Title
Opportunities and Challenges for Long-Term Tracking

ISBN
9783030673215

Authors
Epstein, Daniel A
Eslambolchilar, Parisa
Kay, Judy
et al.

Publication Date
2021

DOI
10.1007/978-3-030-67322-2_9

Peer reviewed

Opportunities and Challenges for Long-Term Tracking

Daniel A. Epstein, Parisa Eslambolchilar, Judy Kay, Jochen Meyer, and Sean A. Munson

Abstract As self-tracking has evolved from a niche practice to a mass-market phenomenon, it has become possible to track a broad range of activities and vital parameters over years and decades. This creates both new opportunities for long term research and also illustrates some challenges associated with longitudinal research. We establish characteristics of very long-term tracking, based on previous work from diverse areas of Ubicomp, HCI, and health informatics. We identify differences between long- and short-term tracking, and discuss consequences on the tracking process. A model for long-term tracking integrates the specific characteristics and facilitates identifying viewpoints of tracking. Finally, a research agenda suggests major topics for future work, including respecting gaps in data and incorporating secondary data sources.

Keywords Self-tracking · Long-term · Personal informatics · Physical activity

D. A. Epstein (✉)
University of California Irvine, Irvine, USA
e-mail: epstein@ics.uci.edu

P. Eslambolchilar
Cardiff University, Cardiff, UK
e-mail: EslambolchilarP@cardiff.ac.uk

J. Kay
University of Sydney, Sydney, Australia
e-mail: judy.kay@sydney.edu.au

J. Meyer
OFFIS Institute for Information Technology, Oldenburg, Germany
e-mail: meyer@offis.de

S. A. Munson
University of Washington, Seattle, USA
e-mail: smunson@uw.edu

E. Karapanos et al. (eds.), *Advances in Longitudinal HCI Research*,
Human–Computer Interaction Series, https://doi.org/10.1007/978-3-030-67322-2_9

1 Introduction

New sensors, miniaturization, the ubiquity of smartphones, networking and the Internet of things present designers with a plethora of new applications and systems that promise to provide people with that data to support and improve their personal health, well-being, and fitness, and, for researchers, opportunities to understand health and well-being longitudinally. Many research and commercial systems aim to promote *personal tracking*, or monitoring of one's habits for self-understanding and self-improvement [32]. Substantial work in HCI has demonstrated benefits of *short-term tracking*, where people collect data about their habits and reflect on them for a couple of weeks or months. Short-term tracking interventions have been designed and evaluated for improving physical activity [31, 46], eating habits [59], workplace productivity [27], and other domains. There is also some growing understanding of the ways people can harness long-term tracking data for self-understanding and self-improvement.

Personal tracking tools now support collecting more and more detailed data about ourselves, with varying levels of effort. Wearable devices or smartphones can passively monitor physical activity as total daily steps, as steps per minute, or heart rate based as exertion. Ambient and interaction-free "install-once-and-forget" devices such as Withings Aura or Beddit can monitor sleep as time in bed and time asleep, but may over-promise in other measures they offer [33]. Nutrition can be monitored manually using either lightweight diaries, detailed database approaches, or photo-based tracking, giving insights into general dietary behavior, calorie consumption, or macro- and micro-nutrient intake. Substantial research continues to explore how sensing in wearable devices can passively automate collection of data (e.g., [7]). In spite of these technological advances, we acknowledge that commercial tracking tools often do not meet standards for clinical accuracy, and the resulting data should not be used to support inferences or decisions it cannot [38].

Use of tracking technology has moved from a promising novelty to a long-term phenomenon. For many, tracking happens not just as an exceptional activity for a limited period of a few weeks or a couple of months. Rather, it is a regular part of life, covering years or decades, or even life-long. We are only slowly starting to understand that there are considerable opportunities from such long-term tracking [35, 36]. For example, the availability of long-term tracked personal data can enable identifying and reflecting on long term trends in behavior, early detection of health risks or diseases, monitoring progress against a long term target, giving a lifelong health support, or enabling repeated N-of-1 experiments.

As tracking technology and the practice of long-term tracking become more ubiquitous, opportunities for studying and leveraging long-term tracking in research increase. Studying how tracking tools align with people's lived experiences can lead to recommendations for improving the design of tracking tools, addressing key barriers or challenges. Analyzing the data people collect about themselves can also be used to understand people's practices, such as understanding exercise or nutritional trends [4], following the progression of illnesses, or longitudinal surveillance of health conditions [3].

