The Secret Life of Data – Uncovering the Diverse, Lived, and More-Than-Human Nature of Personal Data

Rosa van Koningsbruggen
Bauhaus-Universität Weimar
Germany
rosa.donna.van.koningsbruggen@uni-weimar.de

Bart Hengeveld
Eindhoven University of Technology
Netherlands
b.j.hengeveld@tue.nl

Eva Hornecker
Bauhaus-Universität Weimar
Germany
eva.hornecker@uni-weimar.de

ABSTRACT

As the datafication of our personal lives increases, researchers have started to critique what we consider as data. Where data are often seen from a reductionist perspective –as neutral numbers and graphs– theory suggests that data are messy, subjective, and pertaining to more than the human alone. They are experienced and lived with, something which is often not accounted for in personal data. To explore what people themselves consider data, sixteen interlocutors participated in a cultural probe study where they visually documented what they considered data in their everyday lives. Our analysis indicates that data can be disciplinary, social, and extend beyond ourselves, incorporating more-than-human aspects. Data are often regarded for entertainment, work, and wellbeing purposes. As interlocutors actively engaged and lived with data, we end the pictorial with a working definition of personal data, which acknowledges people's active roles in their data creation.

Authors Keywords

cultural probe, physical data

CSS Concepts

• Human-centered computing~Human computer interaction (HCI)~Empirical studies in HCI; ~HCI theory, concepts and models

INTRODUCTION

In attempts to better understand ourselves and the world around us, data are ever more present. New technologies, such as wearable and smart home devices, have enabled us to quantify and data-fy aspects of our lives which previously did not explicitly exist numerically. This process of datafication has resulted in an increase of available personal data [32], both voluntarily created (e.g., counting your steps to get an idea of your activity levels) or as side-effect of using a service (e.g., tracking cookies after visiting a website). However, whereas data's context has changed and expanded –from the scientific to the personal world– our understanding of what personal data are has not co-evolved.

Previously defined as *"any piece of information that can identify or be identifiable to an individual"* [35], personal data stem from medical settings where they are used to get a better understanding of a patient's life. Despite no longer being limited to this context, personal data are still surrounded by an aura of objectivity and efficiency [7,21]. These views are upheld by what we consider to be personal data, and how we track and represent these [7,11]; *"Impartial machines"* in the form of tracking technologies and sensors [21] reduce us to numbers, statistics, and logical patterns, that supposedly show us the path towards better health, productivity, and happiness [13,21]. These data are then represented in visualisations, which often look minimal, sleek, and scientific [23]. The combination of impartial machines and clean-looking visualisations contributes to personal data being experienced as objective and truthful [7,9,27–29].

These perceptions of data have been criticised by researchers, who argue that there are no neutral data [2,19]. It is only through human actions, that data become data. They are small and local, rather than big and universal [27,42]. They are plural, instead of singular [7,27] and come in assemblages—both with us, data, and other nonhuman actors [8,10,29]. These relations to and within the world are what define data beyond their numeric value [20,26]. Therefore, Rooksby et al. introduced the notion of *"lived informatics"* which covers the different ways in which people use, track, and live with data [37]. Going further, recent work suggests that personal data might even have 'post-qualitative' aspects [25,31], meaning that personal data can be *"any kind of matter, both organic and inorganic"* [30]. This can be seen in the field of data physicalisation, which shows that data can be physical and represented beyond the visual modality. For this, Offenhuber has created a framework of physicalisation's various ontological and epistemological purposes [34]: those that stay close to quantitative, digital data and visualisation principles (such as 3D bar graphs [41]), and those that do not use traditional data sets and represent data through traces (e.g., a seismograph) [33,34].

Considering the current mismatch between how research discusses personal data and how they are commonly conceptualised in popular culture (as numerical and objective, such as used in the Quantified
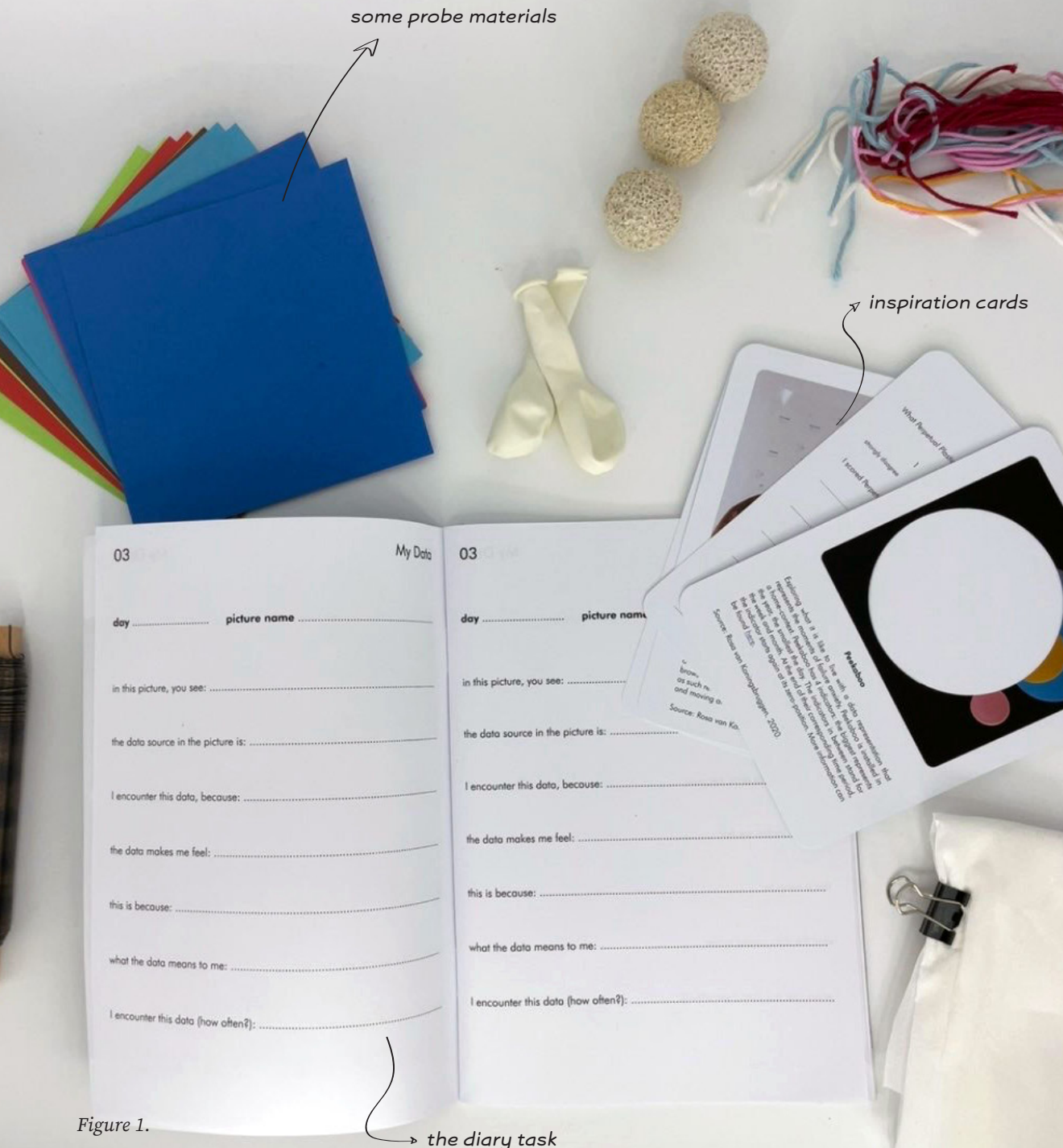
some probe materials

inspiration cards

03   My Data   03

day ....   picture name ....

in this picture, you see: ....

the data source in the picture is: ....

