberlander, J. (2016). Designing from, with and by Data: Introducing the ablative framework. *Proceedings of the the 50th Anniversary Conference of the Design Research Society*.
Stappers, P. J., & Giaccardi, E. (2017). Research through Design. In M. Soegaard & R. Friis-Dam (Eds.), *The Encyclopedia of Human-Computer Interaction* (2nd ed., pp. 1–94). The Interaction Design Foundation.
Star, S. L., & Griesemer, J. R. (1989). Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, *19*(3), 387–420. https://doi.org/10.1177%2F030631289019003001
Stieglitz, S., & Dang-Xuan, L. (2013). Social Media and Political Communication: A Social Media Analytics Framework. *Social Network Analysis and Mining*, *3*(4), 1277–1291. https://doi.org/10.1007/s13278-012-0079-3
Stolterman, E. (2008). The nature of design practice and implications for interaction design research. *International Journal of Design*, *2*(1), 55–65.
Taylor, A. S., Lindley, S., Regan, T., Sweeney, D., Vlachokyriakos, V.,

Grainger, L., & Lingel, J. (2015). Data-in-Place: Thinking Through the Relations Between Data and Community. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2863–2872. https://doi.org/10.1145/2702123.2702558

Tiidenberg, K. (2020). *The SAGE Handbook of Qualitative Data Collection* (By pages 466-479). SAGE Publications Ltd. https://doi.org/10.4135/9781526416070

Tiidenberg, K., & Gómez Cruz, E. (2015). Selfies, Image and the Re-making of the Body. *Body & Society*, *21*(4), 77–102. https://doi.org/10.1177/1357034X15592465

Toolkit Navigator. (2020). *Observatory of Public Sector Innovation*. Retrieved April 18, 2020, from https://oecd-opsi.org/toolkit-navigator/

Tukey, J. W. (1962). The future of data analysis. *The Annals of Mathematical Statistics*, *33*(1), 1–67.

Tukey, J. W. (1977). *Exploratory data analysis*. Massachusetts: Addison-Wesley.

Vallgårda, A., & Fernaeus, Y. (2015). Interaction Design As a Bricolage Practice. *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, 173–180. https://doi.org/10.1145/2677199.2680594

van Boeijen, A., Daalhuizen, J., & Zijlstra, J. (2020). *Delft Design Guide (revised Edition): Perspectives—Models—Approaches—Methods*. BIS Publishers.

van Kollenburg, J., Bogers, S., Rutjes, H., Deckers, E., Frens, J., & Hummels, C. (2018). Exploring the Value of Parent Tracked Baby Data in Interactions with Healthcare Professionals: A Data-Enabled Design Exploration. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 297:1–297:12. https://doi.org/10.1145/3173574.3173871

Van Manen, M. (1990). *Researching Lived Experience: Human Science for an Action Sensitive Pedagogy*. State University of New York Press.

Venturini, T., Jacomy, M., Bounegru, L., & Gray, J. (2019). Visual Network Exploration for Data Journalists. In *The Routledge Handbook to Developments in Digital Journalism Studies*.

Venturini, T., Jacomy, M., & Jensen, P. (2019). What Do We See When

We Look at Networks. An Introduction to Visual Network Analysis and Force-Directed Layouts. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3378438

Vermaas, P. (2016). A logical critique of the expert position in design research: Beyond expert justification of design methods and towards empirical validation. *Design Science*, *2*, e7. https://doi.org/10.1017/dsj.2016.6

Vertesi, J., & Dourish, P. (2011). The Value of Data: Considering the Context of Production in Data Economies. *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*, 533–542. https://doi.org/10.1145/1958824.1958906

Waller, M. A., & Fawcett, S. E. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics*, *34*(2), 77–84. https://doi.org/10.1111/jbl.12010

Wang, T. (2013). *Big Data Needs Thick Data*. Ethnography Matters. Retrieved April 18, 2020, from http://ethnographymatters.net/blog/2013/05/13/big-data-needs-thick-data/

Wickham, H. (2014). Tidy Data. *Journal of Statistical Software, Articles*, *59*(10), 1–23. https://doi.org/10.18637/jss.v059.i10

Wolff, A., Gooch, D., Montaner Cavero, J., Rashid, U., & Kortuem, G. (2016). Creating an Understanding of Data Literacy for a Data-driven Society. *The Journal of Community Informatics*, *12*(3).

