Participation in Data Donation
Co-Creative, Collaborative, and Contributory Engagements with Athletes and Their Intimate Data

Gómez Ortega, Alejandra; Bourgeois, Jacky; Kortuem, Gerd

DOI
10.1145/3643834.3661503

Publication date
2024

Document Version
Final published version

Published in
DIS '24

Citation (APA)
Gómez Ortega, A., Bourgeois, J., & Kortuem, G. (2024). Participation in Data Donation: Co-Creative, Collaborative, and Contributory Engagements with Athletes and Their Intimate Data. In A. Vallgårda, L. Jönsson, J. Fritsch, S. Fdili Alaoui, & C. A. Le Dantec (Eds.), *DIS '24: Proceedings of the 2024 ACM Designing Interactive Systems Conference* (pp. 2388-2402). Association for Computing Machinery (ACM). https://doi.org/10.1145/3643834.3661503

Participation in Data Donation: Co-Creative, Collaborative, and Contributory Engagements with Athletes and their Intimate Data

Alejandra Gómez Ortega
Delft University of Technology
Delft, The Netherlands
a.gomezortega@tudelft.nl

Jacky Bourgeois
Delft University of Technology
Delft, The Netherlands
j.bourgeois@tudelft.nl

Gerd Kortuem
Delft University of Technology
Delft, The Netherlands
g.w.kortuem@tudelft.nl

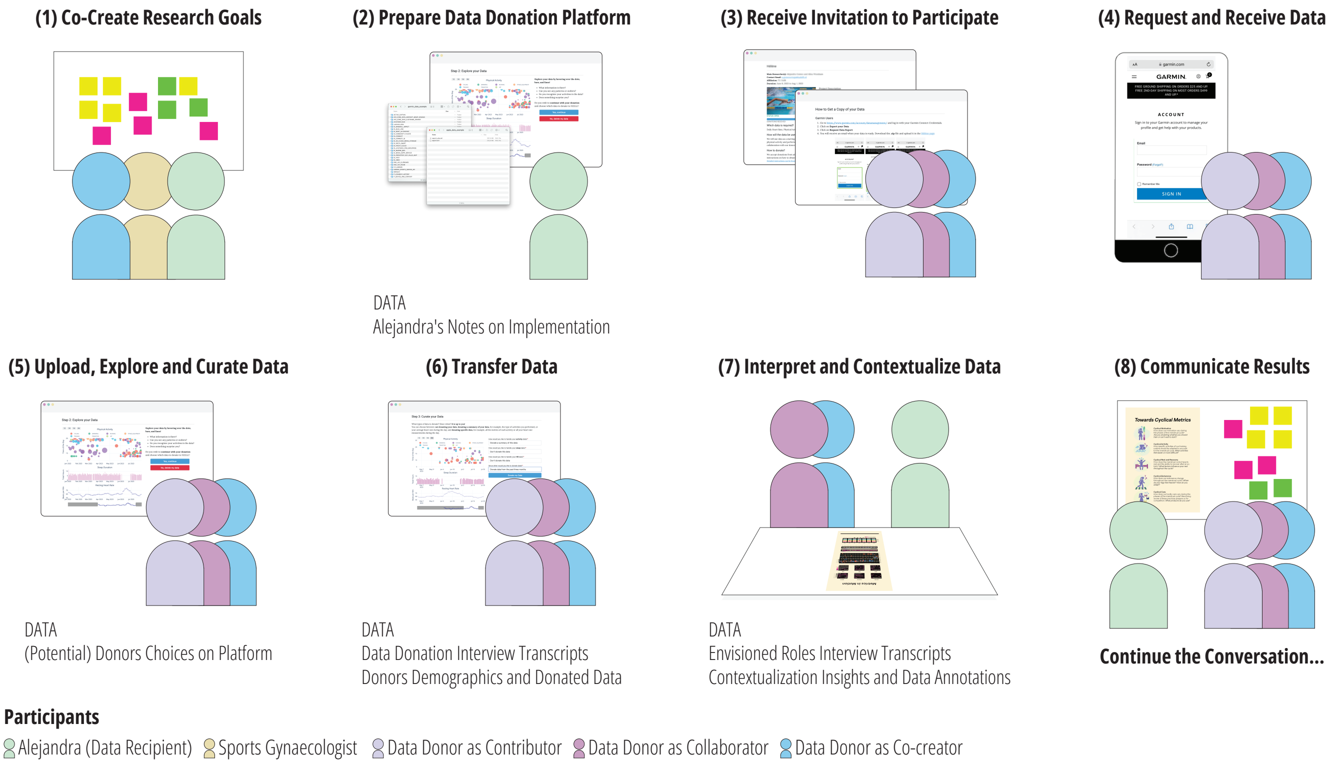


Figure 1: Timeline of the research, including data collected and generated throughout the data donation process according to the degree of participation of donors.

ABSTRACT

Data donation is an emerging practice enabling personal data collection for research. While it offers opportunities to access new insights into people's behavior and experiences through their digital-trace data, the role of individuals – as research participants – is limited in most data donation projects. They primarily contribute with data, limiting the perspectives included and accounted for around critical research-design decisions. In this paper, we explore the opportunity to embed data donation in research processes that are not only contributory but collaborative and co-created. To do so, we present a participatory data donation case study focused on athletes' perceptions of the impact of their menstrual cycle on their sports performance through their physical activity data. Based on the data donation experiences of 20 athletes, our paper provides insights into people's preferences and expectations in participatory data donation processes and discusses considerations for supporting various degrees of participation in future data donation research.

CCS CONCEPTS

• **Human-centered computing** → **HCI design and evaluation methods**; **Empirical studies in HCI**; • **Security and privacy** → *Social aspects of security and privacy*; • **Social and professional topics** → *Government technology policy*.

DIS '24, July 01–05, 2024, IT University of Copenhagen, Denmark

ACM ISBN 979-8-4007-0583-0/24/07
https://doi.org/10.1145/3643834.3661503

KEYWORDS

Data Donation; Data Feminism; Sensitive Data; Participation; Intimate Care; Sports.

ACM Reference Format:

Alejandra Gómez Ortega, Jacky Bourgeois, and Gerd Kortuem. 2024. Participation in Data Donation: Co-Creative, Collaborative, and Contributory Engagements with Athletes and their Intimate Data. In *Designing Interactive Systems Conference (DIS '24), July 01–05, 2024, IT University of Copenhagen, Denmark.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3643834.3661503

1 INTRODUCTION

Access to personal digital-trace data is increasingly valuable in scientific research across several contexts and domains. As a result, various methods and approaches enable and facilitate access: from connected prototypes developed by researchers and deployed in people's homes (e.g., [8, 40]) to the APIs of connected products and services developed and maintained by product-service providers that are available to researchers (e.g., [11, 51]). Data donation is an emerging approach to personal digital-trace data collection aimed at enabling and facilitating access [7, 11]. Similar to the donation of blood or organs, the donation of data is a voluntary transaction from a person who 'has'[1] data to another person (e.g., researcher) or entity (e.g., healthcare institution) who needs it.

In practice, data donation has been facilitated by the implementation of the European General Data Protection Regulation (GDPR) in 2018[2], specifically the *rights of access and data portability* [25, Art. 13 and 14]. These rights allow people to request and receive their personal digital-trace data from data controllers in a machine-readable format and (re)use it. Most empirical data donation projects invite people to request their data from a data controller and once they have received it, upload them to a research repository or a digital data donation platform. For instance, van Driel et al. [70] collected Instagram data from teens through data donation, and instructed participants to upload a copy of their data to a repository. Similarly, communication scholars employed data donation to collect data from YouTube and Google's browser to interrogate algorithmic recommendations during political elections in Germany [1] and Switzerland [6].

While data donation offers the opportunity to collect personal digital-trace data at the individual level, most approaches to data donation (e.g., [11, 60, 70, 75]) don't align with the collaborative and participatory nature of design and Human-Computer Interaction (HCI) research. Empirical approaches to data donation predominantly conduct research *on* people, specifically on their (sensitive) data, instead of *with* people. These approaches rely on data points or datasets disconnected from the specific contexts (i.e., times, places, bodies, devices) where data is generated and shaped [42]. Thus, they limit the perspectives included and accounted for around critical research-design decisions (e.g., What are the questions and objectives? What data can be collected and how?).

In this paper, we introduce a data donation case study where we investigate how to embed data donation in research processes that are not only contributory but collaborative and co-created [65]. Throughout the case study, we strive to conduct research *with* people and their sensitive data as opposed to *on* people, or *on* their sensitive data. Our case study sits at the intersection of sports performance and the menstrual cycle. In collaboration with 20 professional, semi-professional, and amateur athletes, we explore the perceived impact of the menstrual cycle on sports performance and envision ways to generate knowledge about it through their (1) menstrual cycle, (2) sleep, (3) heart rate, and (4) physical activity data. This context is suitable for exploring various degrees of participation as athletes are often misrepresented in research [17, 23]. Their participation as collaborators and co-creators can foreground their knowledge and experience(s) behind their digital-trace data.

This paper contributes with the following: (1) a participatory data donation approach where donors can contribute, collaborate and co-create with researchers through their personal digital-trace data; (2) a potential implementation of data donation in the form of a functional data donation platform; and (3) empirical findings from the case study, including insights into how athletes track their physical activity and interact with their data, and a set of holistic metrics they would like to track. By illustrating different forms of participation through personal data, our work can serve as a starting point for designers and researchers to reflect on the challenges and opportunities of participatory data donation approaches.

2 RELATED WORK

2.1 Degree of Participation in (Design) Research and The Role of Data

Shirk et al. [65] define the *degree of participation* in scientific research as *"the extent to which individuals are involved in the process of scientific research."* They propose a framework for public participation in scientific research. It comprises five types of projects based on the degree of participation: (1) *contractual* projects, where individuals and communities ask professional researchers to conduct a specific research project; (2) *contributory* projects, where researchers invite members of the public to contribute data; (3) *collaborative* projects, where researchers invite members of the public to contribute data, interpret data, and/or disseminate findings; (4) *co-created* projects, where researchers invite members of the public to design the research project and some members are actively involved in most aspects of the research process; and (5) *collegial* projects, where members of the public conduct research independently.

The design and HCI research landscape comprises various degrees of participation: from *people as subjects*, where people contribute insights to designers (i.e., contributors), to *people as partners*, where people have more influence over the design process and bring their expertise towards informing, ideating, and conceptualizing the design or research (i.e., co-creators) [15, 44, 56, 64]. Increasingly, data influences participation in design and HCI research. Previous work has developed several methods to involve participants and their data through design and research processes. These include Data-Enable Design [49, 71], Participatory Data Analysis [14], and Contextual Inquiry [32], among others. With our work,

---

[1] The term is in quotation marks as legal scholars have argued about the limits of ownership – as exclusive use– in the context of data [37, 57].