In this chapter, we first introduce case studies from our personal experiences and prior research which characterize challenges and opportunities for collecting and analyzing long-term tracking data. We then describe conceptual models and theories on people's self-tracking practices. We contrast these models with a model we created to highlight the important feedback loops in long-term tracking and to describe the individuals and entities producing and using data. We then discuss how our model addresses issues identified in the case studies and articulate dimensions that are helpful in teasing apart the ways in which long-term tracking differs from short-term tracking. Finally, we discuss potential future directions for long-term tracking research and considerations when conducting that research.

2 Case Studies in Long-Term Tracking

We present case studies from our prior research and inspired by our personal experiences to characterize challenges in conducting long-term tracking, and therefore in conducting research leveraging long-term tracking. The stories of Jasmin and Joe illustrate difficulties maintaining a long-term record of physical activity. Designing around adherence and the point of lapsing illustrate potential opportunities for designs to make data from long-term tracking more useful for reflection over behavior.

2.1 Jasmin: A Story of Multiple Trackers

Jasmin is a healthy active woman in her forties. Her goal is maintaining a healthy and active lifestyle as her daily job is sedentary and involves sitting for long hours in a confined space. As a result, she invested in an activity tracking watch to monitor her daily activity patterns in 2014. She chose a tracking watch that allowed her to log her sitting time, walking/cycling steps/rides and altitude (for climbing) automatically. Over the course of three months, she collected enough data to build a good picture of her daily/weekly activities via the watch's dashboard. For example, she noticed that some days she was more sedentary than others, and the weather and work deadlines played an important role in her decisions to not cycle or walk to work and not climb. In the following three to six months, Jasmin considered alternatives to increasing activity, e.g., exercising more on weekends and purchased a home trampoline so when the weather was bad, she could exercise at home. Almost one year after buying her first activity tracker, Jasmin felt that she has finally got into a routine that worked for her lifestyle and meant she could maintain her health.

In 2016, while talking to a friend, she learned that it is important to include aerobic exercises in her weekly activities [43]. Her friend recommended buying a watch with heart rate sensor and to start her training in a low heart rate zone and then to include

Fig. 1 Left—Jasmin is following a 5 k running program in 2016. Right—Jasmin is wearing two watches made by two different activity tracking manufacturers

interval training three times a week. Unfortunately, the recommended watch was made by a different manufacturer. For Jasmin, this meant wearing two watches at the same time because she wanted to track with her old tracker while monitoring her heart rate with another (see Fig. 1). The dashboards for both trackers did not talk to each other; she could not close her account down with the old tracker, download her data of 18 months and import it into the new watch's dashboard; she could not find a third-party platform to merge her data either.

Keeping and accessing her old data was important to her for several reasons. Jasmin wanted to track her daily, weekly, monthly and possibly yearly trends and changes to her overall fitness picture, reflect on past activities, and use that data to help her adapt to new situations. Therefore, she had no choice but to wear two watches when she ran. Despite the practical challenges of wearing two watches at the same time, Jasmin found the heart rate training program linked to the watch manufacturer helpful. For example, the watch prompted her if she was running too hard or too slowly or if she had missed a run on a scheduled day. She subsequently participated in a 5 k race in 2016 and completed a 10 k race in 2017 and a half-marathon in 2018. She could not have achieved these without her long-term training program and monitoring her progress on the watch/ dashboard.

Jasmin's old tracker approached the end of its shelf life in 2017, with her battery no longer lasting a full day. The warranty had expired, and the manufacturer was not able to replace the battery or the watch with the same model as tracker had been discontinued. Heartbroken and disappointed with the loss of her first activity companion, she replaced her second tracker with a more advanced model from same manufacturer so she could access her early running training days and benefit from other sensors on the new watch for other activities, e.g., climbing. Jasmin hopes that one day she can download her data from her first tracking device server and combine them with her other data.