Wong, P. C., & Thomas, J. (2004). Guest Editors' Introduction–Visual Analytics. *IEEE Computer Graphics and Applications, 24(5):20-21*, *24*(5), 20–21. https://doi.org/10.1109/MCG.2004.39

Yang, Q., Scuito, A., Zimmerman, J., Forlizzi, J., & Steinfeld, A. (2018). Investigating How Experienced UX Designers Effectively Work with Machine Learning. *Proceedings of the 2018 Designing Interactive Systems Conference*, 585–596. https://doi.org/10.1145/3196709.3196730

Yu, C. H. (1994). Abduction? Deduction? Induction? Is There a Logic of Exploratory Data Analysis? *Annual Meeting of American Educational Research Association*.

Summary

The emergence of the internet and subsequent massive data collection and storage is creating vast opportunities for design research and practice. In this dissertation, we investigate the interrelationship between design and data science practices and explore data as a new creative lens for design inquiry. While digital data has been increasingly used by designers, such as using A/B testing to drive design decisions for internet products, data has been less explored as a resource for inquiry about the world. Despite how data-connected artifacts increasingly facilitate human interactions, designers' repertoire still primarily relies on practices established for inquiring in the physical world. The current industry practice of integrating data scientists into the design team is neither affordable nor feasible to apply across the vast majority of contexts and cases where design operates. To address these problems, in this dissertation, we aim to deepen the theoretical and practical knowledge on the intersection of design and data science, and to develop methodological contributions to support future data-rich design practices.

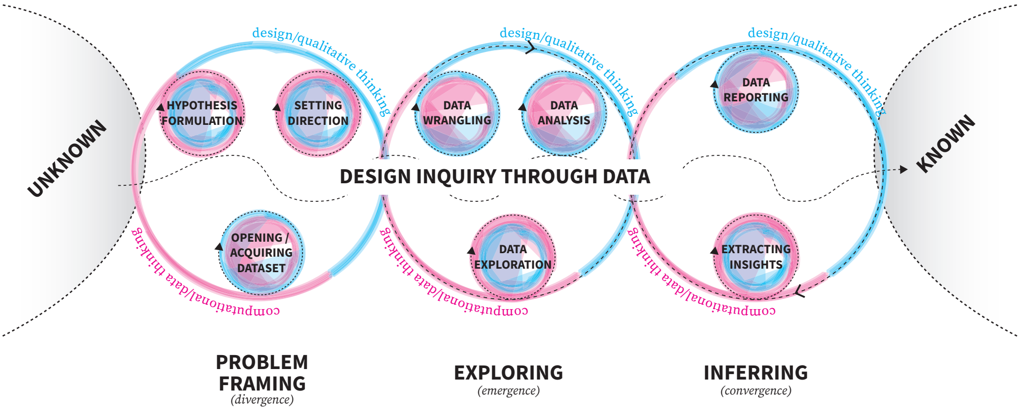
The main research question we pursue in this dissertation is "How can designers integrate data practices into design inquiry?" We address this question through conducting a Research-through-Design program to gain, on the one hand, a better understanding of how the fields of design and data science intersect, and on the other hand, to develop methodological contributions for future data-rich design practices. The resulting conceptual framework of Design Inquiry Through Data has been constructed throughout a series of empirical studies in which data-rich design practices are studied. For each study, practical data methods and techniques have been curated and/or developed.

The first study shows that designers' creativity takes new ways as they appropriate data practices. We observe how designerly sensemaking creatively repurposes non-expert data tools, initially developed for the needs of data analysis, for data exploration. The second study shows that when designers research human experiences through data, the designers' empathy takes new manifestations. In order to incorporate

human needs into design inquiry, designers' empathy needs to be applied to the designing of search queries. These early insights of how designers' creative approaches take new shapes working with data leads the research to investigate data practices as a creative process further. Framing data practices as part of a creative process has enabled the alignment of data exploration practices methodologically into design practice. Based on this alignment we conclude on the Exploratory Data Inquiry methodology, which enables designers to see data practices through a design process lens, and in that way combining data practices into design. Exploratory Data Inquiry establishes three conceptual steps of 'problem framing', 'exploring', and 'inferring', which makes data practices applicable in different types of design activities.