I encounter this data, because: ....

the data makes me feel: ....

this is because: ....

what the data means to me: ....

I encounter this data (how often?): ....

Figure 1.

the diary task

Self (QS) movement), and HCI and Computer Science research, it is time to reconceptualise what personal data are [42], as the current reductionist view does not account for the variety of data [2,34], nor how people live, interact with, and make sense of them—which tends to be qualitative rather than quantitative [37]. To contribute to this reconceptualization, our work introduces a cultural probe study with sixteen interlocutors, in which we explored what people consider data in their lives. We specifically focus on one of the tasks, where interlocutors took pictures of the data they encountered. In line with Feinberg's critique [15], our work shows the variety of data and tracking processes: ranging from mundane and tedious (e.g., reminders to clean), to creative data proxies (e.g., organisms which helped people make sense of time). Moreover, we introduce three types of personal data: (1) *Disciplinary Data*, that indicate or remind what you can(not) do; (2) *Social Data*, that apply to other people, organisms, and things, both close and distant to us; and (3) *More-Than-Human Data*, which describe our relation to non-human entities which interlocutors labelled as personal data. These data were used for three types of purposes: (1) Work and Utility; (2) Entertainment; and (3) Wellbeing.

With these findings we show that current definitions of personal data are too limited, as they do not include the user's active role in creating data nor that all data can become valuable and personal, as they are part of larger assemblages which are meaningful to us [16]. Therefore, we introduce a working definition of personal data. The contribution of our work is two-folded: (1) we show what people themselves consider personal data, and (2) we provide visual examples of the richness of personal data and human data use. Through these examples and our working definition, we aim to redefine how designers and HCI researchers view personal data, and inspire them to explore and create beyond reductionist understandings of personal data.

METHODOLOGY

This work is based on a cultural probe [18] completed by sixteen interlocutors. The aim of this study was to

explore and visualise the data people encounter in their everyday lives, and find out what they consider as personal data. To do so, we reached out in our social networks, for a mix of people who work with data (either professionally or for fun) and people who do not work with data. Furthermore, we looked for a balance between creative occupations (e.g., artists and designers), and scientific or technical occupations (e.g., data scientists), to have interlocutors who either felt comfortable working with data or were open to explore what they are. We speculated that this diversity in backgrounds would result in interesting overlaps and differences in what is considered and understood as data. Interlocutors lived in the Netherlands, Germany, and Denmark. We use the word 'interlocutor' (opposed to 'participant') to reflect the (creative) work conducted by the people who took part in our study and how their views have influenced our own.

After the informed consent procedure, each person received the probe either via post or personally. The probe consisted of a box filled with basic craft supplies (such as tape, origami paper, clay, etc.), a diary with assignments and explanation of the probe, and inspiration cards containing pictures of data and data representations (see Figure 1). The diary and inspiration cards are shared as part of the article's supplemental materials. Through these materials, we intended to sensitise interlocutors to think 'outside the box' when considering what data could be. The probe task had two phases, with an interview in-between, to discuss interlocutors' experiences, pictures, and introduce the new phase.

We here focus on one of the assignments of the first phase, where interlocutors worked with a diary to reflect on what data are, where and when they encounter them, and took pictures of the data. We here specifically focus on the latter. For this assignment, interlocutors were asked to: *"make pictures of all the data you encounter in your everyday life. This can be anything you see or label as data"*, explain the images in their diary (Figure 1), and upload the pictures to a personal folder hosted on the
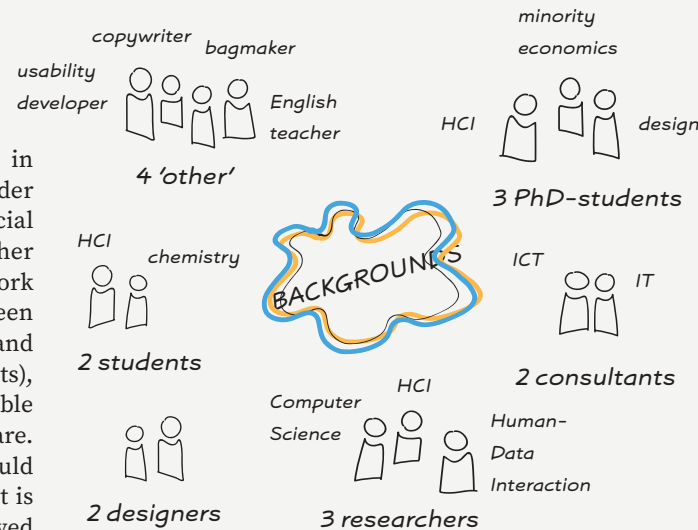
university's cloud service. Interlocutors had five days for this, which did not need to be consecutive; to fit interlocutors' lives and schedules, they could decide on which days they would photograph data, as long as it added up to five days. The assignment resulted in 225 pictures. The minimum number of pictures from an interlocutor was eight pictures, the maximum twenty-seven pictures, with an average of fourteen pictures.

Data analysis

The images were analysed by two people (the first author and a HCI research colleague) using Reflexive Thematic Analysis (RTA) [4]. Following an inductive approach, the researchers first individually looked for patterns and contrasts within the data of an interlocutor. Once the researchers had a good grasp of the data, they looked for patterns and contrasts throughout all interlocutors' pictures, noting down things which seemed relevant to them in initial open coding, such as whether the data was tangible or digital, a recurring or singular event, and where people encounter the data. Once both researchers had gone through all the pictures, they discussed their observations and initial coding in two 3-hour-long sessions. The researchers aligned their codes and created initial themes. These themes were further discussed after a reflection period of a week. Now, the researchers finalised their coding and themes, and arranged them spatially in a thematic map, highlighting relations between themes.