For researchers looking to study or design for long-term tracking, Jasmin's story highlights some challenges for long-term tracking: (1) persistent data access over a long period of time on one platform is nearly impossible, (2) frequently changing goals are not necessarily compatible with one device hence multiple devices may be needed, (3) emotional bonds with devices [26] can influence device choice and abandonment, and (4) "old" data can be wished for years later, taking extreme measures to preserve it and seeking out ways to integrate it with new practices or simply to reminisce.

2.2 Joe: 9 Years of Self Tracking

Joe is a healthy man whose interest in self-tracking arose through both an interest in new technologies and curiosity about his personal health. Joe's general goal is to maintain a healthy and active lifestyle, though this is secondary to his curiosity about the technologies. Joe chose a set of consumer-grade, mass-market products to cover a broad range of activities and vital parameters. As his tracking is primarily incidental, without a specific goal, he sought out devices which collected data passively and required as little additional interaction as possible. After trying many options, he ended up using a stable setup of an activity tracker for daily activity, a sports watch for workouts, an interaction-less, ambient sleep monitor, and a networked body scale. Joe also uses a social network to manually "check in" to places he is visiting, including gyms and other sports facilities.

As of this writing, Joe has collected data for a total of 9 years. Joe describes himself as a "power-user," tracking consistently every day when possible. Therefore data about his daily physical activity, workouts, and body composition is mostly complete. The sleep data is susceptible to errors and gaps due to the need for physical re-adjustment of the sensor after a couple of months, which Joe occasionally missed due to the lack of direct feedback about the measurements. On the other hand, the time of stepping on the scale in the morning proved to be a reasonable indicator for wake-up times under routine circumstances.

Joe's tracking behavior influenced the data in several ways. Wearing both the sports watch and the activity tracker during workouts results in duplicate data; this duplication needs to be taken into account when processing and aggregating the data. Joe's very consistent tracking behavior makes deviations from the routine important information. For example, Joe was able to treat not stepping on the scale as a strong indicator of a time that he was not at home. Due to the intentional lack of manual interaction with some devices, technical failures occasionally went unnoticed, resulting in some incomplete or incorrect data (Fig. 2).

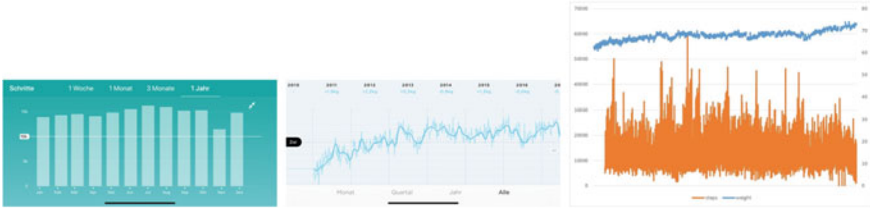

Fig. 2 Trying to make sense of Joe's data: The native apps show the data, steps (left) for one year only, weight (middle) for longer periods. Visualizing both in Excel for the whole period is possible (right), but difficult to comprehend

Joe's experience of tracking points to challenges around maintaining a persistent data record and reflecting on that record amidst so much data from disparate sources. Even as a self-described "power user", Joe's data has gaps, such as not realizing that passively recording devices have stopped syncing [19]. His duplicate data sources may help mitigate this, but require effort to aggregate and analyze. On the other hand, gaps may also tell a story, such as likely absence from home when not stepping on the scale in the morning. Data may well have a secondary, not originally intended use, such as the time of stepping on a scale as a proxy for wake-up time.

For a researcher, Joe's case shows four relevant insights: (1) The choice of devices is very personal, and two different people will probably choose two different set-ups. Even if they chose to use the same devices, their routines for when and how they use them would likely differ. Researchers therefore have to deal with heterogeneous data from disparate sources. (2) Gaps in the data are inevitable; they should not just be considered a normal part of the data, but they may also tell a story on their own. (3) Data will be imprecise or unreliable, but sometimes data may also have some unforeseen value. (4) In spite of all the challenges, the data may provide insights into a person's life that are worth being uncovered and made use of.

2.3 *Interpreting Longitudinal Tracker Data in the Real World: Missing Data, Multiple Interpretations, Human-Machine Collaborative Interpretation*

This case study is based on IStuckWithIt [55, 56], an interface onto long-term data from a wearable activity tracker. There are three key principles that underpin its design. We now describe these, both as they apply for this interface and in terms of aspects that are relevant to managing longitudinal data of many types in the real world.