In the third study, we use the Exploratory Data Inquiry methodology for developing a design method. The corresponding Data Exploration for Design method sets up an outline for using data exploration integrated into a design process and provides tools to guide the learning curve. We use the method to understand the emerging thinking patterns better when data exploration is used for design inquiry. Throughout the studies we have noticed that data exploration influences the framing and reframing practices of designers. To more directly study this change of practices, during the fourth and last study, we investigate how the Exploratory Data Inquiry methodology is adopted in the broader design methodology of Frame Innovation. The last study reveals that designers assume a 'hybrid mindset' of combining computational/data thinking with design/qualitative thinking. The hybrid mindset enables designers to opportunistically mix the two main data traditions of qualitative and quantitative data, introducing creative and novel uses of data analysis tools, and switching inquiries on the same dataset seamlessly. In this creative utilization of data, the different tools and techniques for leveraging data become appropriated in ways unique to design. Visualizations take the role of prototypes and boundary objects in the design process. In such a role, visualizations are not made to create representations for the sake of reporting the data, but for 'greasing' human cognition. In this support for understanding, visualizations are made rapidly for understanding a particular phenomenon through the dataset.

The different findings of the dissertation come together in the Design Inquiry Through Data framework. The framework describes how data practices can be integrated into a generic design inquiry process that moves from an “unknown” state to a “known” state. The framework shows the process how a designer combines computational/data thinking with qualitative/design thinking throughout three conceptual steps of ‘problem framing’, ‘exploring’, and ‘inferring’. Moving through these three stages involves intertwined design and data-related activities, where not only the data is repurposed in creative ways in the design process, but also the data activities themselves are approached in a designerly fashion.



Design Inquiry Through Data. The framework shows the synergy between data practices and design practices, as intertwined steps of computational/data thinking (color pink) and design/qualitative thinking (color blue). In the interim steps, the color mix represents which thinking and tools are dominant (outer color is dominant).

The dissertation provides practical, methodological support to bring data practices into design inquiry. By reframing data practices in a designerly perspective, designers are able to capitalize on existing data and existing data techniques better and without relying on data scientists. In practical terms, the dissertation provides designerly contributions as a methodology, a method, design tools, and a framework. These new design techniques formalize a future data-rich design practice and provide a versatile perspective on how data could be used broadly for design inquiry. These contributions are presented

along with rich empirical insights about how novice designers use design inquiry through data. These insights are also valuable for design educators and designers of non-expert data tools.

Samenvatting

De opkomst van het internet en de daaropvolgende massale dataverzameling en -opslag creëert enorme kansen voor ontwerponderzoek en ontwerppraktijk. In dit proefschrift onderzoeken we de onderlinge relatie tussen de praktijken van de ontwerpdiscipline en datawetenschappen, en verkennen we data als een nieuwe creatieve lens voor ontwerponderzoek. Hoewel digitale data in toenemende mate door ontwerpers worden gebruikt, zoals bijvoorbeeld A/B-testen om ontwerpbeslissingen voor internet-producten te sturen, is er minder onderzoek gedaan naar het gebruik van data als middel om de wereld te begrijpen. Ondanks dat met data-verbonden artefacten menselijke interacties steeds vaker mediëren, is het repertoire van ontwerpers nog voornamelijk gebaseerd op interacties in de fysieke wereld. De huidige industriepraktijk waar datawetenschappers in ontwerpteam geïntegreerd zijn, is vaak niet schaalbaar naar andere ontwerpcontexten en -praktijken. Om deze problemen aan te pakken, willen we in dit proefschrift de theoretische en praktische kennis op het snijvlak van ontwerp- en datawetenschappen verdiepen alsook methodologische bijdragen ontwikkelen om toekomstige datarijke ontwerppraktijken te ondersteunen.