[2] The GDPR applies to the population of the European Union. Yet, the rights to access and data portability are available worldwide, since international companies rarely limit them by geography [9].

we expand on the methods and tools for conducting participatory (design) research with data. We explore how people can contribute to, collaborate, or co-create (design) research projects through data donation.

2.2 Data Donation

Data donation is a promising alternative to platform-centric approaches enabling personal data collection, such as the Application Programming Interfaces (APIs) of platforms and organizations [11, 51, 70]. Especially considering the current landscape, in which access to data through the APIs of certain platforms (e.g., Facebook, Twitter, Reddit) is increasingly restricted [11]. Thus, data donation is considered an individual-centered [11] or user-centered approach [51]; where researchers are able to access data *directly* from individuals – who get to opt-in or consent to their participation[3]. In contrast to the public data available through platform-centric approaches (e.g., all public posts from a person), data available through data donation are individual-level private data (e.g., all public posts from a person, the direct messages they have exchanged with others, and their activity logs) [11, 29, 60, 70]. This data offers researchers access to new insights and opens the way to investigate new research questions across sensitive domains. For instance, Razi et al. [60] collected private Instagram conversations from teens to identify online risks; including nudity and porn, sexual messages or solicitations, harassment, and violence, among others. Yet, requesting and accessing private data at the individual level also poses new challenges. These include informed consent [7, 37–39, 50], how donors can exert their autonomy [35, 37–39, 57] and preferences (e.g., deciding whether/ what/ and to whom to donate [68]), and the practical challenges donors face through the process of donating their data [2, 9, 28, 70].

Studies that apply data donation empirically employ one of two approaches: (1) digital platforms or repositories where donors upload a copy of their data, previously requested by a data controller [5, 11, 28, 29, 52, 60, 70], and (2) applications where donors consent to scrape data using their account(s) [11, 63]. Breuer et al. [11] applied and compared the two approaches using Facebook data and discussed the legal, privacy, and ethical implications. They concluded that the first offers higher transparency for donors but requires more effort and that both result in the collection of sensitive data. Moreover, empirical data donation research projects have been mainly *contributory*: donors contribute to a project by transferring their data and might further contribute by augmenting it with more data. For instance, Breuer et al. [11] integrated a survey on privacy concerns, and digital habits into their Facebook data donation approach, while Razi et al. [60] asked donors to annotate their private conversations on Instagram based on perceived risk. In our previous work [28, 29], a subset of the donors participated as collaborators, as they participated in activities related to interpreting and contextualizing their data. To date, no projects have invited donors to participate as *co-creators*, limiting the perspectives included and accounted for around critical research-design decisions. In this paper, we continue to build on and illustrate the important ethical considerations and practical approaches from the data donation literature. Additionally, we explore and expand on the degree of participation in data donation projects: from contributory to contributory, collaborative, and co-created.

2.3 Sports Performance and the Menstrual Cycle

Through a case study, we situate our participatory data donation research in the intersection of sports performance and the menstrual cycle, where athletes increasingly rely on digital technologies to track their physical activity and increase their performance [16]. These technologies play an instrumental role in measuring performance and feeding data back to athletes (e.g., resting heart rate, respiratory rate, sleep duration), and an experiential role in supporting and enhancing the sports experience, allowing athletes to have a closer connection to their bodies [69, 73]. Although digital technologies provide feedback and advice to athletes in various aspects, including recovery and training, one vital health indicator [24] is often overlooked; the menstrual cycle[4]. Likewise, most of the research in sports science – informing the design and development of these devices – has been conducted with men [17, 18]. Yet, findings are inappropriately applied to women athletes [23].

The impact of the menstrual cycle on physical performance is recognized as a key consideration for sports. Previous research demonstrates that hormonal fluctuations throughout the menstrual cycle affect athletes during training and competition (e.g., [13, 41, 46]). Similarly, physical activity impacts the menstrual cycle. For instance, amenorrhea (i.e., the absence of menstruation) is highly prevalent among athletes, yet, it is often not discussed or reported [72]. Elite sports organizations such as the Chelsea Football Club and the Red Bull High-Performance Division are using FITR Woman [26], a commercial smartphone app to track athletes' performance and bodily symptoms during the various phases of the menstrual cycle.[5] The app invites athletes to self-report their physical activity and symptoms (e.g., cravings, sensitive breasts) and offers training and nutrition advice. Feminist HCI research has widely criticized these types of apps, arguing that by collecting intimate data and translating it into more or less accurate predictions they restrict other types of knowledge [12, 36, 62, 67, 74]. In this paper, we build on this call by exploring the perceived impact of the menstrual cycle on sports and collectively envisioning ways to generate new forms of knowledge about it.

3 METHODOLOGY

In this paper, we investigate how to embed data donation in research processes supporting various degrees of participation, namely *contributors*, *collaborators*, and *co-creators*, as proposed by Shirk et al. [65]. We build upon how data donation has been operationalized and applied in previous research (e.g., [28, 29, 60, 70]) to apply it in a pre-registered case study[6] at the intersection of sports performance

---

[3] Platform-centric approaches don't require individual participants to consent to the use of their data. In some cases, it is assumed they consented to secondary uses of their data through the Terms and Conditions of a specific product or service.

[4] Notably, the menstrual cycle was also disregarded in the initial release of Apple's Health app in 2014 [22]

[5] The menstrual cycle phases include, early follicular, late follicular, ovulatory, early luteal, mid-luteal and late luteal. It starts with menstruation during the early follicular phase [46].

[6] Pre-registration: https://doi.org/10.17605/OSF.IO/QK2MV

and the menstrual cycle. This context is suitable for exploring various degrees of participation in data donation, as athletes who use wearable devices constantly interact with data. They can account for the nuances of these interactions, and their participation as collaborators or co-creators can provide valuable contextual insights – often misrepresented in sports research [17, 23] – and further shape the research. Similar to previous empirical data donation research, we approach data donation through a process where we invite (potential) donors to request their data from a data controller and upload it to a digital data donation platform. We augment this process[7] by incorporating dedicated activities for each degree of participation:

- **Contributors:** Donors who decide to participate as contributors primarily request and receive a copy of their data, explore it, and decide what (not) to transfer through the data donation platform. These activities respond to a call from prior research illustrating how at the time of donation, donors often *"don't know what they don't know"* about their data [28, 38]. Additionally, they participate in an interview with the first author about their data donation experience.
- **Collaborators:** Donors who decide to participate as collaborators, in addition to the above, participate in a one-on-one session with the first author where they are invited to interpret, contextualize, and analyze their data. These activities build on our previous data donation research [28, 29].
- **Co-Creators:** Donors who decide to participate as co-creators, in addition to the above, can opt to participate in a session to scope the research goals and activities (Section 3.1) and in meetings throughout the process to discuss the partial results and research direction.

Following the Feminist tradition of positionality and reflexivity [3, 4, 21], we report that the first author is an amateur athlete who menstruates, with a strong motivation to undertake this case study.

3.1 Co-Creation of the Research Goals and Data Needs

The initial aim of the case study was to investigate: **How do athletes perceive the impact of the menstrual cycle on sports performance?** We further scoped this goal together with four athletes, who later became co-creators, and a sports gynecologist. We contacted them by posting flyers in local sports associations inviting adult athletes with vulvas[8] to a creative session focused on menstruation in sports. We expressed our appreciation to them by inviting them to a Q&A with the sports gynecologists after the session.

Athletes proposed to expand the focus: from investigating the perceived impact of the menstrual cycle in sports to envisioning new "metrics" or ways to generate knowledge about it. We discussed relevant data sources and decided to focus on: (1) menstrual cycle, (2) sleep, (3) heart rate, and (4) physical activity, data as proxies recovery, and performance. Finally, we co-defined a value-gain strategy responding to athletes' wish to learn about their data and how to interpret it. It comprised a (printed or digital) poster representing their data (Fig. 3a) with guiding questions on how to interpret them, as athletes often enjoy displaying their achievements and mementos (e.g., bib numbers and medals) [66], and a commitment to communicate our results throughout the process (Fig. 3b). Thus, aligned with previous research [30, 57, 61], we aimed to honor donors' contributions and efforts by carefully engaging with their data, creating opportunities and artifacts for them to engage with their data, and building and maintaining a relationship.

3.2 Design and Development of the Data Donation Platform and Procedure

We describe the activities conducted to design and develop the digital prototype of a data donation platform[9] and the resulting data donation procedure. We illustrate how we deployed and applied these in the case study. These activities were reviewed and approved by our institution's Human Research Ethics Committee and Privacy Team.

3.2.1 *Providing Choices around Data Disclosure.*

Together with co-creators, we decided to focus on the following data: (1) menstrual cycle, (2) sleep, (3) heart rate, and (4) physical activity. Due to the intimate nature of the menstrual cycle data, we invited donors to self-report the information they considered relevant. The remaining data (i.e., sleep, heart rate, and physical activity) are collected by wearable devices commonly used by athletes. Thus, we invited donors to transfer retrospective data collected by Garmin, Apple Watch, or other devices compatible with the Apple Health ecosystem. Initially, we wanted to focus only on Garmin devices – designed with sport-specific features and most commonly used by athletes [47] – but we opted for more variety to lower the participation threshold and allow for participation by athletes who do not own a Garmin device.

To prepare the data donation platform to parse and visualize the data, the first author made several requests to Garmin and Apple Health and reflected on this process through a research journal[10]. In the platform, we designed an interactive visualization where potential donors could explore their data on a timeline (Fig. 2). The x-axis showed the month and year and the y-axis showed the value for sleep and heart rate data, and the time of the day for physical activity data. In the visualization, potential donors could filter the data by activity (e.g., running), and time (e.g., day, week, month, six months). We opted for this representation as previous research describes how viewing data arranged by time helps people reason about it [55].

On the data donation platform, we invited donors to explore and visualize (Fig. 2) their data locally *before* they decide whether and what to donate. This aimed to promote their interaction and familiarization with the data and support them in identifying and defining initial boundaries around its sensitive elements. We wanted to offer

---

[7] We formalize this approach into a Sensitive Data Donation method in: Gómez Ortega et al. [31]

[8] We use the term "athletes with vulvas" and not "athletes who menstruate" as we wanted to include diverse menstrual experiences, including those of athletes who experience amenorrhea.

[9] The data donation platform was implemented using TypeScript and the Python web framework Django. Potential donors register on the platform by creating an account with an email address to manage their donations. This allows them to update their participation preferences and informed consent over time. URL and source code available at: https://datadonation.ide.tudelft.nl/

[10] Research journal template in the supplementary material.

(a) Overview of the entire dataset.

(b) Overview of a week of data.

Figure 2: Interactive data exploration visualization on the data donation platform.

donors meaningful options around disclosure that would still allow us to conduct the research. Therefore, we explicitly prompted them to decide if and what types of data they wanted to transfer and for how long (i.e., the last three or six months). We intentionally chose not to collect all available retrospective data (i.e., from when donors started using a device to the date of donation) by bounding the time range. Thus, we adhered to the data minimization principle. Furthermore, potential donors could choose to transfer an 'overview' of their data, describing the summary and statistics of each day or activity, or the 'specifics' of the data, with timestamped logs for each variable based on the sampling frequency of the device. Finally, potential donors could remove any given activity or data point from the dataset through the interactive visualization. They participated asynchronously and on their own time in this activity. Giving donors various choices around sensitive information disclosure pushed us to consider how to approach partial, messy, or incomplete datasets.