First, long-term data from a wearable activity tracker is typically incomplete because people do not wear the tracker all day, every day. This is important for interpreting the data. For example, if a person wore their tracker for 16 h in a day, that is likely to give a quite accurate measure of their total step count. But if they only wore the device for 2 h, the step count recorded may be a gross under-estimate of their true activity level. We introduced the term *adherence* [57] to describe this. Intuitively, 100% adherence means that the person used the device, wearing it, with the device operational and with sufficient power, so that it can track all their steps in the day. We explored several ways to define adherence. Essentially, these are based on the broad idea of defining a *valid day*, one with high enough adherence for the step count to be meaningful. To explore the impact of some of the measures in the published literature, we analyzed 12 datasets from diverse classes of people, including those who chose to volunteer their tracker data for analysis, people who had the trackers as part of medical interventions, and participants in a public health study of university students. These datasets had a total of 753 users and over 77,000 days with any data, as well as 73,000 interspersed days with no data. The choice of adherence definition had different effects on step counts for different datasets. The dataset with the largest difference had a low of 6952 and a high of 9423. This is a clinically significant difference. The core message is that the interpretation of data from long-term sensing based on wearable devices requires careful consideration of adherence.

A second challenge is that the data are not homogeneous over long periods of time, as Jasmin's case study points out. Our evaluations of IStuckWithIt were restricted to data from Fitbit activity trackers. These have been available since 2009, with some early adopters having over 10 years of data, covering several models of the tracker.

One other challenge that we tackled in the IStuckWithIt project was the need to support people in making rich and flexible interpretations of their own data. Once the analysis of long-term data has taken account of the challenges above, it is important to create mechanisms for human-in-the-loop interpretation. To do this IStuckWithIt can be seen as an example of an interface that offers some flexibility in the choices of interpretation available as well as scaffolding to help a person "see" the aspects that are important and that can draw on that individual's own knowledge.

Figure 3 shows an example IStuckWithIt screenshot for a hypothetical user we will call Alice. The label (A) indicates that Alice has selected the steps view of her data that comes from a Fitbit. Other views of the same underlying data can show her activity in terms of the number of moderately active minutes she had each day, very active minutes per day and distance.

The area marked (B) shows her activity in the first part of 2014. The cells are bright blue on days she met her target of 10,000 steps. The light blue cells are for days she was below the target but was above 5000 steps and white indicates days with at least 1 step recorded but less than 5000. In the months of February and March, she has quite a lot of blue cells.

A notable feature of the IStuckWithIt interface is the careful handling of days with no data. Days where there is no data are gray, as can be seen at (C). To help Alice interpret her data, the visualization also communicates adherence. (D) marks the display of the average hours she wore the tracker per day. So, for example, in February, she consistently averaged more than 10 h a day of wear. However, in March she had weeks where she averaged less than 6 h a day of wear. This means that the

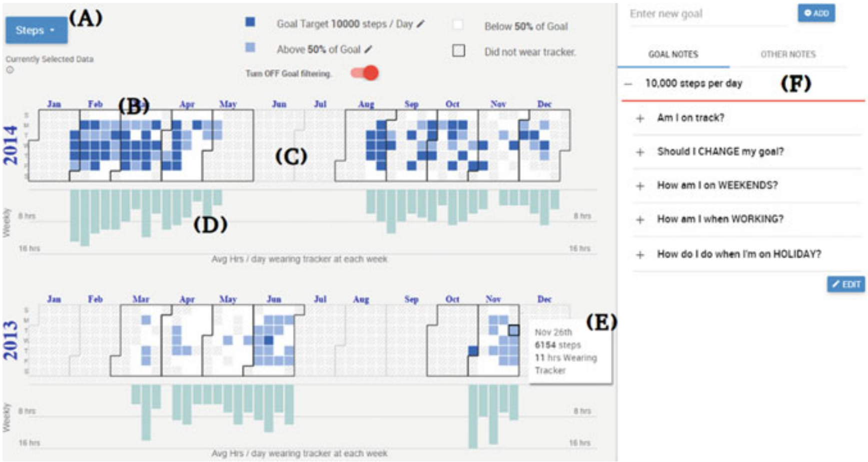

Fig. 3 Example screenshot of IStuckWithIt for a user, Alice

counts visible here are likely to be under-estimates of her actual level of activity. A user can see the precise information for any cell by mousing over it, as shown at (E).