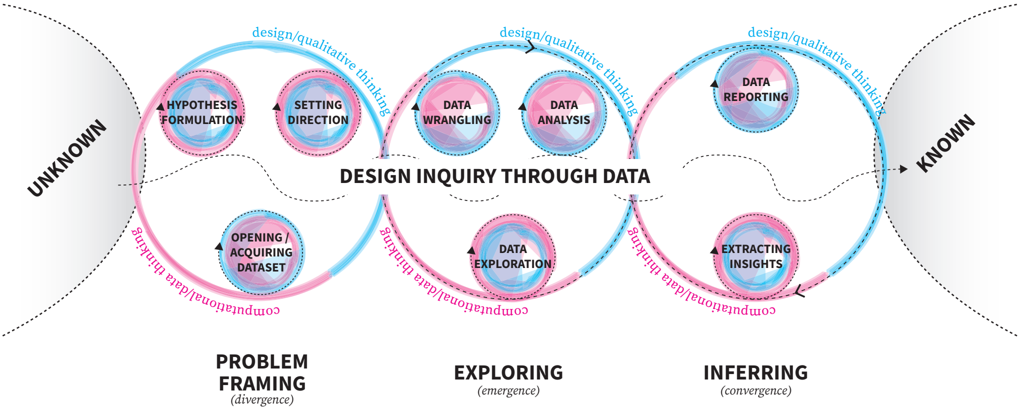
De onderzoeksvraag die in dit proefschrift centraal staat is: "Hoe kunnen ontwerpers datawetenschapspraktijken integreren in ontwerponderzoek?" We pogen deze vraag te beantwoorden met behulp van een Research-through-Design aanpak zodat we enerzijds een beter begrip krijgen van hoe de velden van ontwerp en data science met elkaar overlappen, en anderzijds om methodologische bijdragen te ontwikkelen voor toekomstige data-rijke ontwerppraktijken. Het resulterende conceptuele raamwerk Design Inquiry Through Data is ontwikkeld gedurende een reeks empirische studies waarin datarijke ontwerppraktijken zijn bestudeerd. Voor elke studie zijn methoden en -technieken samengesteld en ontwikkeld, die in de praktijk te gebruiken zijn.

Uit de eerste studie blijkt dat de creativiteit van ontwerpers nieuwe

wegen inslaat bij het toepassen van datawetenschapspraktijken. We zien hoe ontwerpers op creatieve wijze beschikbare data-technieken, die oorspronkelijk ontwikkeld zijn voor dataanalyse en -verkenning op andere wijzen inzetten en hergebruiken. De tweede studie laat zien dat wanneer ontwerpers menselijke ervaringen onderzoeken met behulp van data, de empathie van ontwerpers nieuwe vormen aanneemt. Om menselijke behoeften in het ontwerponderzoek een plek te geven moet de empathie van ontwerpers worden toegepast op het ontwerpen van zoekopdrachten. Deze vroege inzichten over hoe de creatieve benaderingen van ontwerpers nieuwe vormen aannemen door met data te werken, leiden ertoe dat het benutten van data voor een creatief proces verder worden bestudeerd. Door datawetenschapspraktijken te beschouwen als onderdeel van een creatief proces, wordt het mogelijk om methoden voor dataverkenning methodologisch af te stemmen specifiek op de ontwerppraktijk. Deze afstemming heeft geleid tot de Exploratory Data Inquiry-methodologie, die ontwerpers in staat stelt om datawetenschapspraktijken door de lens van het ontwerpproces te beschouwen, en op die manier deze te combineren in de ontwerppraktijk. Exploratory Data Inquiry stelt drie conceptuele stappen voor van ‘probleemdefinitie’, ‘verkennen’ en ‘interpreteren’, waardoor datawetenschapspraktijken toepasbaar zijn in verschillende soorten ontwerpactiviteiten.