After exploring and curating their data, donors were prompted to provide initial informed consent through the data donation platform, transfer their data to researchers, and decide on their degree of participation (i.e., contributors, collaborators, collaborators). Donors participated asynchronously and on their own time in this activity. We automatically sent an email to donors who decided to participate as collaborators or co-creators inviting them to schedule a one-on-one session with the first author to further explore, situate, and contextualize their data. From this moment on, donors could modify or revoke their consent through the data donation platform and verbally in direct contact with the researchers. For example, by deciding to stop sharing a type of data or all data. If and when necessary, we recorded these changes through the platform. Following this activity, we conducted semi-structured interviews with donors who decided to participate as contributors about their data donation experience[11].

3.2.2 Supporting Interpretation and Contextualization.

We invited collaborators and co-creators to a one-on-one session to interpret and contextualize their data and reflect on their data donation experience. During the interview, we used the personal data poster (Fig. 3a) to facilitate interpretation and reflection[12]. We represented (1) the resting heart rate together with the physical activity types and duration in a calendar view; (2) the type and duration of physical activity for each hour of the day over time; (3) the resting heart rate and sleep duration in a timeline; and (4) the physical activity types and average duration for each day of the week. When any of this data was not available, we left it blank. With the calendar view, we aimed to support donors in identifying patterns and outliers in their data. While with the timeline used to represent physical activity, resting heart rate, and sleep data, our goal was to help donors identify correlations between the different types of data and their menstrual cycle over time. During this session, we also discussed with co-creators their envisioned roles and ways forward for collaboration. Finally, we explicitly prompted donors to (re)assess their participation (i.e. modify or revoke their consent) after the interviews and while sharing partial results with them.

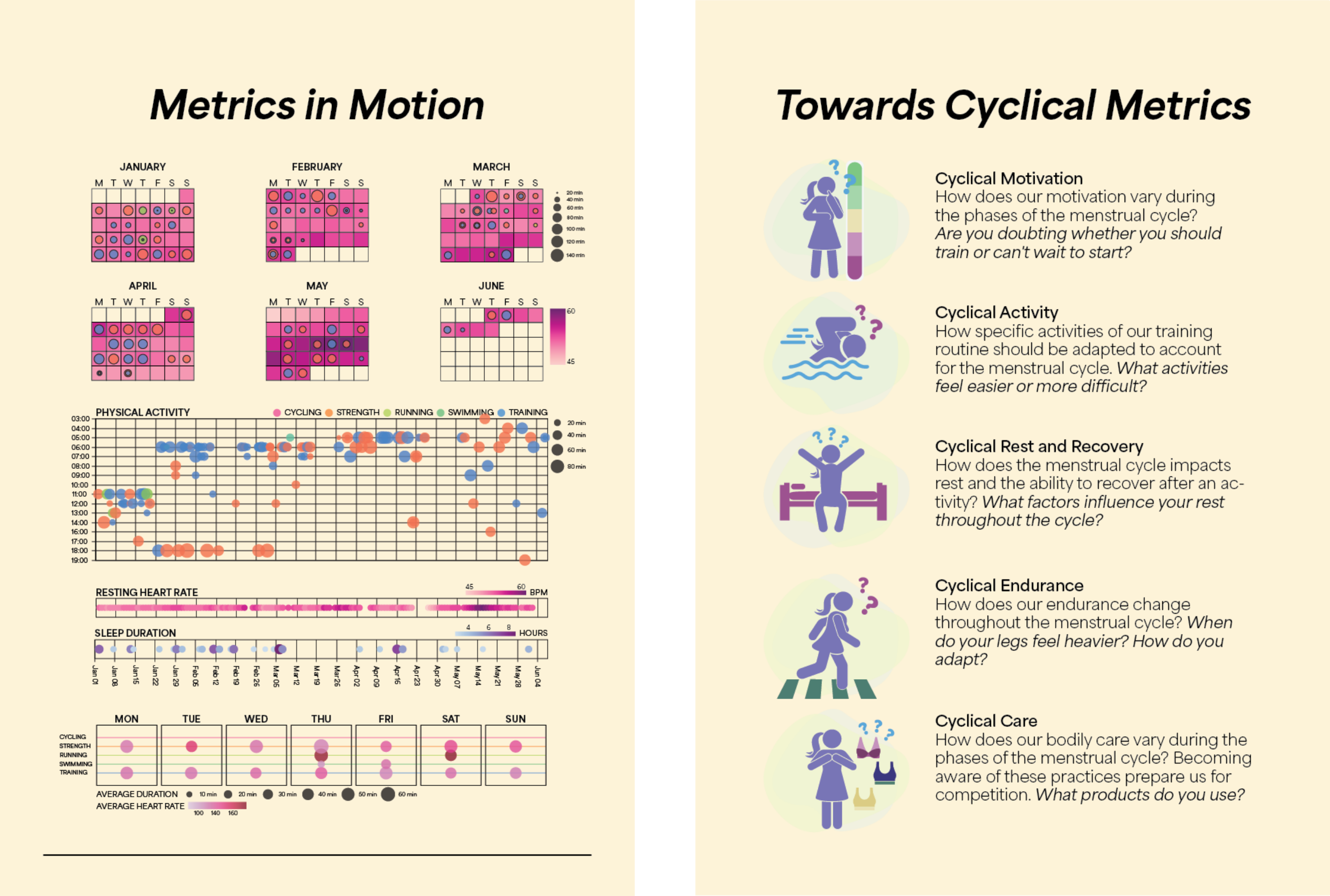
3.3 Participation in Data Donation Process

To invite donors to participate in the case study, between June and July 2023, we reached out to adult athletes with vulvas who were active users of Garmin, Apple Watch, or other wearable devices compatible with Apple Health worldwide, and we invited them to donate their data. We are aware that not all people with vulvas have a menstrual cycle. Yet, we wanted to include diverse menstrual

---

[11] Data donation experience interview protocol in the supplementary material.

[12] Contextualization interview protocol in the supplementary material.



Cyclical Motivation
How does our motivation vary during the phases of the menstrual cycle? *Are you doubting whether you should train or can't wait to start?*

Cyclical Activity
How specific activities of our training routine should be adapted to account for the menstrual cycle. *What activities feel easier or more difficult?*

Cyclical Rest and Recovery
How does the menstrual cycle impacts rest and the ability to recover after an activity? *What factors influence your rest throughout the cycle?*

Cyclical Endurance
How does our endurance change throughout the menstrual cycle? *When do your legs feel heavier? How do you adapt?*

Cyclical Care
How does our bodily care vary during the phases of the menstrual cycle? Becoming aware of these practices prepare us for competition. *What products do you use?*

(a) Example of a personal data poster, excluding the menstrual cycle data. Shown with permission of the donor.

(b) Example of one of the infographics with partial results returned to all donors.

Figure 3: Elements designed for donors to interact with their data, the research outputs, and to honor their contributions.

experiences including those of athletes experiencing amenorrhea – common among high-performing athletes [72]. We used purposive and snowball sampling by reaching out to sports associations, teams, and individual athletes with a strong social media presence, regardless of the discipline, which shared our 'call to donate'[13] with their communities through social media, newsletters, and mailing lists. Additionally, we distributed flyers at local sports events. While reaching out to donors, we specified who was conducting the research, what the goal was, who could participate and how, and we shared the URL of the data donation platform. On the data donation platform, we provided potential donors with a statement describing how we intended to use the data and detailed visual instructions describing how to obtain their data from Garmin or Apple Health. Potential donors could contact us if they had any questions or needed additional information. We intentionally decided not to incentivize donors financially due to the ethical precedent of not compensating research subjects as it could limit their ability to offer consent voluntarily [34, 43, 53]. We determined that voluntariness and autonomy should prevail around transactions of sensitive data.

20 athletes, professional (n=2), semi-professional (n=8), and amateur (n=10), (referred to in the paper as donors D1-20) volunteered to participate in our research by donating their data. They were active users of a Garmin Wearable (n=13), an Apple Watch (n=5), or a third-party smartwatch synchronized with Apple Health (n=2). All donors identified as women, they ranged in age from 22 to 55 years (mean=29, median=32). Donors were located in the European Union (EU) and South America – obtaining a copy of the data, enabled by the GDPR, was also possible for donors outside the EU. Six donors participated in the case study as co-creators, ten donors participated as collaborators, and the remaining four participated as contributors.

3.4 Research Data and Analysis

We conducted semi-structured interviews with all donors around their expectations, information gained from autonomously exploring their data, and the perceived value derived from participation,

---

[13] The exact phrasing we used to communicate our project can be found in the supplementary material.

and with the ten collaborators and six co-creators focused on collaboratively interpreting their data. Additionally, we had follow-up discussions with the six co-creators regarding the preliminary results and the directions of the case study. We anonymized and transcribed the interviews using MS Office and manually reviewed each transcript.

In sum, we generated and collected the following: (1) the researcher's journal entries and reflections, (2) the donor's choices on the data donation platform, including their preferred degree of participation, (3) the donated menstrual cycle, sleep, heart rate, and physical activity data, (4) the donors' perceptions of their data donation experience, in the form of interview transcripts, and (5) the specific insights derived from the contextualization of the data, in the form of interview transcripts. We analyzed the interview transcripts using Reflexive-Thematic Analysis within a constructionist view [10]. The first author went through the transcripts to familiarize with the data and inductively coded the entire dataset using ATLAS.ti; reviewed the codes, and grouped them into tentative themes; and iteratively reviewed and refined the themes[14]. The resulting themes and sub-themes were discussed and refined with co-authors. We report on these in the following sections.

4 CASE STUDY RESULTS: MENSTRUAL CYCLE, TRACKING, AND SPORTS PERFORMANCE

4.1 Tracking and Interacting with the Data

Donors use smartwatches and wearable devices to plan their day-to-day activities and monitor them as they go (e.g., to make sure they are running at the planned speed). Once an activity is over, they glance at the data to review the summary and their overall progress. For specific activities, such as intense training or competitions, they compare with their past performance – stored in the form of Personal Records (PRs) : *"I don't look back* [at the data]. *Not really. I just look for PRs"* (D3, co-creator). Outside specific activities, they engage with prescribed metrics (e.g., body battery[15], stress) throughout the day to better understand what is happening with their bodies. These metrics are sometimes a poor quantification of their experiences, which becomes a source of doubt and concern.