CONSTRUCTED INSIGHTS

Next, we show and discuss three types and three purposes of personal data. While the initial task required interlocutors to picture the data encountered in their lives, insights from diaries and interviews revealed that the data that participants noticed, were primarily personal data—data either stood out because they were personal, or by noticing them, the data became personally relevant. Types and purposes of personal data cannot be disentangled from each other. Often pictures of data sources could be assigned to multiple categories, showing the plurality of the data source and its aim. For clarity purposes, we show pictures under one type or purpose. However, we acknowledge that the following insights are interconnected with each other.

Moreover, as stated by Segal (2021), data are simplifications of the rich reality around us and do not have any meaning on their own [39]. It is only by giving shape to data and representing them (e.g., in a visualisation), that we can start to make sense of the data and discover their meaning. This can also be seen in the next sections, where interlocutors captured data representations, autographic visualisations [33], and data proxies, opposed to 'raw' data. Similarly, participants photographed data sources which could be considered information rather than data. As the distinction between data and information is blurry [5,40] and personal –what is information for one is data for someone else [7]– we adhere to our interlocutor's labelling of the image as data.

Each section starts with an introductory page, followed by examples and detailed explanation. Pictures are referenced to the interlocutor who took them (e.g., I3) and discussed in clockwise direction or from left-to-right. Some images are annotated with the interlocutors' diary explanations or notes from the researchers. Interlocutor explanations can be recognised through the indicator after the quote.

Type 1: Disciplinary Data

The largest category of data types (70 pictures) are *'disciplinary data'*. Disciplinary data refer to a group of data which tell or are perceived to tell what people should (not) do. This can be literal –such as payment letters with a due date (I1) and calendar entries (I8)– or more abstract, such as reminders of what to do (or not do), e.g., I7's picture of a water bottle. Reminders often concerned *sports* or *cleaning*. Furthermore, interlocutors took pictures of data which determined whether they could do something. This applied to checking the weather forecast or the weather outside, to see whether they could go for a run or which coat to wear, thermostats, to check whether it was really as cold as they perceived it to be and could heat more, and COVID-19 rapid tests, to validate whether they could go outside. Sometimes, disciplinary data have social aspects, such as a pet indicating the end of a working day, what music to listen to, or a reminder to take care of an artefact.

location is essential
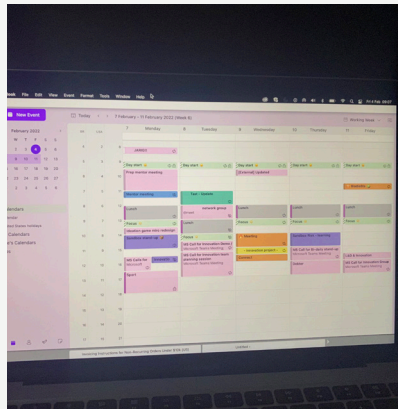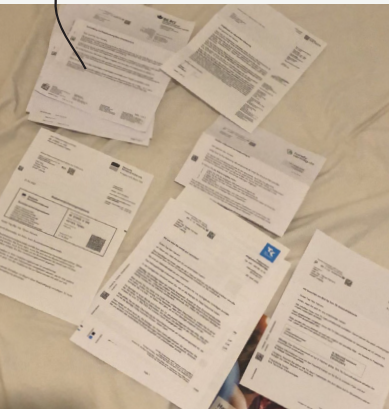
Left: picture of letters on the bed by I1. Every time I1 receives a letter which mentions important information, such as payments that need to happen or action items, they place them on the bed in the guest room—indicating that they require action. Right: Picture of I8's calendar stating what they need to do throughout the week.

I7's picture of a water bottle, which reminds them to drink enough throughout the day.

"I try to drink a lot of water" – I7

all data that influences our behaviour

1.

"Is it really freezing in here?" – I2

*digital encouragement*

Can I?
The pictures on the left show the most popular types of data within the category of *'Can I?'*. In total, we received five pictures of thermostats, as people check to verify whether the perceived temperature is correct or decide whether they were allowed to adjust the temperature—as was the case in this picture by I2. Next, is a negative COVID-19 test by I15. In total, we received seven pictures of such tests, which results influenced whether interlocutors were allowed to go outside. Finally, we have a picture of the outside weather by I8. Interlocutors either looked outside (two pictures) or at their weather app (four pictures) to get an idea of whether they could go outside for sports or what to wear.
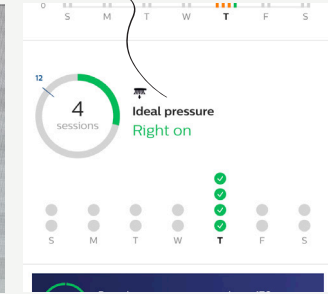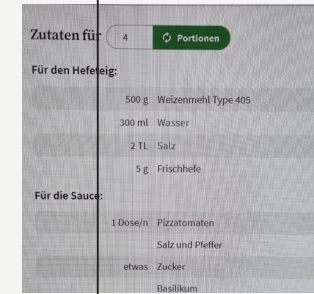
2.

Dictating Data
Data could 'dictate' interlocutors. For example, recipes (depicted left by I13) which mention how much of each ingredient has to be used. Dictating data often subtly hinted interlocutors to change or improve their behaviour, such as I2's screenshot of their toothbrush app which gives feedback regarding their technique.

kintsugi-inspired gold glue

4.

3.

Reminders to (not) do something
This category consists of varied data sources, such as calendar entries, speed indicators above the road, emails regarding deliveries, or the fuel gauge reminding you to fill up the tank. Depicted below are the most popular reminders: sports and cleaning. Physical (gym) equipment (as captured by I11) as well as health apps and booklets documenting your progress reminded interlocutors to work out. Similarly, messy situations, such as dust or children's toys scattered over the floor captured here by I13, and cleaning schedules reminded to clean.

physically in the way

Social Aspects
Although disciplinary data often concerned only the interlocutors themselves, they sometimes showed social aspects. For example, I4 photographed their dog waiting in front of the door to indicate that they should end their working day and a vase which is personally important to them, is staged so that the vase reminds them to repair and take care of it. Finally, I1 captured an excel sheet filled with concept albums. The excel sheet is a shared project with a friend and based on all the concept albums listed on Wikipedia. In randomised order, the friends listen and rate these albums, where the list dictates what they listen to and when.

Type 2: Social Data

Social data concern data sources which regard other people, society, and more-than-human aspects. They cover ways to connect with others, activities we do for and with our dear-ones, and helping or taking care of others. Thus, going beyond the idea of social data only constituting the sharing of quantitative data to establish and foster relations (cf. [38]). Social data play a role on micro (family and friends) and macro levels (retrieving information on a warzone or how your country is doing at the Olympic Games), and are not limited to living organisms, as they extend to things and artefacts.