In het derde onderzoek gebruiken we de Exploratory Data Inquiry-methodologie voor het ontwikkelen van een ontwerpmethode. De bijbehorende methode Data Exploration for Design biedt een leidraad om het verkenning van data in het ontwerpproces te integreren alsook richtlijnen om het leerproces te faciliteren. We gebruiken de Data Exploration for Design methode om beter te begrijpen wat de denkpatronen zijn die opkomen wanneer dataverkenning wordt gebruikt voor ontwerponderzoek. Gedurende de studies hebben we gemerkt dat data-exploratie de framing- en reframing-praktijken van ontwerpers beïnvloedt. Om deze verandering directer te bestuderen, onderzoeken we tijdens de vierde en laatste studie hoe de Exploratory Data Inquiry-methodologie wordt toegepast in de bredere ontwerpmethodologie Frame Innovation. Uit de laatste studie blijkt dat ontwerpers veelal uitgaan van een ‘hybride mentaliteit' die kwantitatief data-denken met kwalitatief design-denken combineert. De hybride mentaliteit stelt ontwerpers in staat om de twee belangrijkste tradities van kwalitatieve en kwantitatieve data opportunistisch te combineren, creatief en nieuw gebruik van data-analysetools te introduceren, en op dezelfde dataset naadloos te wisselen van soort vraag. Bij dit creatief gebruik van data worden de verschillende tools en technieken voor het benutten van data toegeëigend op een manier die uniek is voor het ontwerpproces. Visualisaties spelen de rol van prototypes en 'boundary objects' in het ontwerpproces. In een dergelijke rol worden visualisaties niet gemaakt om representaties te creëren omwille van het rapporteren van de data, maar om de menselijke cognitie te 'smeren'. In deze praktijk gericht op het ondersteunen van begrip worden visualisaties snel gemaakt om een bepaald fenomeen te bevatten via de dataset.

De verschillende bevindingen van het proefschrift komen samen in het Design Inquiry Through Data raamwerk. Het raamwerk beschrijft hoe datawetenschapspraktijken kunnen worden geïntegreerd in een generiek ontwerponderzoeksproces dat daarmee van een "onbekende" staat naar een "bekende" staat gaat. Het raamwerk laat zien hoe een ontwerper kwantitatief data-denken met kwalitatief design-denken combineert in de drie conceptuele stappen van 'probleemdefinitie', 'verkennen' en 'interpreteren'. Het doorlopen van deze drie fasen gebeurt aan de hand van met elkaar verweven ontwerp- en data-activiteiten, waarbij niet alleen de data op een creatieve manier worden hergebruikt in het ontwerpproces, maar ook de data-activiteiten zelf op een met een ontwerpaanpak worden benaderd.



Design Inquiry Through Data raamwerk. Het raamwerk toont de synergie tussen datawetenschapspraktijken en ontwerppraktijken als verweven stappen van kwantitatief data-denken (kleur roze) en kwalitatief design-denken (kleur blauw). In de tussenstappen geeft de kleurenmix weer welk denken en gereedschap dominant is (buitenste kleur is dominant).

Het proefschrift biedt praktische, methodologische ondersteuning om datawetenschapspraktijken mee te nemen in ontwerponderzoek. Door datawetenschapspraktijken opnieuw vorm te geven maar nu vanuit een ontwerpperspectief, kunnen ontwerpers beter profiteren van bestaande datasets en dataverwerkings technieken zonder daarbij afhankelijk te zijn van de beperkte beschikbaarheid van datawetenschappers. Praktisch gezien draagt het proefschrift bij aan de ontwerpdiscipline in de vorm van een toevoeging aan methodologie, methode, ontwerptools en een raamwerk. Deze nieuwe ontwerptechnieken formaliseren een toekomstige datarijke ontwerppraktijk en bieden een veelzijdig perspectief op hoe data breed kunnen worden gebruikt voor ontwerponderzoek. Deze bijdragen worden gepresenteerd samen met rijke empirische inzichten over hoe beginnende ontwerpers ontwerponderzoek door middel van data gebruiken. Deze inzichten zijn tevens waardevol voor ontwerpdocenten en ontwerpers van datatools voor de niet-data-wetenschappers.

Acknowledgements

I am sitting in the Estonian countryside right now, trying to recollect all the support I have gotten and influential experiences I was part of that shaped this project as it turned out. Time to acknowledge the village that made this work happen. Thank you all for being part of the journey.

The first thanks go to my supervisory team.
Ingrid, I am indebted you gave me a blind chance many years ago to start in your team in Rotterdam. It has been what a ride since, full of consortium meetings, calls, travels, hackathons together. When there was space and time for research, you tirelessly hammered into my mind how to straighten my thinking and communicate my research so it's foolproof. You persistently challenged me to do my best and not to settle. You kept me to drill down to my core interests and beliefs that resulted in a dissertation that is 100% authentic me. Thank you.