> "I don't know if it's an error, but sometimes I feel like I slept very well but my *body battery* doesn't recharge that much." (D9, contributor)

Other than these brief interactions, donors rarely look (back) at their data and never see an overview of their past data, *"it's like, oh, what did I do yesterday? What do I have for tomorrow? But I don't check anything else."* (D2, collaborator). For professional and semi-professional athletes, coaches and training staff have direct access to data and interpret them. They are assumed to have the knowledge and tools to correctly do so *"I like that because there are many things that I ignore or do not know how to interpret well"* (D7, co-creator). In these cases, the menstrual cycle is rarely discussed or considered, in part because the coaches of all athletes we interviewed were cisgender men.

4.2 Perceived Impact of the Menstrual Cycle in Sports Performance

Donors perceived that *physical*, *mental*, and *other* factors have an impact on sports (Fig 4). We categorize them in this way, as it is common for athletes to train and address them through different activities. Yet, these are interrelated. For instance, during the early follicular phase, with the onset of menstruation, water retention can lead to feeling bloated (i.e., *physical factors*). This, together with bleeding can increase tension or anxiety over physical appearance and spotting and decrease training motivation (i.e., *mental factors*), which in turn, might shape the choice of activity, clothes, or menstrual products (i.e., *other factors*).

> "I won't be swimming that day, but maybe in two days, you know? So, I will, like, reschedule my week to avoid the strongest day of my period being the day of swimming." (D4, co-creator)

Physical factors impact (1) the range and types of activities that athletes can perform – for instance, sensitive breasts make it difficult to perform activities that require running at high speeds or jumping; (2) how athletes perform and how they perceive their performance – for instance, a workout can be considered conditionally good or productive considering the phase of the cycle; and (3) the type of activity that athletes can participate in – for instance, whether a lightweight boxer will be allowed to compete or must modify her diet before competition to account for weight fluctuations.

> "When I'm training for a fight it is much harder because I need to keep checking my weight every day and then it's frustrating when you go and check and you did a hard training, and you were eating well, and it's like nothing changes, or even it [weight] goes up." (D11, co-creator)

Notably, a prominent physical factor is amenorrhea, or the absence of menstruation, especially for athletes competing in lightweight categories (e.g., boxing, rowing, weightlifting) or athletes who have recently lost weight. *Mental factors* impact (1) whether and how athletes approach sports – for instance, if they want to train and where *"it doesn't feel comfortable to go to the gym"* (D2, collaborator); and (2) their feelings before, during, and after sports. Notably, donors emphasized that performing any sport positively impacts how they feel after – independently of any other factor, *"I think it helps me feel energized"* (D18, contributor). *Other factors*, such as period products and clothes, affect the type of physical activity, training, and competition that athletes can and are comfortable doing.

> "I will never be able to do an Ironman [long-distance triathlon] with my period, never. Nothing will resist. Not even a tampon. So, when I was in my first Ironman I got an injection to stop it." (D3, co-creator)

4.3 Towards Cyclical 'Metrics'

Collaborators and co-creators were involved in crafting metrics that they would like to track as part of their routine. We use the term 'metric' loosely, as these don't aim to measure the body, but represent a subjective experience. These are not intended to (more or less accurately) predict but rather to foster understanding. Donors

---

[14] The overview of the thematic structure can be found in the supplementary material.

[15] Body Battery is a metric designed by Garmin to "show the effects of physical activity, stress, relaxation, and the restorative power of sleep together in a single place" [27]. It goes from 0 to 100.

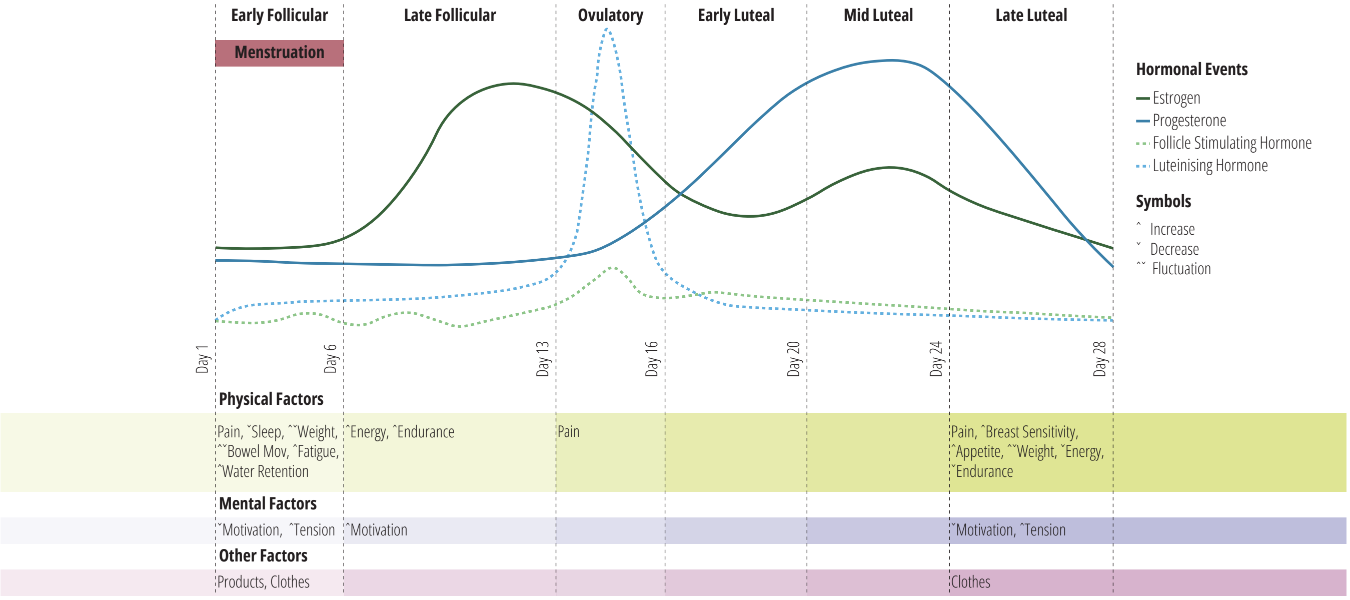


Figure 4: Factors that donors perceive impact physical activity, training, and competition; represented on the phases of a eumenorrheic 28-day menstrual cycle. Adapted from Carmichael et al. [13].

choose to craft these metrics on the cyclical nature of the menstrual cycle. They envisioned cyclically reflecting on them, through the menstrual cycle and after multiple cycles.

- **Cyclical Motivation:** How motivation to train varies through the cycle. It can be represented through a subjective ordinal scale in which athletes reflect on how much they look forward to a specific activity and the factors influencing this, *"I think that's really an indicator of how you are feeling well because if you are not feeling well, you are not looking forward to it. [...] For me, the low scale is «I'm doubting whether I should train» and then the high scale is «I can't wait to get started»"* (D12, co-creator)
- **Cyclical Activity:** How specific activities (e.g., flexibility, strength) and types of activity (e.g., running, swimming) from the training routine should adapt to account for the menstrual cycle. It can be represented through a qualitative reflection process that accounts for the experiences with different activities throughout the cycle and the factors influencing these; *"maybe it also has something to do with the body. Like when I have my period, running, for me it's very hard like I feel weak and not that motivated, but maybe doing some yoga will help me."* (D3, co-creator).
- **Cyclical Endurance:** How the menstrual cycle impacts endurance (i.e., the capacity to withstand a training of competition) and performance (i.e., how well athletes execute a specific training or competition). It can be represented quantitatively, for instance through a proxy for effort derived from existing data, *"you can see if I made more or less effort comparing if the heart rate was higher or lower, and at the same time it could be correlated with the speed"* (D5, co-creator); and qualitatively, for instance by reflecting on one's feelings: *"the data is not a measure of how good or how fit I'm feeling for exercising [...] but there is a clear difference that not every day I'm feeling the same."* (D4, co-creator).
- **Cyclical Rest and Recovery:** How the menstrual cycle impacts rest and recovery after sports. It goes beyond existing pre-defined metrics such as sleep duration and body battery and aims to represent athletes' bodily knowledge around sleep, rest, recovery, and fatigue. It can be represented through qualitative observations on specific factors, *"I feel that the rest is different [during menstruation], I think that my sleep is like heavier"* (D2, collaborator) and documenting occurrences, such as pain: *"when the pain is very strong I wake up"* (D5, co-creator).
- **Cyclical Care:** How donors care for their bodies throughout the menstrual cycle, especially during training, to account for it during competitions. For instance, how they nurture their bodies, manage and train with pain, account for breast sensitivity by wearing 'the right' bra, or choose a menstrual product based on the type of activity and its duration. It can be represented through qualitative observations over time. *"Sometimes I was really in pain for my training, and the guys at the gym used to say yeah, have a day off. But if it* [pain] *happened on the day* [of the fight], *I can't have a day off, so I need it just to go anyway."* (D11, co-creator).

5 RESULTS: REFLECTIONS ON THE DATA DONATION PROCESS

5.1 Motivations, Expectations, and Degrees of Participation of Data Donors

The main factor steering donor's willingness to contribute to the case study was interest in the specific research context. Interest derives from three motives; each carries underlying expectations.

The first motive is relatedness, identifying oneself in the context and goals of the project and the community it is directed towards. Relatedness spans multiple experiences, such as being a woman, being an athlete, and having specific experiences (e.g., absent periods) and anxieties (e.g., leaking) with menstruation in sports contexts. Donors who relate to the research perceive it as highly relevant to their personal experience and are therefore eager to contribute. Their underlying expectation is to learn through the research process and from its outcomes and apply these learnings to their individual experiences.

> "This is a topic that is very, very interesting for me because I'm involved with the two subjects or topics that you are touching on within your research which are sports and menstruation. Because I have had during my whole life problems with that. So, it's very, very interesting to me to see what the relation is between these two and how my whole data, that is recorded in these gadgets that I use for sports, how can like help me to understand how my body works." (D4, co-creator)

The second motive is solidarity, recognizing oneself as part of a group and wanting to help others within that group; in this case, by contributing to research. Similar to relatedness, solidarity stems from feelings of sameness and belonging to the community the case study is directed towards (i.e., professional, semi-professional, and amateur athletes). Donors whose primary motivation is solidarity experience feelings of satisfaction or a 'warm glow' – an emotional reward from contributing to research. Their underlying expectation centers on the outcomes of the research and the impact they can have on the daily experience of other athletes.

> "Since I was a little girl I have been playing sports. (...) So, if I can help other generations to have more knowledge and to be part of a world with more sports, happy to help." (D3, co-creator)

The third motive is familiarity, belonging to a specific community (e.g., team, group of friends) already involved with the project and wanting to contribute due to the closeness with this community. Donors whose primary motivation is familiarity already trust the specific community they belong to, and as a result, have increased trust in the research. Their underlying expectation stems from trust and has to do with accountability; they expect researchers to communicate about the next steps and the outcomes of results of the research.