We conducted a study of IStuckWithIt to gain insights about the ways that long-term Fitbit trackers had been using their data and whether they could gain new insights from IStuckWithIt as described above [55]. This study recruited 21 people (7 women), who had an average of 23 months of Fitbit data. Many of these participants were committed to tracking to maintain their level of activity and were in the maintenance phase of the Transtheoretical Model of Behavior Change [45]. This was expected, given we recruited people who had long-term data on the promise of new ways to see it. In terms of broader research on long-term data, these participants have much to offer in terms of their motivations for collecting it, their experiences in doing so, their uses of it, and their insights about the data and how it enabled them to harness it to serve their needs.

The interview asked people about how often they looked at their step count. If needed, this was followed by more detailed probing in terms of the timescales shown in Fig. 4. Despite having long-term data, few of these participants made use of the longer-term data. They indicated that this was too hard to do.

All participants made discoveries about themselves, in terms of their wearing behavior, such as reflecting on what caused them to not have data at certain times, or understanding how many hours a day they tended to wear it. For example, one participant noted from the visual representation: *"That's when I lost it, at end of July, that's why also there is a gap would make sense in that case. I think the gap really affected me, I got out of habit."*. The insights from these dedicated, long-term activity trackers highlight that awareness of their adherence behavior needs to be considered in designing research around such long-term data.

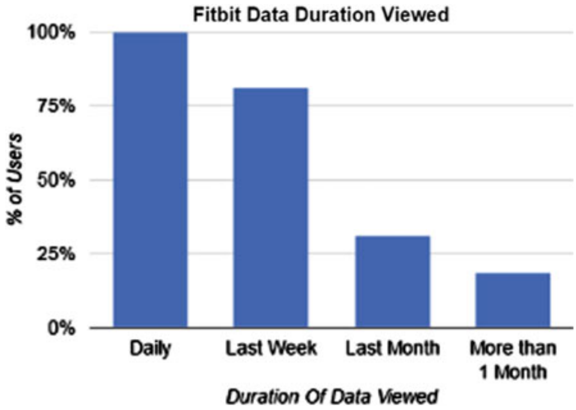

Fig. 4 How frequently participants viewed their data

Most participants made discoveries from their long-term data, such as the influence of their environment on their activity levels (e.g., living in a city vs. a more rural area) or changes in their activity levels because of tracking. Some participants recognized their vacations as the reason for a period with high steps counts. Some consulted their calendars to figure out why step count deviations were notable—this represents a flexible, on-the-fly integration of another long-term data source by the user. Since we did not ask people to provide that data, the control of these separate data stores remained in the hands of the users. In planning studies, researchers should consider what other data sources could triangulate their primary data sources, facilitating answering their research questions, and whether those data can be collected while also balancing participating burden and privacy.

After each participant had explored the main IStuckWithIt interface, the interviewer revealed the scaffolding labeled (F) in Fig. 3. This scaffolding helped them gain new insights [56]. For example, before use of IStuckWithIt, most participants had clearly not considered whether their activity levels changed between weekdays and weekends (even though they stated that they looked at their data each day and we had asked them about this in the earlier interview, potentially priming them to be aware of this). Public health researchers have established that most people are less active on weekends, and consider it important to account for this when measuring activity levels. Even with prompting, 7 participants said they had no idea and could not make an estimate. Of the 6 who thought they were more active on weekdays, three were much more active (by 18%, 49% and 70%), 1 much less so (−90%) and two were about the same (both 2% more). A similarly diverse picture appeared for the 6 people who thought they were less active on weekdays. Once the scaffolding in IStuckWithIt was revealed, 10 participants made new insights about themselves (8 about weekday vs. weekend wear), 2 about workdays versus others, 4 about the impact of holidays and 5 about rethinking the goals). Some of the other participants already had clear goals and intense use of the tracker and did not need this scaffold; they suggested the interface should personalize the scaffolds. For researchers using long-term tracking data to elicit memories and experiences in studies, this highlights that people typically need scaffolding to build self-awareness of key aspects of their data. Our work involved people who were rather dedicated trackers and some had very clear goals; for data from broader user populations, one would expect this is an even stronger factor.