Note, as mentioned under *"Constructed Insights"*, there are overlaps between the types of personal data. This is especially the case for social and disciplinary data. For example, I3's dog on the previous page could be seen as social data. For us, the difference between social data and social aspects is the main aim of the data. In case of the dog, it is to help the person stop working, thus being disciplinary rather than social data to us.

relation to own nationality   *relation to other nationalities*

Left: I10 captured the moment they were watching skating at the winter Olympics, both for the data portrayed about the skaters and to keep up with how their country is performing. Right: I15 included an image of how the war in Ukraine is progressing on a map. Every day they checked the map in the hope that Ukraine's position has improved.
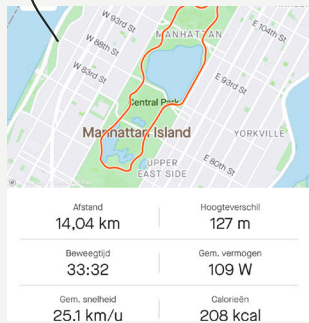
numerical data

is this data for the robin as well?

I3 took a picture of a robin eating seeds from their birdfeeder. I3 considered this data, as the image tells them that it is cold outside and indicates a friendly connection with the robin.

"it tells me that it is really cold outside, otherwise the robin isn't interested in the seeds I put outside" – I3

Quantified Self-data

"Sunrise, it is rare to see this."- I9

1.

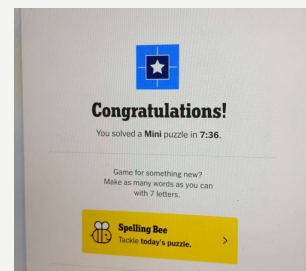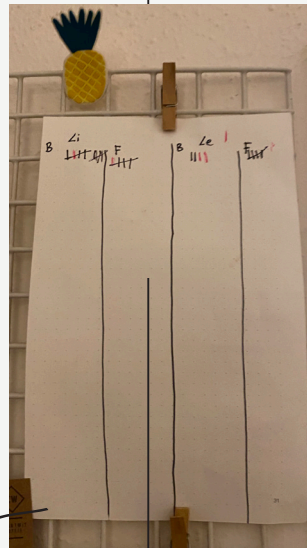*I9 shares 'rare' moments with family?*

plants are arranged so they can socialise

3.

Activities With and For Friends

Besides connecting with others, social data encompassed activities. For example, I8's picture of a *"burps and farts"*-list (left), which documents the number of burps and farts two friends release, to conclude who *"pollutes the most"*. Another activity is I7's picture of making filter coffee for their partner. Social activity data were often tangible and physical, but extended to the digital realm as well—as demonstrated by I15's picture of having finished a mini crossword, something they do at the start of each day with a friend to see who wins (the fastest person) that day.
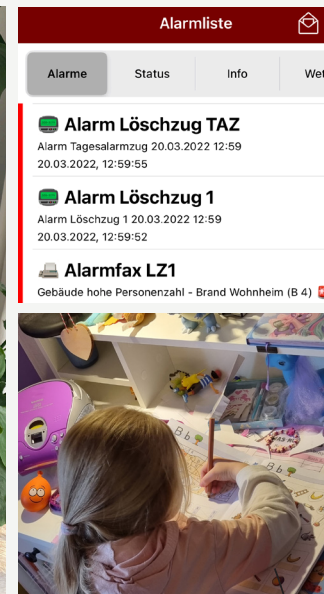
2.

Helping Others

Lastly, social data include instances of helping others. This could be on a societal scale, such I11's picture of the notifications regarding fires they receive as a voluntary firefighter (top-right), a personal scale, such as I13's picture of helping their child doing their homework, or relate to non-humans. The latter can be seen in I8's picture of taking care of their plants.

Connecting with Others and the Past

Social data are often used to connect with others or the past. To connect with people, digital communication –e.g., WhatsApp notifications– and sharing –e.g., LinkedIn posts– are popular. A platform that was often pictured was Strava, which allowed interlocutors to share their workout data (such as duration, and cycling or running route as shown in I5's picture on the top-left) with friends and respond to each other's workouts. Besides digital communication, pictures played an important role in fostering connections and reminiscing: I9 takes pictures of sunsets and views to share with their family back home, and I7 included two polaroid pictures of their friends, reminding them of their activities together. Finally, some objects have the power to remind of others, such as the pen captured by I4—a gift from their mother-in-law. When I4 uses the pen, they are reminded of her and how she was.

data = "How often I have to ask her when she is starting her homework."-I13

expressing love by making coffee. Thus, data can be love-able?

data can be funny!