> "Well, [a friend] mentioned to me that you were doing this research and I think she already helped you before with it and she said it was good because she learned as well." (D11, co-creator)

The degrees of participation also carry expectations about what participation entails. The main goal of contributors is to support research; they expect to do so through data; being part of the data and increasing the amount of data. For instance, in response to the 'gender data gap' [19]: *"that's why I did it because I think there is not enough data"* (D13, contributor). Collaborators, in addition to supporting research, also expect to learn something new by interpreting and contextualizing their data; about the research, their data, or themselves, *"I was thinking on how like to learn about this data and how I can have those records to like be useful for my day by day"* (D2, collaborator). The intention to interpret and contextualize the data shapes how they engage in the other stages of the data donation process. For instance, several collaborators (D2, D4, D10, D12, D18) expressed that they did not explore their data on the platform because they knew they were going to do it with us and preferred our guidance. Co-creators, in addition to the above, expect to engage in a more active role throughout various research activities. It includes scoping future iterations of the project, *"I might have some, I don't know, like suggestions for you or like interesting topics that we would like to address in a second version of the project"* (D4, co-creator) and building a community for athletes to share experiences and discuss with others, *"I would love to know other women doing sports, maybe create a community and help each other"* (D3, co-creator).

5.2 Obtaining, Exploring, Curating, and Contextualizing Sensitive Data

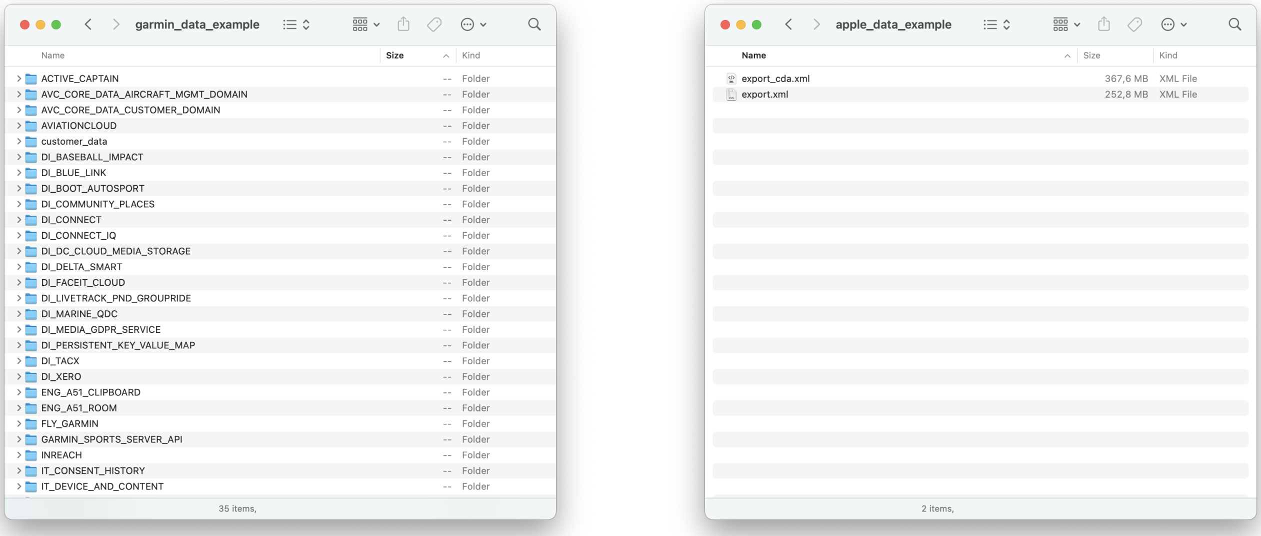
For most donors, being part of a data donation process is a new, and almost accidental, experience. First, because of the process itself. Second, because of the decisions that the process invites.

> "I had never done it, and I never used the watch with any intention, like thinking it might be useful to someone, just for myself. And well, it's great that it can be useful for someone, for research." (D9, contributor)

The process starts with requesting and obtaining a copy of their data – an already unfamiliar procedure that most donors did not know was possible. The majority of donors perceived this process as easy and straightforward, *"it was like exactly the step by step that you described"* (D2, collaborator); although some ran into practical difficulties. For instance, forgetting their username or password, or having to wait a long time (i.e., more than six days) for their files. Throughout this process, donors remain largely in the dark regarding their data. Even if they technically 'have' a copy, it is opaque and meaningless as it is either in a large ZIP file containing multiple folders with no structure or guidance (Fig. 5a), or in a large ZIP file containing files in a format donors are unfamiliar with (Fig. 5b). Most donors did not open the ZIP files as they were not sure how to open, read or interpret them; those who did found them just as intimidating.

> "I saw in the beginning that the ZIP [file] has so many things. There are so many folders that I was like OK this is too much to look at. So no." (D15, contributor)

Hence, when donors upload their (sensitive) data into the data donation platform they are doing so blindly. This is awkward and confronting, especially because of the type of information, which they considered "*very personal*, *sensitive*, *confidential*, *internal* information" (D2, D6, D8, D20). Privacy-concerned donors experienced this process as swimming against the tide as they are increasingly aware of the importance of data protection, *"I felt awkward, actually, because I know that my personal data should be protected"* (D3, co-creator). Here, two main factors helped mitigate their concerns: (1) the research context, and (2) the exploration and curation process built into the platform. The data exploration process helped



Name	Size	Kind
ACTIVE_CAPTAIN	--	Folder
AVC_CORE_DATA_AIRCRAFT_MGMT_DOMAIN	--	Folder
AVC_CORE_DATA_CUSTOMER_DOMAIN	--	Folder
AVIATIONCLOUD	--	Folder
customer_data	--	Folder
DI_BASEBALL_IMPACT	--	Folder
DI_BLUE_LINK	--	Folder
DI_BOOT_AUTOSPORT	--	Folder
DI_COMMUNITY_PLACES	--	Folder
DI_CONNECT	--	Folder
DI_CONNECT_IQ	--	Folder
DI_DC_CLOUD_MEDIA_STORAGE	--	Folder
DI_DELTA_SMART	--	Folder
DI_FACEIT_CLOUD	--	Folder
DI_LIVETRACK_PND_GROUPRIDE	--	Folder
DI_MARINE_QDC	--	Folder
DI_MEDIA_GDPR_SERVICE	--	Folder
DI_PERSISTENT_KEY_VALUE_MAP	--	Folder
DI_TACX	--	Folder
DI_XERO	--	Folder
ENG_A51_CLIPBOARD	--	Folder
ENG_A51_ROOM	--	Folder
FLY_GARMIN	--	Folder
GARMIN_SPORTS_SERVER_API	--	Folder
INREACH	--	Folder
IT_CONSENT_HISTORY	--	Folder
IT_DEVICE_AND_CONTENT	--	Folder

(a) Files returned by Garmin when requesting the data.

Name	Size	Kind
export_cda.xml	367,6 MB	XML File
export.xml	252,8 MB	XML File

(b) Files returned by Apple Health when requesting the data.

Figure 5: Structure of the files returned from the data controllers in the context of the case study.

donors see their data and better understand the types of data being transferred. It also invited them to challenge and have questions about their data; how it was encoded, *"I was really confused that all my rowing activities were generic activities"* (D8, contributor), and the reasons behind specific gaps, patterns, or outliers *"I saw like my training times, mornings and nights. I saw once that I trained like at midday and I was like what happened that day?"* (D5, co-creator). The data curation process prompted donors to pay attention to and care for the transaction being made. In addition, it explicitly invited donors to choose what data to donate and for how long, which fostered a sense of security and trust.

> "At first, I did it very automatically and then I saw that there was like a choice about the data and the amount of time. There I felt that I should pay more attention because suddenly there can be sensitive information. [...] So as I progressed, I became more alert because I was super calm at first. I don't know. It's like a form. When you are filling out a form and they need your name and surname, well you are calm, you give your name and surname, but then, when they have more fields, more information, like your phone number, or email address, then you begin to pay more attention. Like, wait, this is getting a bit long. Are they going to call me? What I am doing? That is how I felt it." (D13, contributor)

Nonetheless, when it came to curating their sensitive data and defining their personal boundaries most donors chose to donate as much data as possible. This decision stems from three motivations. First, full openness and not caring much about privacy considering the various entities that already have access to data about themselves, including Garmin or Apple, *"I was like, well, everything I'm seeing I'm willing to share. I'm always quite open"* (D8, contributor). Second, wanting to contribute to the research as much as possible, *"it was super specific, like until when, for what, which type. So I felt secure, and because of that, I tried to give my best. Like, they're doing a good job. I want to help too"* (D3, co-creator). Third, wanting to see and explore as much data as possible, or to compare it or relate it to a specific event (e.g., injury, competition). Especially, for collaborators and co-creators with the intention to participate in the contextualization interview. In fact, some of them had already a specific goal for this activity, *"I would like to learn how to interpret the data we have available in these phones and watches and decide, well, what interests me and what doesn't."* (D7, co-creator). Through the collaborative interpretation of their data, collaborators and co-creators (re)discovered patterns and events reflected by the data, confirmed their expectations using the retrospective data as a reference point, and challenged their assumptions from seeing the data over time. In addition, they gained more clarity about their data, awareness about their bodies, and new questions for the future.

> "I was very interested in the visualization of my data and identifying these patterns that we just discussed. (...) After the interview, now I have like these open questions. Yeah, I feel motivated also to help research. And I like the fact that right now I'm motivated to be part of it and not like OK, I donated my data, and I went to the interview. No, I'm like in a feedback process and I really want to be part of it." (D4, co-creator)

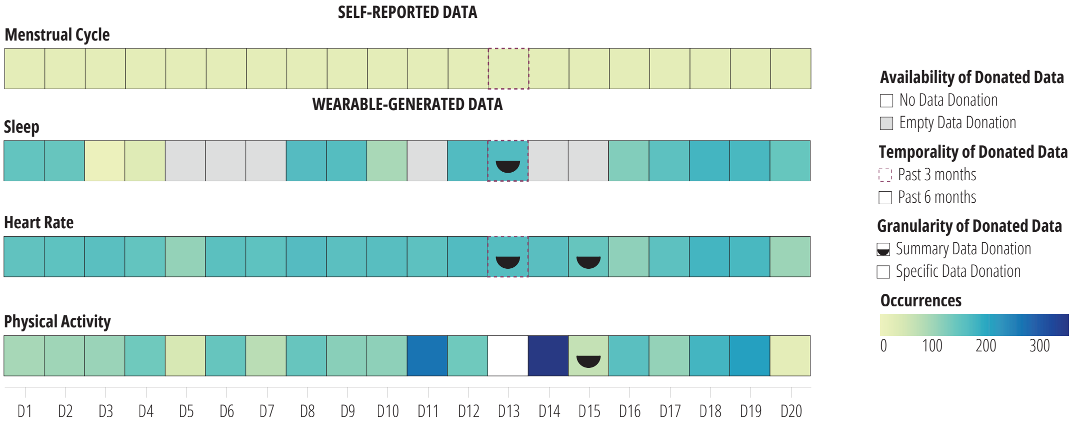


Figure 6: Summary of the data donated by each donor and how it is configured through their choices.