Gerd, you started on this project and the organization a few months later than me, and I had absolutely no idea what to expect from you. At first you were a great sparring partner to figure out how data and design could overlap. You tried to steer me away from overcomplicating my research, nevertheless played along to help me pursue my chosen path. I have appreciated your listening, calm paraphrasing of feedback until I got it, and constructive thinking-along to better all my work. And last, I have appreciated your mentorship to set me up in academia and to have a realistic sense of such career. Thank you.

My committee members, Annika, Alessandro, Maaike, Marcus, and Paul. Thank you for assessing my dissertation, and for the thoughtful comments. I hope you and I are not done with each other, and there are collaborations to come. Alessandro, you knew this project for years. Some of your comments from the 12 months meeting has sticked with me since. Annika, you and I have had some history already, hope to keep working together.

The Open4Citizens consortium. Nicola, Amalia, you are giving me shelter for my next chapter in life, that is something to look forward. Louise, Anne Sofie, that O4C time together was hilarious and special, thank you. Anne, Francesco, Grazia, Ingrid, Janice, Mehdi, Kjartan, Marc, Petter, Tomas, Tomasz, thank you for the great collaboration, team spirit, and fantastic consortium meetings. It was fun. The Rotterdam pilot, Jelle, Robbert, Geertje, Mercedes, Anne, Jochem, Gerard and the rest. It was great collaboration to get those hackathons running.

The graduation students, Mercedes and Qianqian, you don't know, but your projects were instrumental to get my thinking on the PhD further. Mercedes, you turned a very hard topic something tangible. Qianqian, I took on mentoring you at such a late stage of my PhD, that I really shouldn't have had. No regrets though, your work was very inspirational and energizing to explore the boundaries of my research.

Dear Studiolab people, you were my second home for many years. Aadjan, Abhi, Alev, Anna, Berit, Bob, Chen, Eefje, Evert, Fenne, Froukje, Haian, Holly, Ianus, Iohanna, Jay, Jantien, Jasper, Lorenzo, Lyè, Maarten, Mafalda, Makiko, Marco, Maria Luce, Marian, Marie Rose, Martin, Natalia, Patrizia, Pelin, Peter v W., Ricardo, Richard, Roy, Sacha, Sam, Stella, Susanne, Taylor, Tessa, Tomasz, Xueliang, and the ones I unfortunately may forgot. The lunches, the coffees, the pantry encounters, the department events, the weird discussions from parenting to toilets or world politics, it was unique.
Anna P., who would have known many years ago that we will be desk neighbors and talk a lot about our kids? Boud, you and I are dangerously alike. I cherish the DRS bus trip, the long and deep talks about life instead of writing, your daring humor and sharing fatherhood experiences together. Holly, one month into my PhD, we had a transformative conference time in Eindhoven, and you have been an outstanding expat friend in so many ways since. Maria Luce, you were a dangerous office mate for a short while, but we had so much to discuss struggling with expat life, academia, and parenthood :) Ianus, you whirled me up in the IDE Academy machine very early, and it had a large impact on my PhD; for better or worse, thank you. PJ, knowing that one can remain a nerd even high in the ranks

comforted me to stay true to my strange interests. Elisa, thank you for the senior advice while traveling back from Zurich and grabbing coffee with me; they brought me clarity. Tomasz, despite the rough start, I am glad to had you around as a fellow bratanki, always generously willing to help. Aadjan, Martin, I wish we hacked more, maybe next time. Berit, I am very glad I gained a Nordic friend, see you in Copenhagen :) Peter v W., there *is* an end to this! A special mention goes to my PhD cohort, Bob, Boud, Chen, Frithjof, Holly, Lyè, Marian, Patrizia, Tessa, Xueliang, I am glad for the camaraderie how we have been helping each other out, providing safe space for venting, and generally doing this together. The Double Blind Reviewers, those jams kept me smiling and were absolutely worth the scheduling hell. Amanda, Charleyne, Daphne, Denise, Joost, thanks for running the department support machine over the years.