Type 3: More-Than-Human Data

As the previous theme shows, interlocutors encountered personal data that went beyond themselves and that covered more-than-human perspectives (e.g., I4's pen). Here, we refer to Lupton's use of more-than-human data, where personal data are described as human—non-human data assemblages [30]. These assemblages show that personal data do not only concern the creator or the 'self' [36,44]. More-than-human data cover organic and inorganic data sources, which people use to make sense of their lives (aspects such as time and rituals) or living conditions (e.g., I7 who bought a pair of new shoes as a gift since things are going so well). Important to this theme is *"thing-power"* [3]: *"the strange ability of ordinary, man-made items to exceed their status as objects and to manifest traces of independence or aliveness, constituting the outside of our own experiences"* [4]. It refers to moments where humans and thinghood overlap. In case of the shoe example, it is the specific pair of shoes with their history, origin, time, and aesthetics that have a relation with the human, that results in a sentimental affect-loop. It is these specific shoe's thing-power that constitutes their relation to the human [3].
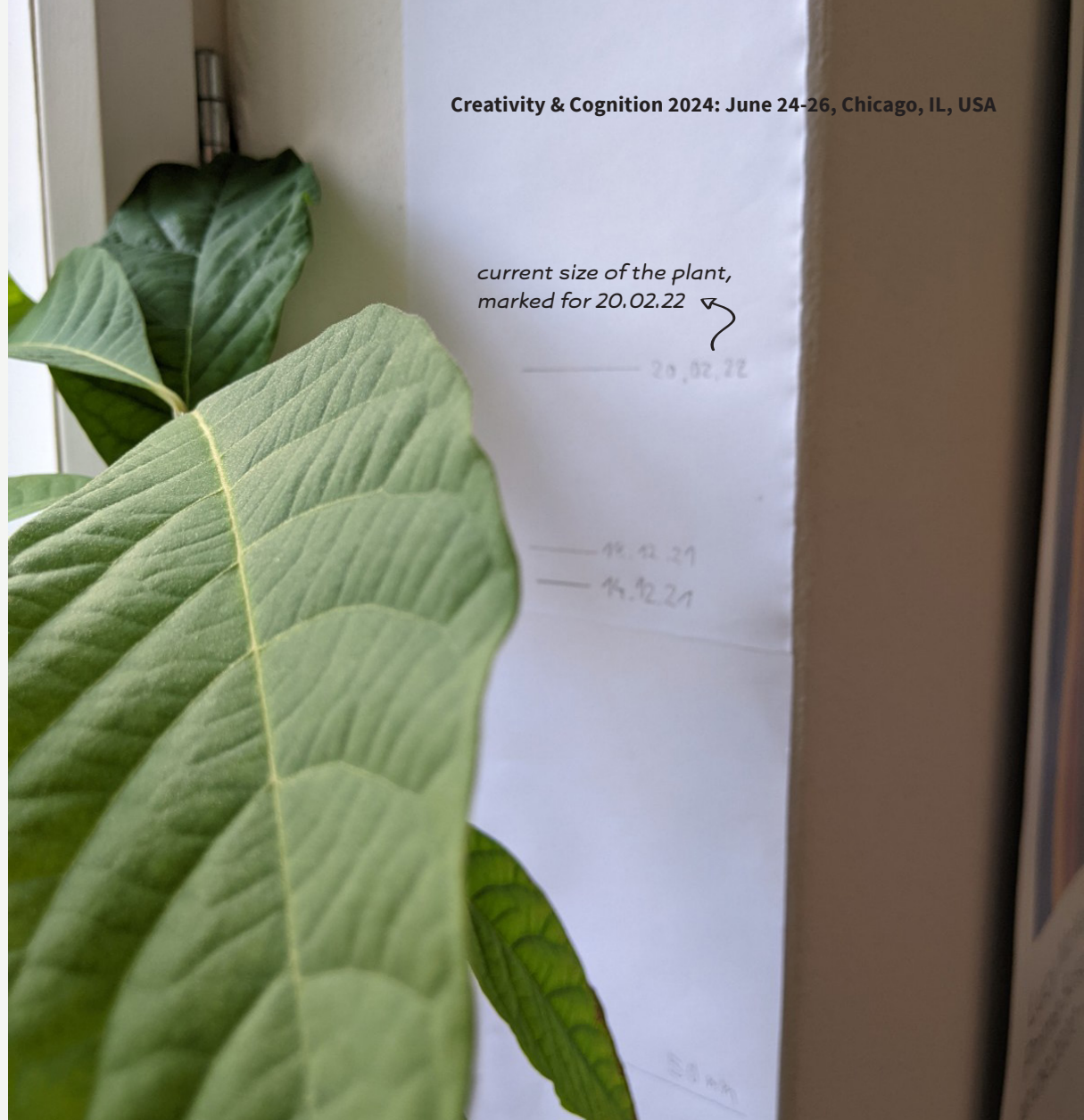
More-than-human-data was the second largest theme, consisting of 43 pictures. Within this category, interlocutors considered plants, animals, views, and artefacts as data sources.

current size of the plant, marked for 20.02.22

"The dog is pure softness and love" – I3

I15 took a picture of a self-drawn ruler behind one of their plants. Part of a friendly competition, I15 and one of their friends measure whose plant is growing faster. This offers both social data and data on how well the plant is doing.

Left: I3's picture of their dog taking a nap, representing unconditional love and softness to them. Right: I7's captured a pair of shoes they bought themselves as a personal gift; a memento of being proud of themselves.

"One tiny data source of giving myself nice gifts that cost a lot of money" – I7

1.

"The cat has its winterbody and lies on the blanket, thus it is winter" – I2

swans returning in spring

representation of I7's mental health

3.

Relation to and with Things

As shown for disciplinary and social data as well, interlocutors often photographed artefacts as a data source. In context of more-than-human-data, these data sources represent long-term relations – such as I3's water bottle, which broke after roughly twenty years (top-right)–, or allow one to express themselves, such as I15's drum machine which they use in their music practise or I14's interior decoration. Not only are the chair and Christmas lights a way to express themselves, they also influence I14's mood by creating a comfortable atmosphere. These can be seen as indirect representations of the self, thus saying something about us, becoming identifiable to us, and becoming personal data.

2.

"I have it for ±20 years. Always in my bag, all children took sips from it." – I3

Animals and Plants

Interlocutors considered living organisms as data, as they helped them deduce time on various scales. For example, I6 photographed their dog begging for a piece of cheese (top left)—a daily ritual which always happens at the same time. On a longer scale is I3's picture of two swans. To I3, this picture indicates the start of spring, as it shows the yearly ritual of the swans returning to start a new nest.
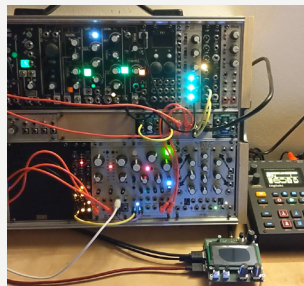
Besides being an indicator for time, instances of this data category were used for inspiration and understanding. I2 photographed their cat on a blanket looking very relaxed, which inspired them to unwind themselves, and I15's bulbs which begin to become visible above the ground. For I15, this is a special moment: when growing bulbs in water, you can see exactly what they are doing, whereas underneath the earth do not have this feedback and data. When the stems are starting to show, the data they use to make sense of the plant's wellbeing is there again.

personal expression and allow you to set the mood

Views

Views were considered data for three reasons. For I5 and I9, views facilitate social connection. I5 takes the same picture every day during the train ride to their work, where they cross an area with parked cranes (top-right). This view is then sent as an inside joke to one of their friends. Less frequent, but also fostering connection, is I9's picture of their bike in front of a mountainous background during their holidays. Not only sparked this view the desire to connect with others, I9 also took it for reminiscing on their holiday. Finally, I7's picture of trees in a park was considered data, as this particular walk through their city confirmed the feeling that they did not like where they are currently living: the city felt unwelcoming and distant. This view embodied those aspects according to I7, thus becoming a data point to them.

Purposes of Personal Data

The types of data were used for various purposes. In our analysis, we constructed three clusters of common purposes. Below, we introduce these from largest to smallest cluster, starting left, then in clockwise direction.