5.3 Types and Characteristics of the Donated Data

The data donation process results in different types of data with characteristics that are shaped by the degree of participation. The baseline is the contributor's data (Fig. 6). It is spatiotemporal and retrospective, going back up to six months from the date it was requested. It is decontextualized and its content and specificity depend on the choices made by donors. In turn, these are influenced by how donors use their devices, *"I donated everything I could except the sleep [data] because I don't sleep with my device"* (D5, co-creator). In the case of collaborators and co-creators, this baseline is augmented, annotated, situated, and contextualized through three strategies. The first strategy is relating it to other sources of data (e.g., calendar, menstrual tracking app) and to specific events *"I can check the actual date. Yeah, eight and nine. We had a weekend to, like, train for the competition"* (D17, contributor). The second strategy is to reflect on the experience lived around the data, here several factors come into account, including motivation, perceived effort, training preferences, and stress, *"this represents the* [event] *stress"* (D5, co-creator). The third strategy is to underline variability and temporality and describe the reasons behind it.

> "Those are the days I have the most time and I always train for an hour and a half, a little bit more. The other days of the week it is extremely restricted because I have one hour to train in the morning." (D16, co-creator)

Data is challenged and rectified. For instance, by noticing outliers and challenging these values, *"actually, sometimes it is earlier, like 7:00 a.m."* (D2, collaborator), or by differentiating activity types that were recognized poorly by the device or miscategorized, *"when you select like cardio in Garmin, in my case it can be either like cardio in the gym, like bike or a treadmill, or basketball because I also use it in that setting"* (D4, co-creator).

Moreover, through exploration, curation, and interpretation, data also plays different roles for donors. Here, it should be noted that before uploading the data to the platform, donors had never seen an overview of their data over an extended period of time. It is seen as evidence of their (lack of) consistency and compliance with a specific routine or training program, *"seeing the big picture. It's like I should be so consistent and should be attached to the plan"* (D3, co-creator). It is seen as a medal, commemorating the effort, *"this is also kind of proof to myself like you are doing really well."* (D20, collaborator). It provides satisfaction and it is appreciated. Finally, it is seen as an alarm, signaling disruption or change and motivating donors to care differently for their bodies or adapt their routines, *"that is then a wake-up call for myself. Like, oh, this isn't right. I need to change my habits"* (D8, contributor).

6 DISCUSSION

6.1 Towards Participation and Meaningful Choices in Data Donation

In this paper, we explored how to support different degrees of participation in data donation research. Our research process and activities aimed to promote and support meaningful decisions around participation and (sensitive) information disclosure. Therefore, donors underwent a data donation process involving various choices: whether to participate in the research, how to participate (i.e., as contributors, collaborators, or co-creators), and what (types) of data to donate and for how long. For donors, being able to choose how to participate raised expectations about the process and allowed them to interact with the research(ers) in a way that suited their interests and availability; while being able and invited to choose what (types) of data to donate increased their trust in the research and enabled them to reflect on their boundaries and clearly define them.

For us, as researchers, promoting and supporting these choices introduced certain tensions: more choices might result in less data, different degrees of participation might lead to heterogeneous datasets. Do we want (to deal with) this? For instance, D13 decided not to donate her physical activity data because she considered it sensitive information. This would have been considered a "consent error" [7] in other data donation approaches. For us, it was an opportunity to consider upfront and throughout the process how to harness partial or incomplete datasets. Similarly, promoting different forms of participation resulted in heterogeneous and messy datasets, as these are created from boundaries and exploration and account differently for contextual factors around data. We found how not all donors wish to participate in the same way and how expectations vary depending on the degree of participation. Consequently, the resulting data is not always "complete" nor situated and contextualized. This might also be inconvenient, and opposite to the tidy and clean datasets expected in most data science projects [20, 59]. Especially when donors participate as contributors the data has gaps and outliers, leaving us with questions, assumptions, and possible (mis)interpretations. We resisted the urge to clean up the data and instead chose to question them together with donors where possible and record our questions when not. This invited us to embrace the messiness and heterogeneity of data and determine what data is useful and when. For example, the decontextualized data from contributors can be used to assess the cyclical endurance metric, but this metric would not have been developed without situated and contextualized knowledge from contributors and co-creators.

In contrast to traditional data donation approaches, where donors contribute aggregated anonymous digital-trace data (e.g., [11, 70]), promoting choices and different degrees of participation leads to (dis-)aggregated data that might be openly discussed with collaborators and co-creators. This calls for critical ethical considerations, including transparency (e.g., communicating the benefits and risks associated with data donation and participation [45, 70]), privacy (e.g., explicitly inviting donors to limit the disclosure of sensitive and intimate information [48, 54]), and accountability (e.g., continuously informing donors about the process, progress, and limitations of the research [29, 58]). Additionally, it requires researchers to build and maintain mechanisms to collaborate, communicate, and be accountable to donors throughout the research process beyond what reviewers and ethics committees expect.

Overall, promoting and supporting different choices and degrees of participation invited us to question the characteristics of data collection processes and methods prioritized by researchers. Similarly, previous research has challenged these practices through a critical perspective. Gould [33] illustrates several aspects that render data collection a *consumption experience* such as speed (i.e., being able to collect data quickly) and off-the-shelf tools (e.g., standard questionnaire tools). He invites us to reflect on the trade-offs we make around research design decisions. D'Ignazio and Klein [20] underline the power differentials in traditional data collection practices and invite us to challenge them. For example, by recognizing the people and labor involved in data practices. While Loukissas [42] argues for creating (data) interfaces that cause friction. We echo these perspectives as we argue for a different mindset in data donation and similar practices around the collection and use of personal data for research. Instead of prioritizing speed, could we prioritize building and maintaining a relationship with interested people and communities? Instead of prioritizing the quantity and "heterogeneity" of data could we prioritize meaningful choices that stem from the supporting friction and interaction between people and their data?

6.2 Implications for Policy Makers, Researchers, and Potential Donors

We translate our experience conducting the case study into implications for policy makers, researchers, and potential donors (i.e., individuals interacting with connected products and services that collect digital-trace data).

- **Implications for Policy Makers:** The data donation process described in this paper required donors to exercise their rights to access and data portability. In doing so, they experienced some practical challenges: forgotten login credentials when requests are handled through dedicated platforms, unexpected delays due not opening an email in due time, lengthy waiting times, and receiving data in confusing formats. Above all these, the biggest obstacle remains how data is returned as illustrated by Figure 5 (i.e., unstructured files with no guidance, confusing formats) [2, 9]. Policy makers could develop policies that bridge the access gap: from obtaining a machine-readable copy of our data to being able to understand it and use it.
- **Implications for Researchers:** In this paper, we argue for promoting and supporting different choices and degrees of participation in data donation and similar practices around the collection and use of personal data. This perspective might require for us to change the way we approach research and implement new processes. For instance, our interactions with Human Research Ethics Committees or Institutional Review Boards do not allow for open-ended projects or projects with changing goals [71]. In our case, we submitted our project for review in two instances. First, before the co-creation session. Second, before the data donation process. As a research community, we could develop more dynamic processes that account for the open-endedness that comes from co-creation. Additionally, we should strive to develop practices and safeguards to remain accountable to the individuals who trust us with their data.
- **Implications for Potential Donors:** The data donation process described in this paper illustrated some of the uncertainties and knowledge gaps around personal data collection. Although data is about us, it is often opaque and presented in ways that make it hard to interpret or to benefit from it. As individuals, who interact with connected products and services, we could and should be more aware of the rights around our data and the value it has for product and service providers, and potentially for researchers and ourselves. For instance, requesting and receiving a copy of our data means we could potentially use them.

6.3 Practical Recommendations for Approaching Data Donation

Based on our challenges and experience approaching data donation, we provide the following practical recommendations:

- **Become familiar with practicalities:** The donation of data requires individuals to request and receive their data from data controller(s). There are many ways in which this process can be delayed or go wrong. Additionally, the practicalities of this process could change throughout the project. Similar to van Driel et al. [70], we found that the content and file structures that potential donors requested from the same data controller changed over time. We recommend becoming familiar with this process and being alert to potential changes to support donors when necessary and anticipate challenges in the platforms and systems enabling data donation.

- **Balance trade-offs:** Are more data better data? Are more choices better choices? When defining the relevant sources of data there are choices involved entailing practical trade-offs. In our case, supporting data from various devices (i.e., Garmin and the Apple Health ecosystem) increased the number of people available to donate, but it also increased the technical complexity of the platform and the data donation process. We recommend being aware of these trade-offs and balancing them with the needs and goals of the project.
- **Invite and expect messiness:** Data donation research projects and data needs should be flexible enough to enable various choices from donors, such as: how to participate and what information to disclose. We, as researchers, can gain valuable insights even from messy and incomplete datasets. We recommend anticipating messiness and incompleteness by considering upfront how to promote meaningful choices around information disclosure and how to use the resulting heterogeneous data.

7 LIMITATIONS AND FUTURE CHALLENGES

In this paper, we introduced a case study illustrating how to include different degrees of participation in data donation research processes. Yet, there are limitations to this approach. First, not all individuals can or wish to participate in scoping the research process and goals; which might limit the perspectives included in the early stages of the research – for example, by only including those of highly motivated participants. Second, there are several factors that limit people's willingness and ability to donate personal data. For instance, owning a specific device (e.g., Garmin or Apple Watch), having used the device for a given time (e.g., three or six months), having sufficient digital abilities to request and transfer the data, and having sufficient trust in the research, among others. Third, a group of donors might likely differ from one recruited through different means (e.g., crowdsourcing) or responding to different incentives (e.g., money); which might bias the research process and outcomes. Fourth, we demonstrated how donors initially approach data donation blindly, underlining the importance of supporting exploration and interpretation. To do so, we opted for a timeline data visualization by following guidelines proposed by Pins et al. [55], however, we did not evaluate its effectiveness or accuracy. Fifth, we delayed the initial moment of informed consent until after potential donors were invited to explore their data. However, inviting them to explore and inspect their data does not mean they will. We found that some donors who participated as collaborators and co-creators did not explore their data on the platform as they preferred to do so with us. Does this limit their initial ability to assess and consent to their participation at the time of the transaction? Finally, we explicitly asked donors to (re)evaluate their participation, but none withdrew their consent or data in any stage of the data donation process. Future research should further explore the implementation of dynamic consent practices.

8 CONCLUSION

The aim of this paper was to explore the opportunity to embed data donation in research processes that support and promote various degrees of participation: contributors, collaborators, and co-creators. Therefore, we conducted a data donation case study at the intersection of the menstrual cycle and sports performance where, in collaboration with 20 professional, semi-professional, and amateur athletes, we explored the perceived impact of the menstrual cycle on sports performance and envisioned ways to generate knowledge about it through their (1) menstrual cycle, (2) sleep, (3) heart rate, and (4) physical activity data. We discussed their data donation experience and reflected on the challenges and tensions we faced engaging in a research process aimed at promoting and supporting different choices and degrees of participation. We argue for different mindsets and priorities around data donation and similar practices that entail the collection of personal data for research. We conclude discussing the implications of our findings for policy makers, researchers, and individuals and providing practical recommendations for researchers applying data donation.