2.4 *Designing to Surface Trends at the Point of Lapsing*

Lapsing in the act of tracking is a well-known and studied phenomenon, as surfaced in Joe's personal experiences and a history of HCI work [9, 16, 19, 30, 52]. Together with colleagues and published previously [18], we have explored how tracking technology can treat the point of lapsing as an opportunity for self-reflection. As a case study, we explored designs for tracking physical activity, specifically collected by the Fitbit device.

Our technique presented people who had lapsed with visual representations of their data combined with captions. We used *visual cuts*, an approach we had developed previously [17], to surface trends which answer questions people often have about their data, but tools typically do not answer. Cuts typically focus on longer-term trends rather than daily or weekly logs, aligning with people's desire to reflect over their behavior rather than review recent behavior. For example, cuts highlighted when throughout the day people tended to have their activity, or a timeline of people's average activity grouped by month and year that they tracked. Figure 5 shows two cuts we designed.

We paired each visual cut with a *framing technique* derived from taxonomies and strategies for designing persuasive technology and facilitating behavior change [22, 37]. For example, framing techniques drew attention to circumstances where a person was particularly active or an opportunity for improvement, prompting them to consider what prevents them from walking more or comparing their performance to others.

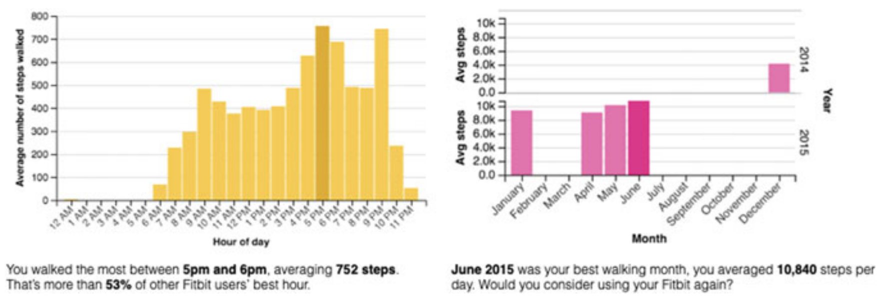

Fig. 5 Two visual cuts and framing captions we created to help people reflect on their data after a lapse in tracking [18]

In an experimental survey, we asked 141 people to rate a series of cuts paired with framing techniques according to how informative and appropriate they found them and describe what they thought of the visualization in a short sentence. We found that cut preference varied by use pattern. Participants who had tracked for a short amount of time (3 months or less) prior to lapsing tended to prefer cuts which aggregated their data by hour or day (e.g., Fig. 5 Left), whereas participants who had tracked for longer preferred aggregations which highlighted their long use (e.g., Fig. 5 Right). Participants with more long-term use described that they had already learned their daily and weekly activity trends from having worn their Fitbit and reflected on their data for a long time. Participant's preference toward framing techniques tended to align with their perspective on whether or not they wanted to return to tracking in the near future. Participants who felt they had learned enough from tracking for the time being preferred framing techniques which surfaced the times they were most active, while those looking to return appreciated framing techniques which nudged them in that direction.

We imagine that in the future, designs can tailor such a presentation to match people's experiences and perspectives. Our study showed that we can leverage properties about people's data to infer what they might find interesting, but designs might benefit from explicitly asking people for their perspective on their tracking experience. For example, the left image in Fig. 6 emphasizes high activity days drawn from a long-term tracking history, while the right highlights the day of the week a tracker averaged the most steps, an approach that can be effective with even a relatively short tracking history.

We anticipate it is relatively easy to detect whether a person has lapsed in tracking. Standalone devices or apps on devices could stop syncing with the cloud servers or local backends where data is stored, or a recent sync will return no data. It is also plausible to detect whether a person has recently reviewed their tracked data by opening the app which collects the data [23] or glancing at the device [24]. Once a