body as a data carrier

LARGEST CLUSTER
pictures of media

ENTERTAINMENT
As entertainment data, pictures of Netflix (as depicted here by I5), online games, making and listening to music, and learning something new by reading books and engaging with podcasts were common. Entertainment data could be a solitary (e.g., reading a book) or a social activity (e.g., multiplayer games). Moreover, we encountered a rare type of entertainment data, which we label *"funny data"* (three pictures). Funny data describe data which are meant to facilitate or trigger laughter. On the next page, we show examples.
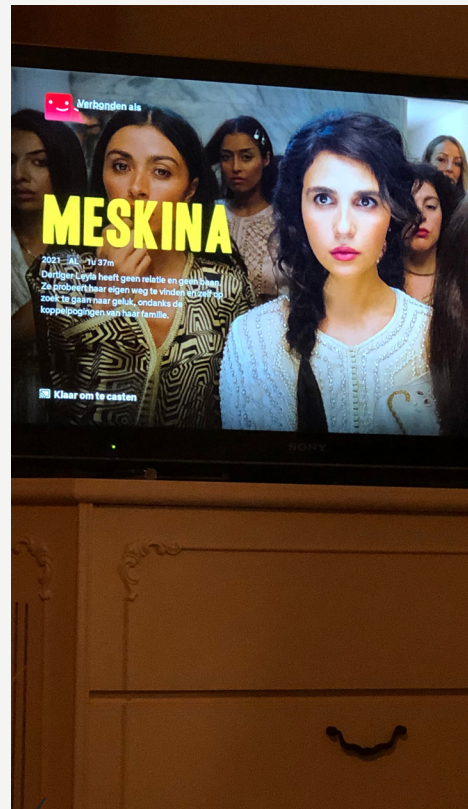
"The data is the Netflix recommendations." – I5

body as data (representation)
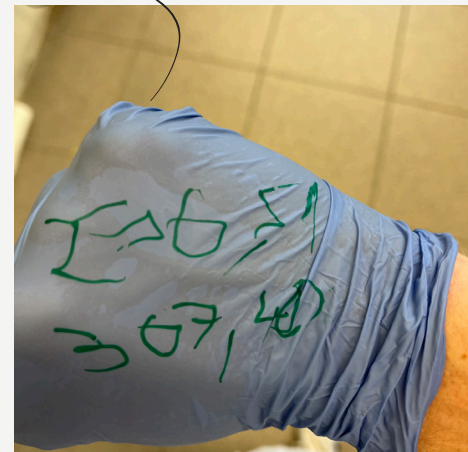
WORK
Twenty-eight pictures were labelled as related to work or for work purposes. Depicted here is I11 wearing a glove in the laboratory. The glove has written data on it, which I11 needed for an experiment—resulting in a wearable data representation. However, more commonly, the data were either encountered online (e.g., online databases) or physical (e.g., books). Noteworthy to us, and as we demonstrate on the next page, for each instance of digital work data, a tangible version existed as well in our data set.

WELLBEING
Lastly, nineteen pictures were categorised as wellbeing data. Although eight pictures show screenshots from wellbeing apps on a phone (e.g. the Health app), this was diversified with physical examples of food (I9), workout equipment (e.g., I11's picture on page 5), and physical reminders for a healthy lifestyle (e.g., I7's water bottle on page 4). Mental health was included as well: I7's picture of their city view on page 9 and as depicted here by I8. When stressed, I8 bites off the ends of their hair. By checking their tips, I8 can see how they have been mentally doing the last weeks.
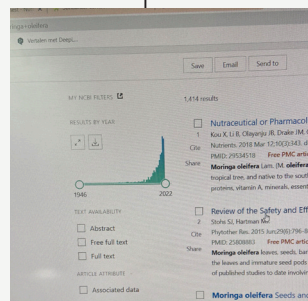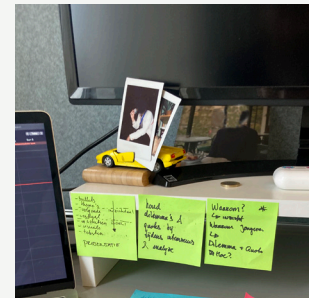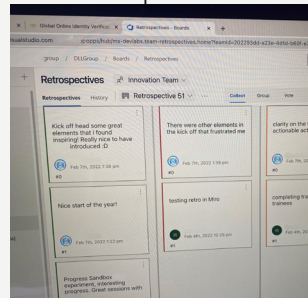
1.

"The data makes me feel happy/silly"– I5

Entertainment — Funny Data

Three instances of 'funny data' were the cranes as an inside joke captured by I5 (top-right), the 'burps-and-farts list' by I8, and the not-yet discussed posters on the wall in I11's home. When I11 enters the hallway to their apartment, they walk past a wall filled with the honorary certificates ('Ehrenurkunde') in floral decoration of their landlords. Every time they walk past it, they sympathetically laugh internally about the pride and enjoyment their landlords get from their work.

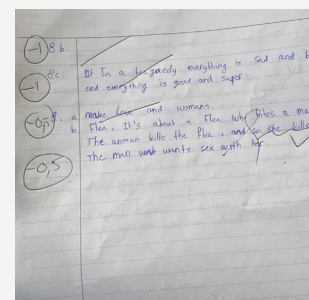*all thing you pass by. Is temporality essential?*

2.

Work

For all subcategories labelled as 'work data' in our data set, there was always a digital and a tangible example. No subcategory was restricted to only one of these categories—aligning with Offenhuber's framework [34]. For example, in the top row, we see the online reflection space captured by I8 and the tangible reflection space created by I7 on their desk, using post-its. The middle row shows I3 looking for articles online, whereas I10 photographed an printed article, and the bottom row shows I15 grading student's work in an Excel sheet, whereas I2 opts for marking physically on paper.
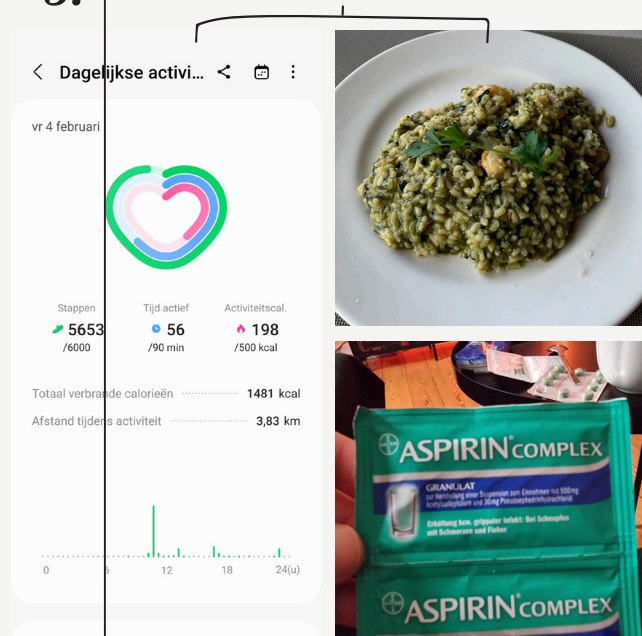

3.

both numeric and qualitative wellbeing data (did I like the food?)

Wellbeing

Included here are a screenshot of a wellbeing-app by I10 (left), a plate of food by I9, and a picture of aspirins taken by I12. The aspirins offer numerical data (the nutrient values listed on the package), a list of ingredients, and indicated that I12 was not feeling well. Thus, this image embodies multiple types of wellbeing data.