CREDIT STATEMENT

Alejandra Gómez Ortega: Conceptualization, Methodology, Formal Analysis, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Software, Project Administration. **Jacky Bourgeois:** Software, Resources. **Gerd Kortuem:** Formal Analysis, Writing - Review & Editing.

ACKNOWLEDGMENTS

We thank all the contributors, collaborators, and co-creators who kindly shared their data, time, and experiences, and steered this project in ways we could not anticipated. We thank the reviewers for their encouraging and constructive feedback that helped strengthen this contribution and Carlo van der Valk for his support with the data donation platform.

REFERENCES

[1] Algorithm Watch. 2021. DataSkop: Investigating YouTube's algorithm during Germany's election campaign. https://algorithmwatch.org/en/dataskop-investigating-youtubes-algorithm-during-germanys-election-campaign/

[2] Fatemeh Alizadeh, Timo Jakobi, Jens Boldt, and Gunnar Stevens. 2019. GDPR-Reality Check on the Right to Access Data. In *Proceedings of Mensch und Computer 2019*. ACM, New York, NY, USA, 811–814. https://doi.org/10.1145/3340764.3344913

[3] Shaowen Bardzell. 2010. Feminist HCI: Taking stock and outlining an agenda for design. *Conference on Human Factors in Computing Systems - Proceedings* 2 (2010), 1301–1310. https://doi.org/10.1145/1753326.1753521

[4] Shaowen Bardzell and Jeffrey Bardzell. 2011. Towards a feminist HCI methodology: Social science, feminism, and HCI. In *Conference on Human Factors in Computing Systems - Proceedings*. 675–684. https://doi.org/10.1145/1978942.1979041

[5] Susanne E. Baumgartner, Sindy R. Sumter, Vladislav Petkevič, and Wisnu Wiradhany. 2022. A Novel iOS Data Donation Approach: Automatic Processing, Compliance, and Reactivity in a Longitudinal Study. *Social Science Computer Review* 0, 0 (2022), 1–17. https://doi.org/10.1177/08944393211071068

[6] Sina Blassnig, Eliza Mitova, Nico Pfiffner, and Michael V. Reiss. 2023. Googling Referendum Campaigns: Analyzing Online Search Patterns Regarding Swiss Direct-Democratic Votes. *Media and Communication* 11, 1 (2023), 19–30. https://doi.org/10.17645/mac.v11i1.6030

[7] Laura Boeschoten, Jef Ausloos, Judith E. Möller, Theo Araujo, and Daniel L. Oberski. 2022. A framework for privacy preserving digital trace data collection through data donation. *Computational Communication Research* 4, 2 (2022), 388–423. https://doi.org/10.5117/ccr2022.2.002.boes

[8] Sander Bogers, Joep Frens, Janne van Kollenburg, Eva Deckers, and Caroline Hummels. 2016. Connected Baby Bottle. In *Proceedings of the 2016 ACM Conference*

on Designing Interactive Systems. ACM, New York, NY, USA, 301–311. https://doi.org/10.1145/2901790.2901855

[9] Alex Bowyer, Jack Holt, Josephine Go Jefferies, Rob Wilson, David Kirk, and Jan David Smeddinck. 2022. Human-GDPR Interaction: Practical Experiences of Accessing Personal Data. In *CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–19. https://doi.org/10.1145/3491102.3501947

[10] Virginia Braun and Victoria Clarke. 2013. *Successful Qualitative Research. A Practical Guide for Beginners*. SAGE Publications Ltd, London. 400 pages.

[11] Johannes Breuer, Zoltán Kmetty, Mario Haim, and Sebastian Stier. 2022. User-centric approaches for collecting Facebook data in the 'post-API age': experiences from two studies and recommendations for future research. *Information, Communication & Society* (7 2022), 1–20. https://doi.org/10.1080/1369118X.2022.2097015

[12] Nadia Campo Woytuk, Marie Louise Juul Søndergaard, Marianela Ciolfi Felice, and Madeline Balaam. 2020. Touching and Being in Touch with the Menstruating Body. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery (ACM), Honolulu, HI, USA, 1–14. https://doi.org/10.1145/3313831.3376471

[13] Mikaeli Anne Carmichael, Rebecca Louise Thomson, Lisa Jane Moran, and Thomas Philip Wycherley. 2021. The impact of menstrual cycle phase on athletes' performance: a narrative review. *International Journal of Environmental Research and Public Health* 18, 4 (2021), 1–24. https://doi.org/10.3390/ijerph18041667

[14] Charlotte L. Clarke, Heather Wilkinson, Julie Watson, Jane Wilcockson, Lindsay Kinnaird, and Toby Williamson. 2018. A Seat Around the Table: Participatory Data Analysis With People Living With Dementia. *Qualitative Health Research* 28, 9 (2018), 1421–1433. https://doi.org/10.1177/1049732318774768

[15] Spinuzzi Clay. 2005. The methodology of participatory design. *Technical Communication* 52, 2 (2005), 163–174.

[16] Tamara Clegg, Daniel M. Greene, Nate Beard, and Jasmine Brunson. 2020. Data Everyday: Data Literacy Practices in a Division I College Sports Context. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376153

[17] Joseph T Costello, Francois Bieuzen, and Chris M Bleakley. 2014. Where are all the female participants in Sport and Exercise Medicine research? *European Journal of Sport Science* 4, 8 (2014), 847–851.

[18] Emma S Cowley, Alyssa A Olenick, Kelly L McNulty, and Emma Z Ross. 2019. "Invisible Sportswomen": The Sex Data Gap in Sport and Exercise Science Research. *Women in Sport and Physical Activity Journal* 29, 2 (2019), 146–151.

[19] Caroline Criado Perez. 2019. *Invisible Women: Data Bias in a World Designed for Men*. Harry N. Abrams.

[20] Catherine D'Ignazio and Lauren F. Klein. 2020. *Data Feminism*. MIT Press. 328 pages. https://doi.org/10.7551/mitpress/11805.001.0001

[21] Catherine D'Ignazio, Rebecca Michelson, Alexis Hope, Josephine Hoy, Jennifer Roberts, and Kate Krontiris. 2020. "The Personal is Political": Hackathons as Feminist Consciousness Raising. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020). https://doi.org/10.1145/3415221

[22] Arielle Duhaime-Ross. 2014. Apple promised an expansive health app, so why can't I track menstruation? https://www.theverge.com/2014/9/25/6844021/apple-promised-an-expansive-health-app-so-why-cant-i-track

[23] Stacey Emmonds, Omar Heyward, and Ben Jones. 2019. The Challenge of Applying and Undertaking Research in Female Sport. *Sports Medicine - Open* 5, 1 (2019). https://doi.org/10.1186/s40798-019-0224-x

[24] Jordan Eschler, Amanda Menking, Sarah Fox, and Uba Backonja. 2019. Defining Menstrual Literacy with the Aim of Evaluating Mobile Menstrual Tracking Applications. *CIN - Computers Informatics Nursing* 37, 12 (2019), 638–646. https://doi.org/10.1097/CIN.0000000000000559

[25] European Parliament. 2018. General Data Protection Regulation (GDPR). https://gdpr.eu/

[26] FITR Woman. 2023. FITR Woman: Track Your Period & Train Smarter. https://www.fitrwoman.com/

[27] Garmin. 2021. Body Battery™ Energy Monitoring. https://www.garmin.com/garmin-technology/health-science/body-battery/

[28] Alejandra Gomez Ortega, Jacky Bourgeois, Wiebke Toussaint Hutiri, and Gerd Kortuem. 2023. Beyond data transactions: a framework for meaningfully informed data donation. *AI and Society* 0123456789 (2023). https://doi.org/10.1007/s00146-023-01755-5

[29] Alejandra Gómez Ortega, Jacky Bourgeois, and Gerd Kortuem. 2022. Reconstructing Intimate Contexts through Data Donation: A Case Study in Menstrual Tracking Technologies. In *Nordic Human-Computer Interaction Conference*. ACM, New York, NY, USA, 1–12. https://doi.org/10.1145/3546155.3546646

[30] Alejandra Gómez Ortega, Jacky Bourgeois, and Gerd Kortuem. 2023. What is Sensitive About (Sensitive) Data ? Characterizing Sensitivity and Intimacy with Google Assistant Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery. https://doi.org/10.1145/3544548.3581164

[31] Alejandra Gómez Ortega, Jacky Bourgeois, and Gerd Kortuem. 2024. Sensitive Data Donation: A Feminist Reframing of Data Practices for Intimate Research Contexts. In *In Proceedings of the 2024 ACM Designing Interactive Systems Conference (DIS '24), July 1–5, 2024, IT University of Copenhagen, Denmark*. Association for Computing Machinery, 1–14. https://doi.org/10.1145/3643834.3661524

[32] Katerina Gorkovenko, Daniel J Burnett, James Thorp, Daniel Richards, and Dave Murray-Rust. 2019. Supporting Real-Time Contextual Inquiry Through Sensor Data Supporting Real-Time Contextual Inquiry Through Sensor Data. In *Ethnographic Praxis in Industry Conference Proceedings*. Edinburgh, UK, 1–29.

[33] Sandy J. J. Gould. 2022. Consumption experiences in the research process. In *CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–17. https://doi.org/10.1145/3491102.3502001

[34] Ruth W. Grant and Jeremy Sugarman. 2004. Ethics in human subjects research: Do incentives matter? *Journal of Medicine and Philosophy* 29, 6 (2004), 717–738. https://doi.org/10.1080/03605310490883046

[35] Ernst Hafen. 2019. Personal Data Cooperatives – A New Data Governance Framework for Data Donations and Precision Health. *The Ethics of Medical Data Donation. Philosophical Studies Series* 137 (2019).

[36] Sarah Homewood. 2018. Reframing Design Problems Within Women's Health. *DRS2018: Catalyst* 2 (2018), 25–28. https://doi.org/10.21606/drs.2018.337

[37] Patrik Hummel, Matthias Braun, and Peter Dabrock. 2019. Data Donations as Exercises of Sovereignty. In *Philosophical Studies Series*. Vol. 137. Springer International Publishing, Cham, 23–54. https://doi.org/10.1007/978-3-030-04363-6{_}3

[38] Kerina H. Jones. 2019. Incongruities and Dilemmas in Data Donation: Juggling Our 1s and 0s. In *Philosophical Studies Series*. Vol. 137. 75–93. https://doi.org/10.1007/978-3-030-04363-6{_}5

[39] Jenny Krutzinna, Mariarosaria Taddeo, and Luciano Floridi. 2019. An Ethical Code for Posthumous Medical Data Donation. In *Philosophical Studies Series*. Vol. 137. 181–195. https://doi.org/10.1007/978-3-030-04363-6{_}12

[40] Albrecht Kurze, Andreas Bischof, Sören Totzauer, Michael Storz, Maximilian Eibl, Margot Brereton, and Arne Berger. 2020. Guess the Data: Data Work to Understand How People Make Sense of and Use Simple Sensor Data from Homes. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376273

[41] Brianna Larsen, Kirstin Morris, Karlee Quinn, Mark Osborne, and Clare Minahan. 2020. Practice does not make perfect: A brief view of athletes' knowledge on the menstrual cycle and oral contraceptives. *Journal of Science and Medicine in Sport* 23, 8 (2020), 690–694. https://doi.org/10.1016/j.jsams.2020.02.003

[42] Yanni Loukissas. 2022. *All Data Are Local: Thinking Critically in a Data-Driven Society*. The MIT Press.