The larger IDE and TU Delft network. Ana, you were one of my favorite colleagues ever, and a role model in how to be a compassionate human. Erica, you were a great boss some lifetime ago, and such a generous friend for my family. Mieke, our collaboration was brief, but I am very glad it happened; after all our meetings, your story and passions left me energized. Milene, the few times we talked about research instead of kids had a large impact on this PhD, thanks for giving me the vocabulary to frame creativity in this design and data context. Jacky, wonderful chats in your office door, your no-nonsense attitude was always refreshing. Jeff, you helped enormously on navigating the ethical dimensions of data collection, and your easygoing and honest attitude made some stressful times easier to cope with. Jotte, our chats and laughter kept me sane.

Two exceptional Dutchmen need some extra words. Jasper, our talks about design and data over the years are all dispersed in this thesis, and your perspectives have largely influenced my work. Your hunger for knowledge, get-go attitude and teaching style are all a reminder what kind of professional I am aspiring to be. Kars, our collaboration on machine learning workshops were such a great way to get to know you, and then had you as a colleague. Thanks for translating the summary of this thesis, and for the superb chats over the years.

The years at Chalmers were among the most stimulating times and left a mark on me as a human and a designer. Sus, and the rest of the staff, thanks for the wonderful learning environment. Niels, Robert, Kaspar, and the rest of the batch – I miss doing weird projects and late night soldering.

Jason, you are the best neighbor one can wish for. I wish academia was more agile :) Friends in the lowlands, Marius, Inez, Laura, Gert, Mik, Wieke, Darina, Lucien, Cristina, Anca, thanks for the visits, the parties, and all the joy to make our Netherlands time more complete. Friends around Europe, vTeele, Paal, Lenno, Telku, Kert, Filip, Borys, Chris, Mário, Morten, Monica, Piotr, Marcos, Pika, Olga, Bea, and the accidentally forgotten ones. What a history we share, and looking forward to make more stories happen :) Please come and visit asap!

Drága jó Andris, András, Catch, Gergő, Pogi, Robi, Sali, Tamó papik, tűkön ülve várom a következő horgászatokat amikor szétröhögjük magunkat. Bookmark ide. Ármin, Marci, Pó, Józsi, Norbi, és a többi jászberényi haver – bár szétszéledtünk, de a több évtizednyi barátság mély nyomokat hagyott rajtam, amit a külföldi távollét csak még szentimentálisabbá tett az évek során :)

Anya, Apa, Gábor, köszönöm a rengeteg önzetlen segítséget, még akkor is amikor fogalmatok sem volt hogy tulajdonképpen mivel foglalkozom. Megannyi támogatást kaptam a PhD előtt és alatt tőletek a gyerekekkel meg mindennel, nehéz leírni mennyire hálás vagyok. Gréti, csak így tovább a testvéremet kordában tartani :)

Karola, Andor, you were part of this PhD from the beginning to the end. Your smiles, laughter, and unconditional love is a constant reminder to focus on what is truly important in life. Anet, thank you for making the space for me to finish this work, and to make me a better human every day. You make me complete. The best is yet to come.

Somewhere near Võru, Estonia, July 2020
Péter Kun

About the author

Péter Kun was born in the small Hungarian city of Jászberény in 1986. Péter joined Delft University of Technology in January 2016 to pursue a PhD at the Faculty of Industrial Design Engineering. In September 2020 Péter is joining the Service Design Lab at Aalborg University in Copenhagen as a post-doc.

Prior to the PhD, Péter received a masters degree in Interaction Design and Technologies from Chalmers University of Technology in Gothenburg, Sweden. Péter also received a bachelor's degree in Industrial Engineering and Management at Budapest University of Technology and Economics, specializing in Product Management and Human-Computer Interaction. During his student years, Péter was active in an international student organization, and among other accomplishments organized a pioneering summer course on physical computing with Massimo Banzi (Arduino) in 2007, lead the organization's internal education group for the year 2009/2010, founded the Trainer's Forum collaboration platform for youth educators, and worked as a trainer-facilitators with youth NGOs and the Graduate School of Delft University of Technology.

Prior joining Delft University of Technology, Péter was a lecturer-researcher in interaction dseign at Rotterdam University of Applied Sciences and a researcher at the Creating 010 knowledge center.