DISCUSSION

To get insights in what people themselves understand as data, we conducted a cultural probe study. From our findings we learned that most of the data that we notice are personal data, as they give new insights into our bodies, habits, environment, and social relations, which in turn contribute to a renewed understanding of ourselves [30]. In our study, interlocutors encountered three types of personal data: disciplinary, social, and more-than-human, and captured three use purposes (entertainment, work, and wellbeing). Although the disciplinary data and purposes align with what personal data are currently used for (e.g., *"entertainment medicine"* [17]), the social and more-than-human data show that our interlocutors have a far more diverse and complex understanding of what personal data are, where we encounter them, and what they mean to us. Furthermore, the more-than-human and social data show that defining personal data as anything that can identify someone or be identifiable to them, is too limited. For example, 'views' (page 9) seem rather general: multiple people will have access to the same views. Nonetheless, these data sources were meaningful and emotional to our interlocutors—stressing the active role they play in the meaning-making of the data.

Looking at the purposes of data and zooming in on the (small, but nonetheless important) category of funny

I15 took a picture of touching the coffee machine to feel whether it is hot, as the indicator is broken.

data, we also see a stark contrast to the illusions that data are sterile and clean [21]. In reality, data can inspire laughter, foster connections, and messily live with us in our daily lives.

A New Definition of Personal Data

Considering these perspectives, data are more than the *"agglomerations of small, discrete signals, represented as 0s and 1s in computer memory"* [27]. They extend beyond numerical and statistical perspectives, and go into what Offenhuber c alls ontological perspectives where the physical thing itself is considered data [34]. Furthermore, our interlocutors show that personal data blur the boundaries between personal/other and human/non-human [30], with *"thing-power"* and *"animism"* (the attribution of life or human characteristics to phenomena that are considered non-living or nonhuman, such as a water bottle or shoes [4]) contributing to data sources beyond ourselves being considered as personal. Lastly, our interlocutors show that personal data are both actively generated (sometimes in creative ways, such as measuring your plant) or implicitly there. For the latter, these data become personal through active perception, resonating Drucker's claim that all data are *"capta"* that are actively generated, rather than given [11].

As this pictorial shows a stark contrast to common definitions of personal data, we argue that it is time to rethink how we define them. Popular definitions, such as that of the GDPR: *"any information relating to an identified or identifiable natural person"* [14], do not focus on users who actively create and engage with data, nor do they acknowledge aspects which are not directly relatable to a person, but which are relevant to them on a personal level. Therefore, we propose the following working definition: *"personal data are any piece of information that are actively created by the user, which they can identify as stemming from themselves or their environment, and which are meaningful to them, or which can be used to identify the user"*. To further protect our data, we hope that in the future, a definition like this will be used for legal purposes and prevent a wider range of personal data being extracted by companies for data capitalist purposes.

a soft, loving interaction?

At the same time, we see opportunities for design: as we have an advanced understanding of how to design for and represent data's numerical and epistemological side (cf. traditional data visualisations) [34], we hope that our working definition helps designers and HCI researchers to explore and design for data's ontological, post-qualitative side. For example, designing technologies that allow us to foster relations to personal objects through data (e.g., I3's water bottle could communicate the memories it holds to us), or for tracking your plants to make sense of time and your social relationships (see I15's picture on page 8). A promising direction are physical data representations, such as autographic visualisations [33] and data physicalisations [22], that inherently encompass an *"interpretivist view"* on data [1], and seem to better acknowledge the subjectivity and messy nature of data [1,16,24,34].

Limitations

Of course, the work presented here has its limitations. Our research took place in the Global North, with researchers and most interlocutors stemming from this part of the world. Therefore, the considerations of what personal data are, are limited to a small set of cultures. More research with more diverse groups of interlocutors is needed to get an understanding of what people consider personal data, and where our understandings differ and align. Moreover, although we asked interlocutors with diverse occupations and educational backgrounds –some that liked working with data and others who felt less confident approaching this topic–, we learned that all our interlocutors at some point had worked with data—thus having a certain level of data literacy [45]. Therefore, future studies should explore how our understanding of personal data differs with a more varied group of data literate interlocutors.

REFERENCES

[1] S. Sandra Bae, Clement Zheng, Mary Etta West, Ellen Yi-Luen Do, Samuel Huron, and Danielle Albers Szafir. 2022. Making Data Tangible: A Cross-disciplinary Design Space for Data Physicalization. CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA 1, 1. https://doi.org/10.1145/3491102.3501939

[2] Gabrielle Benabdallah, Maya A Kaneko, and Audrey Desjardins. 2023. A Notebook of Data Imaginaries. In Proceedings of the 2023 ACM Designing Interactive Systems Conference, 431–445. https://doi.org/10.1145/3563657.3596025

[3] Jane Bennett. 2004. The Force of Things: Steps toward an Ecology of Matter. Political Theory 32, 3: 347–372. Retrieved July 23, 2022 from http://www.jstor.org/stable/4148158

[4] Jane Bennett. 2009. Vibrant Matter: A Political Ecology of Things. Perspectives on Politics 9, 1: 118–120. https://doi.org/10.1017/S1537592710003464

[5] Jay H. Bernstein. 2011. The Data-Information-Knowledge-Wisdom Hierarchy and its Antithesis. NASKO 2, 1: 68. https://doi.org/10.7152/nasko.v2i1.12806

[6] Virginia Braun and Victoria Clarke. 2021. Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health 11, 4: 589–597. https://doi.org/10.1080/2159676X.2019.1628806

[7] Catherine D'Ignazio and Lauren F. Klein. 2020. Data Feminism. The MIT Press, New York. https://doi.org/10.7551/mitpress/11805.001.0001

[8] Manuel Delanda. 2016. Assemblage Theory. Edinburgh University Press. Retrieved March 21, 2022 from https://edinburghuniversitypress.com/book-assemblage-theory.html

[9] Audrey Desjardins and Heidi R. Biggs. 2021. Data Epics. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 1–17. https://doi.org/10.1145/3411764.3445241

[10] Paul Dourish. 2017. The Stuff of Bits - An Essay on the Materialities of Information. The MIT Press, Cambridge, MA. Retrieved June 25, 2021 from https://mitpress.mit.edu/books/stuff-bits

[11] Johanna Drucker. 2011. Humanities Approaches to Graphical Display. Digit. Humanit. Q. 5.