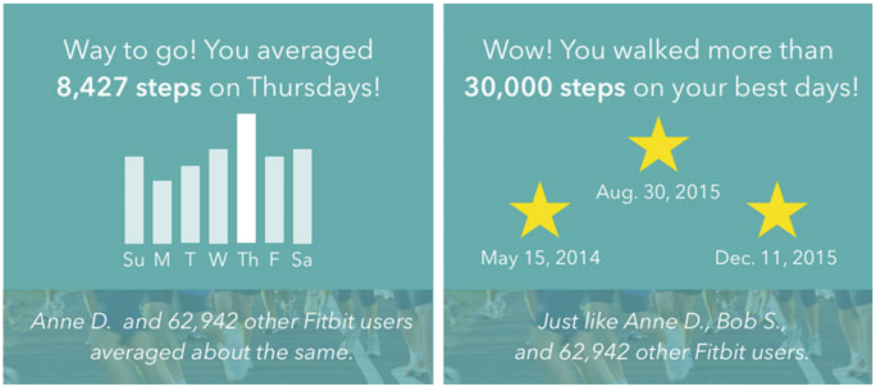

Fig. 6 Designs can surface different information for lapsed trackers who do and do not want to return to tracking

lapse of a reasonable duration is detected, mobile notifications or emails could be used to prompt people to reflect over their data. More difficult, however, is inferring *why* someone has lapsed and sending appropriate messages. For many lapses, frequent prompts or notifications could be overwhelming or annoying. Further study could yield insight into the opportune time to send such a message and further understand how to present it.

This line of research surfaces opportunities for further research on *promoting* long-term self-tracking by identifying and designing interventions for when a person is beginning to stop tracking. We reiterate that lapses in tracking should be expected. There is value in research examining both how designs can encourage people to re-engage in tracking and how designs can provide utility after a person has decided to abandon tracking. It also should be a consideration for studies that seek to use tracking for longitudinal research or public health surveillance. Sustaining engagement in such studies is challenging [29, 44]. While individuals may initially be motivated to participate in these studies to gain personal insights from tracking, they may lose interest or gain the insights they sought, and subsequently chose to lapse even though researchers' goals have not yet been achieved.

3 Theoretical Perspectives on Use of Self-tracking Technology

Aspects of people's experiences collecting long-term data, like Jasmin and Joe, have been characterized in theoretical frameworks. Because these frameworks describe how and why people use tracking technology, building on them is important to studying and supporting long-term tracking. Here, we review key models informing HCI research into tracking.

3.1 Conceptual Models of Personal Tracking Use

Early understanding of how people use tracking technology focused on how collecting data could support linear progress toward a singular goal or decision, such as becoming more physically active or more productive. Inspired by Prochaska & Velicer's Transtheoretical Model of Behavior Change [45] which is very widely used, Li et al. develop a stage-based model describing people's use of tracking technology [32] as an ideally linear progression where people first prepare and collect data, then integrate and reflect on it, and ultimately act on their findings and improve their habits. People often iterate to track new dimensions, and encounter barriers which impact their progress.

Many early adopters of personal tracking systems were professionals in technology-related fields, such as software development, information technology, and data analytics [6, 32]. Participants in the study that informed Li et al.'s model had similar professions. Many were hobbyists in the Quantified Self movement, a group primarily made up of scientists and engineers who sought to build self-knowledge through the collection of numbers about their behaviors [62].

As mobile phones and wearable devices enabled technology with tracking capabilities to further pervade society, who is tracking data and how people collect and engage with data has changed. Though experts and data analysts continue to self-track, today people track for more diverse reasons.

Rooksby et al. characterize people's use of personal tracking tools as "lived informatics", emphasizing that people often do not track with a goal of action, often use multiple tools simultaneously, and are sometimes more socially motivated to track than personally motivated [47]. Epstein et al. draw from this notion to develop the Lived Informatics Model, which characterizes how people use personal tracking tools in their everyday lives [19]. The Lived Informatics Model suggests a more cyclic tracking process where people's varied goals inform the tools they select, collecting and reflecting on data happens simultaneously, and lapses and resuming tracking are frequent occurrences.

The Lived Informatics Model points out that tracking tools often fail to account for the realities of everyday life. People often want or need to migrate between devices or apps as life progresses and as their goals and needs change. In addition to maintaining a continuous record of data to allow people to reflect on long-term trends in their data when possible, tracking tools should account for curiosity-driven goals evolving to self-improvement goals, or from self-improvement to self-maintenance. Tracking over years, compared to weeks or months, particularly highlights a need for designs to account for and address these challenges.