[43] Yuhan Luo, Young-Ho Kim, Bongshin Lee, Naeemul Hassan, and Eun Kyoung Choe. 2021. FoodScrap: Promoting Rich Data Capture and Reflective Food Journaling Through Speech Input. In *Designing Interactive Systems Conference 2021*. ACM, New York, NY, USA, 606–618. https://doi.org/10.1145/3461778.3462074

[44] Tuuli Mattelmäki and Froukje Sleeswijk Visser. 2011. Lost in Co-X: Interpretations of Co-design and Co-creation. In *In Proceedings of IASDR2011, the 4th World Conference on Design Research*, N Roozenburg, LL Chen, and PJ Stappers (Eds.). TU DElft/IASDR, Delft, The Netherlands, 1–12. http://designresearch.fi/courses/ds11/wp-content/uploads/2011/10/Mattelmaki_Lost-in-CoX_fin1.pdf

[45] Matthew S. McCoy, Anita L. Allen, Katharina Kopp, Michelle M. Mello, D. J. Patil, Pilar Ossorio, Steven Joffe, and Ezekiel J. Emanuel. 2023. Ethical Responsibilities for Companies That Process Personal Data. *American Journal of Bioethics* 23, 11 (2023), 11–23. https://doi.org/10.1080/15265161.2023.2209535

[46] Kelly Lee McNulty, Kirsty Jayne Elliott-Sale, Eimear Dolan, Paul Alan Swinton, Paul Ansdell, Stuart Goodall, Kevin Thomas, and Kirsty Marie Hicks. 2020. The Effects of Menstrual Cycle Phase on Exercise Performance in Eumenorrheic Women: A Systematic Review and Meta-Analysis. *Sports Medicine* 50, 10 (2020), 1813–1827. https://doi.org/10.1007/s40279-020-01319-3

[47] Nicole Nguyen. 2022. The Cult of Garmin : Why Athletes Stick With These Smartwatches Over Apple or Samsung. *The Wall Street Journal* (2022). https://www.wsj.com/articles/the-cult-of-garmin-why-athletes-stick-with-these-smartwatches-over-apple-or-samsung-11667088984

[48] Helen Nissenbaum. 2004. Privacy as Contextual Integrity. *Washington Law Review* 79, 1 (2004), 101–139.

[49] Renee Noortman, Peter Lovei, Mathias Funk, Eva Deckers, Stephan Wensveen, and Berry Eggen. 2022. Breaking up data-enabled design: expanding and scaling up for the clinical context. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM* 36 (2022). https://doi.org/10.1017/S0890060421000433

[50] Jakob Ohme and Theo Araujo. 2022. Digital data donations: A quest for best practices. *Patterns* 3, 4 (4 2022), 100467. https://doi.org/10.1016/j.patter.2022.100467

[51] Jakob Ohme, Theo Araujo, Laura Boeschoten, Deen Freelon, Nilam Ram, Byron B. Reeves, and Thomas N. Robinson. 2023. Digital Trace Data Collection for Social Media Effects Research: APIs, Data Donation, and (Screen) Tracking. *Communication Methods and Measures* 00, 00 (2023), 1–18. https://doi.org/10.1080/19312458.2023.2181319

[52] Jakob Ohme, Theo Araujo, Claes H. de Vreese, and Jessica Taylor Piotrowski. 2021. Mobile data donations: Assessing self-report accuracy and sample biases with the iOS Screen Time function. *Mobile Media and Communication* 9, 2 (2021), 293–313. https://doi.org/10.1177/2050157920959106

[53] Jessica A. Pater, Rachel Pfafman, and Amanda Coupe. 2021. Standardizing reporting of participant compensation in hci: A systematic literature review and recommendations for the field. *Conference on Human Factors in Computing Systems - Proceedings* (2021). https://doi.org/10.1145/3411764.3445734

[54] Sandra Petronio. 2002. *Boundaries of privacy: Dialectics of disclosure.* State University of New York Press, Albany.

[55] Dominik Pins, Timo Jakobi, Alexander Boden, Fatemeh Alizadeh, and Volker Wulf. 2021. Alexa, We Need to Talk: A Data Literacy Approach on Voice Assistants. In *Designing Interactive Systems Conference 2021.* ACM, New York, NY, USA, 495–507. https://doi.org/10.1145/3461778.3462001

[56] C. K. Prahalad and Venkat Ramaswamy. 2004. Co-creation experiences: The next practice in value creation. *Journal of Interactive Marketing* 18, 3 (2004), 5–14. https://doi.org/10.1002/dir.20015

[57] Barbara Prainsack. 2019. Data Donation: How to Resist the iLeviathan. In *Philosophical Studies Series*. Vol. 137. 9–22. https://doi.org/10.1007/978-3-030-04363-6{\protect\T1\textdollar}{_}2

[58] Barbara Prainsack and Seliem El-Sayed. 2023. Beyond Individual Rights: How Data Solidarity Gives People Meaningful Control over Data. *American Journal of Bioethics* 23, 11 (2023), 36–39. https://doi.org/10.1080/15265161.2023.2256267

[59] KATIE RAWSON and TREVOR MUÑOZ. 2019. Against Cleaning. *Debates in the Digital Humanities 2019* (2019), 279–292. https://doi.org/10.5749/j.ctvg251hk.26

[60] Afsaneh Razi, Ashwaq Alsoubai, Seunghyun Kim, Nurun Naher, Shiza Ali, Gianluca Stringhini, Munmun De Choudhury, and Pamela J. Wisniewski. 2022. Instagram Data Donation: A Case Study on Collecting Ecologically Valid Social Media Data for the Purpose of Adolescent Online Risk Detection. In *Conference on Human Factors in Computing Systems - Proceedings*. 1–18. https://doi.org/10.1145/3491101.3503569

[61] Janet C. Read, Daniel Fitton, and Matthew Horton. 2014. Giving ideas an equal chance: Inclusion and representation in participatory design with children. *ACM International Conference Proceeding Series* (2014), 105–114. https://doi.org/10.1145/2593968.2593986

[62] Lara Reime, Vasiliki Tsaknaki, and Marisa Leavitt Cohn. 2023. Walking Through Normativities of Reproductive Bodies: A Method for Critical Analysis of Tracking Applications. *Conference on Human Factors in Computing Systems - Proceedings* (2023). https://doi.org/10.1145/3544548.3581450

[63] Robert Koch Institut. 2020. Corona-Datenspende. https://corona-datenspende.de/science/en/

[64] E Sanders and PJ Stappers. 2008. Co-creation and the new landscapes of design. *CoDesign* 4, 1 (2008), 5–18. https://doi.org/10.1080/15710880701875068

[65] Jennifer L. Shirk, Heidi L. Ballard, Candie C. Wilderman, Tina Phillips, Andrea Wiggins, Rebecca Jordan, Ellen McCallie, Matthew Minarchek, Bruce V. Lewenstein, Marianne E. Krasny, and Rick Bonney. 2012. Public participation in scientific research: A framework for deliberate design. *Ecology and Society* 17, 2 (2012). https://doi.org/10.5751/ES-04705-170229

[66] Lauren Reichart Smith and Jimmy Sanderson. 2015. I'm Going to Instagram It! An Analysis of Athlete Self-Presentation on Instagram. *Journal of Broadcasting and Electronic Media* 59, 2 (2015), 342–358. https://doi.org/10.1080/08838151.2015.1029125

[67] Marie Louise Juul Søndergaard. 2020. Troubling design: A design program for designing with women's health. *ACM Transactions on Computer-Human Interaction* 27, 4 (2020). https://doi.org/10.1145/3397199

[68] Veronika Strotbaum, Monika Pobiruchin, Björn Schreiweis, Martin Wiesner, and Brigitte Strahwald. 2019. Your data is gold – Data donation for better healthcare? *it - Information Technology* 61, 5-6 (10 2019), 219–229. https://doi.org/10.1515/itit-2019-0024

[69] Jakob Tholander and Stina Nylander. 2015. Snot, sweat, pain, mud, and snow - performance and experience in the use of sports watches. *Conference on Human Factors in Computing Systems - Proceedings* 2015-April (2015), 2913–2922. https://doi.org/10.1145/2702123.2702482

[70] Irene I. van Driel, Anastasia Giachanou, J. Loes Pouwels, Laura Boeschoten, Ine Beyens, and Patti M. Valkenburg. 2022. Promises and Pitfalls of Social Media Data Donations. *Communication Methods and Measures* 16, 4 (2022), 266–282. https://doi.org/10.1080/19312458.2022.2109608

[71] Janne van Kollenburg and Sander Bogers. 2019. *Data-enabled design : a situated design approach that uses data as creative material when designing for intelligent ecosystems.* Technical Report 2019. Eindhoven University of Technology, Eindhoven, The Netherlands. https://research.tue.nl/en/publications/data-enabled-design-a-situated-design-approach-that-uses-data-as-

[72] Saskia J. Verhoef, Merel C. Wielink, Edwin A. Achterberg, Marlies Y. Bongers, and Simone M.T.A. Goossens. 2021. Absence of menstruation in female athletes: why they do not seek help. *BMC Sports Science, Medicine and Rehabilitation* 13, 1 (2021), 1–11. https://doi.org/10.1186/s13102-021-00372-3

[73] Laia Turmo Vidal, Hui Zhu, Annika Waern, and Elena Márquez Segura. 2021. The design space ofwearables for sports and fitness practices. *Conference on Human Factors in Computing Systems - Proceedings* (2021). https://doi.org/10.1145/3411764.3445700

[74] Nadia Campo Woytuk, Joo Young Park, Jan Maslik, Marianela Ciolfi Felice, and Madeline Balaam. 2023. Tactful Feminist Sensing : Designing for Touching Vaginal Fluids. In *In Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS '23).* 2642–2656. https://doi.org/10.1145/3563657.3595966

[75] Savvas Zannettou, Olivia-Nemes Nemeth, Oshrat Ayalon, Angelica Goetzen, Krishna P. Gummadi, Elissa M. Redmiles, and Franziska Roesner. 2023. Leveraging Rights of Data Subjects for Social Media Analysis: Studying TikTok via Data Donations. *ArXiv* (2023). http://arxiv.org/abs/2301.04945