[12] Chris Elsden, Bettina Nissen, Andrew Garbett, David Chatting, David Kirk, and John Vines. 2016. Metadating: Exploring the romance and future of personal data. In Conference on Human Factors in Computing Systems - Proceedings, 685–698. https://doi.org/10.1145/2858036.2858173

[13] Chris Elsden, Mark Selby, Abigail Durrant, and David Kirk. 2016. Fitter, happier, more productive. Interactions 23, 5: 45–45. https://doi.org/10.1145/2975388

[14] European Parliament. 2016. Art. 4 GDPR: Definitions. General Data Protection Regulation (GDPR). Retrieved from https://gdpr.eu/article-4-definitions/

[15] Melanie Feinberg. 2017. A design perspective on data. In Conference on Human Factors in Computing Systems - Proceedings, 2952–2963. https://doi.org/10.1145/3025453.3025837

[16] Mikhaila Friske, Jordan Wirfs-Brock, and Laura Devendorf. 2020. Entangling the Roles of Maker and Interpreter in Interpersonal Data Narratives. In Proceedings of the 2020 ACM Designing Interactive Systems Conference, 297–310. https://doi.org/10.1145/3357236.3395442

[17] Katleen Gabriels and Tania Moerenhout. 2018. Exploring Entertainment Medicine and Professionalization of Self-Care: Interview Study Among Doctors on the Potential Effects of Digital Self-Tracking. Journal of Medical Internet Research 20, 1: e10. https://doi.org/10.2196/jmir.8040

[18] Bill Gaver, Tony Dunne, and Elena Pacenti. 1999. Design: Cultural probes. Interactions 6, 1: 21–29. https://doi.org/10.1145/291224.291235

[19] Lisa Gitelman and Virginia Jackson. 2013. "Raw Data" Is an Oxymoron. The MIT Press. https://doi.org/10.7551/mitpress/9302.001.0001

[20] Donna Jeanne Haraway. 2003. The companion species manifesto: dogs, people, and significant otherness. The University of Chicago Press, Illinois.

[21] Sun-ha Hong. 2020. Technologies of Speculation -The Limits of Knowledge in a Data-Driven Society. NYU Press, New York.

[22] Yvonne Jansen, Pierre Dragicevic, Petra Isenberg, Jason Alexander, Abhijit Karnik, Johan Kildal, Sriram Subramanian, and Kasper Hornbæk. 2015. Opportunities and Challenges for Data Physicalization. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 3227–3236. https://doi.org/10.1145/2702123.2702180

[23] Helen Kennedy, Rosemary Lucy Hill, Giorgia Aiello, and William Allen. 2016. The work that visualisation conventions do. Information Communication and Society 19, 6: 715–735. https://doi.org/10.1080/1369118X.2016.1153126

[24] Rosa van Koningsbruggen, Hannes Waldschütz, and Eva Hornecker. 2022. What is Data? - Exploring the Meaning of Data in Data Physicalisation Teaching. In Sixteenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '22). https://doi.org/10.1145/3490149.3501319

[25] Patti Lather and Elizabeth A. St. Pierre. 2013. Post-qualitative research. International Journal of Qualitative Studies in Education 26, 6: 629–633. https://

doi.org/10.1080/09518398.2013.788752

[26] Dan Lockton, Delanie Ricketts, Shruti Aditya Chowdhury, and Chang Hee Lee. 2017. Exploring Qualitative Displays and Interfaces. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17), 1844–1852. https://doi.org/10.1145/3027063.3053165

[27] Yanni Alexander Loukissas. 2019. All data are local: thinking critically in a data-driven society. MIT Press, London.

[28] Deborah Lupton. 2017. Digital bodies. In Routledge Handbook of Physical Cultural Studies, David L. Andrews, Michael L. Silk and Holly Thorpe (eds.). Routledge, London, 200–208.

[29] Deborah Lupton. 2018. How do data come to matter? Living and becoming with personal data. Big Data & Society 5, 2: 205395171878631. https://doi.org/10.1177/2053951718786314

[30] Deborah Lupton. 2019. Data Selves: More-than-Human Perspectives. Polity Press, Cambridge, UK.

[31] Maggie MacLure. 2013. The Wonder of Data. Cultural Studies - Critical Methodologies 13, 4: 228–232. https://doi.org/10.1177/1532708613487863

[32] Viktor Mayer-Schönberger and Kenneth Cukier. 2014. Big Data: A Revolution That Will Transform How We Live, Work, and Think. American Journal of Epidemiology 179, 9: 1143–1144. https://doi.org/10.1093/aje/kwu085

[33] Dietmar Offenhuber. 2020. Data by Proxy — Material Traces as Autographic Visualizations. IEEE Transactions on Visualization and Computer Graphics 26, 1: 98–108. https://doi.org/10.1109/TVCG.2019.2934788

[34] Dietmar Offenhuber. 2020. What We Talk About When We Talk About Data Physicality. IEEE Computer Graphics and Applications 40, 6: 25–37. https://doi.org/10.1109/MCG.2020.3024146

[35] Luci Pangrazio and Neil Selwyn. 2019. 'Personal data literacies': A critical literacies approach to enhancing understandings of personal digital data. New Media & Society 21, 2: 419–437. https://doi.org/10.1177/1461444818799523

[36] Amon Rapp and Maurizio Tirassa. 2017. Know Thyself: A Theory of the Self for Personal Informatics. Human–Computer Interaction 32, 5–6: 335–380. https://doi.org/10.1080/07370024.2017.1285704

[37] John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers. 2014. Personal tracking as lived informatics. In Conference on Human Factors in Computing Systems - Proceedings. https://doi.org/10.1145/2556288.2557039

[38] John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers. 2014. Personal tracking as lived informatics. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1163–1172. https://doi.org/10.1145/2556288.2557039

[39] Adrien Segal. 2022. Data Objects Thinking With Your Hands. In Making Data. Materializing Digital Information, Ian Gwilt (ed.). Bloomsbury Publishing, London, UK, 9–24.

[40] Nikhil Sharma. 2008. The origin of data information knowledge wisdom (DIKW) hierarchy. Preuzeto 25: 2021.

[41] Faisal Taher, John Hardy, Abhijit Karnik, Christian Weichel, Yvonne Jansen, Kasper Hornbæk, and Jason Alexander. 2015. Exploring Interactions with Physically Dynamic Bar Charts. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 3237–3246. https://doi.org/10.1145/2702123.2702604

[42] Alex S. Taylor, Siân Lindley, Tim Regan, David Sweeney, Vasillis Vlachokyriakos, Lillie Grainger, and Jessica Lingel. 2015. Data-in-Place: Thinking through the Relations Between Data and Community. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 2863–2872. https://doi.org/10.1145/2702123.2702558

[43] Stephen Wilson. 2003. Information Arts. MIT Press, London, UK.

[44] Emily Winter, Bran Knowles, Daniel Richards, Kim Snooks, and Chris Speed. 2022. Multitudes: Widening the research agenda for personal informatics design. In DRS2022, 1–16. https://doi.org/10.21606/drs.2022.415

[45] Annika Wolff, Daniel Gooch, Jose J. Cavero Montaner, Umar Rashid, and Gerd Kortuem. 2016. Creating an Understanding of Data Literacy for a Data-driven Society. The Journal of Community Informatics 12, 3. https://doi.org/10.15353/joci.v12i3.3275