List of publications

PEER-REVIEWED PUBLICATIONS

Kun, P., Mulder, I., & Kortuem, G. (Under review). Developing a Design Inquiry Method for Data Exploration. *Interaction Design and Architecture*.

Kun, P., Mulder, I., De Götzen, A., & Kortuem, G. (2019). Creative Data Work in the Design Process. In *Proceedings of the 2019 ACM SIGCHI Conference on Creativity and Cognition.* ACM. : https://doi.org/10.1145/3325480.3325500

Kun, P., Mulder, I., & Kortuem, G. (2018). Data Exploration for Generative Design Research. In *Proceedings of Design Research Society 2018* (pp.1342-1356). London: Design Research Society. https://doi.org/10.21606/drs.2018.565 **(Best Paper Award)**

De Götzen, A., **Kun, P.**, Simeone, L., & Morelli, N. (2018). Making Sense of Data in a Service Design Education. In *Proceedings of the ServDes.2018 Conference* (pp. 177-187). Linköping University Electronic Press.

Kun, P., Mulder, I., & Kortuem, G. (2018). Design Enquiry Through Data: Appropriating a Data Science Workflow for the Design Process. In *Proceedings of the 32nd International BCS Human Computer Interaction Conference (HCI 2018)*. BCS Learning and Development Ltd. https://doi.org/10.14236/ ewic/HCI2018.32

Kun, P., & Mulder, I. (2015). From Quantified Self to Quantified Neighbourhood. In *Proceedings of PIN-C 2015: Reframing Design, the 4th Participatory Innovation Conference* (pp. 138-143). The Hague University of Applied Sciences.

BOOK CHAPTERS

Mulder, I., & **Kun, P.** (2019). Hacking, Making, and Prototyping for Social Change. In M. de Lange, & M. de Waal (Eds.), *The Hackable City: Digital Media and Collaborative City-Making in the Network Society* (pp. 225-238). Singapore: Springer. https://doi.org/10.1007/978-981-13-2694-3_12

JURIED AND LIGHTLY-REVIEWED PUBLICATIONS

Kun, P. (2019) Tools for Capacity Building. In Workshop on Exploring, Defining, & Advancing Community-Driven Design for Social Impact at ACM Designing Interactive Systems Conference, 23 Jun 2019, San Diego, United States.

Kun, P. (2019) Empowering Diffuse Designers Using Data In Design. In Workshop on Civic Data Literacies for Bottom-up Data Innovation at 9th Communities and Technologies Conference, 3 Jun 2019, Vienna, Austria.

Wolff, A., **Kun, P.**, Knutas, A., Palacin, V., Seidelin, C., & Mulder, I. (2019) Civic Data Literacies for Bottom-up Data Innovation. In Extended Abstracts of the 9th Communities and Technologies Conference, 3-7 June 2019, Vienna, Austria.

Kun, P. (2019) Creative Data Work – Designers Appropriating Data Work for Design Inquiry. In Workshop on Human-Centered Study of Data Science Work Practice at ACM SIGCHI Conference, May 2019, Glasgow, United Kingdom.

Kun, P. (2018) End-user Data Tool Assemblages for Design Inquiry. In Workshop on Digital Tools in Collaborative Creative Work at NordiCHI'18, 29 Sep, 2018, Oslo, Norway.

Concilio, G., Molinari, F., Morelli, N., Torntoft, L. K., De Götzen, A., Mulder, I., Jaskiewicz, T., **Kun, P.**, Pedersen, J., Ammentorp, P., Aguilar, M., Edman, T. & Sorensen, A. S. (2017). OpenDataLabs: New Infrastructures for Open Data Commons. Poster session presented at ECDG 2017: 17th European Conference on Digital Government, Lisbon, Portugal.

Kun, P., & Mulder, I. (2016) Prototyping for Citizen Engagement. Workshop at the Design and the City Conference, 22 Apr, 2016, Amsterdam, The Netherlands.

Fohlin, R., **Kun, P.**, Rebolo, L., Jansson, M., & Swinkels, N. (2013). Music Co-Creation in Public Spaces via Interactive Coffee Tables (poster). In 9th Student Interaction Design Research conference (SIDeR).