3.2 Modeling Relationships Among Stakeholders and Data in Long-Term Self-Tracking

Long-term tracking introduces complex inter-dependencies between stakeholders, data, users, and applications. We identify five entities: the **primary user** who is collecting data and who has a long-term goal, the **tracking devices and sources** used by the primary user to collect data, the underlying **data** itself coming out of the devices and sources, the **applications** which process and present the data to the user, and finally a **secondary user** who may have reason to access the primary user's data.

Going through the long-term self-tracking loop, we can identify two extreme viewpoints: *purposeful tracking*, going clockwise in the direction of the requirements, and *incidental tracking*, going counter-clockwise following the logical data flow. The purpose of tracking tends to define the tracking behavior the user requires.

Purposeful tracking is the point of view taken in the design of most of today's tracking-based apps. Tracking is often driven by a need for a certain support or service, such as initiating weight loss with a persuasive app or when a person with

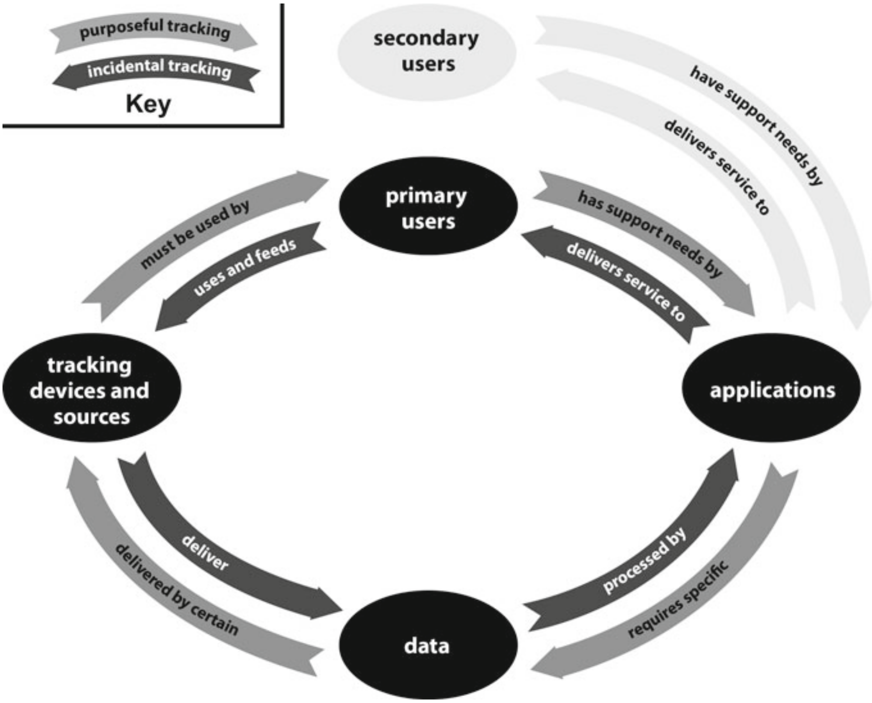

Fig. 7 Long-term self-tracking feedback loops. The outer loop shows purposeful tracking driven needs of the users. The inner loop shows flows in the opposite direction as happens when the user's tracking is an incidental side-effect of their technology use and behaviors

diabetes needs to monitor their blood insulin level. Applications that such purposeful users need demand that the user does a certain amount of work to ensure that there is enough quality data collected for the app to be effective. Therefore the (typically primary) user must use certain tracking devices or dedicated apps to collect that data (e.g., an activity tracker, a scale, a nutrition diary for weight loss case; a glucometer for diabetes).

In contrast, in incidental tracking shown in the inner loop of Fig. 7 and then separately at the right in Fig. 7, the user does not have a specific need. Perhaps they have developed a routine of use [30], tracking for potential later use, or passively collecting data, as happens with mobile phones tracking steps. This incidental tracking determines the type, amount, and quality of data that can be made available for an application, perhaps years after the data was first collected. Such data may have gaps, be imprecise or unreliable.

In practice, researchers should think in terms of both directions of flow. Only then can they really harness long-term data to answer research questions while supporting participants throughout studies and their daily life.

The roles of users Secondary users play an important role, shown at the top of the figure. These people need to make use of the data of the person who needs to ensure their data is collected. They may have many roles and differing relationships with the user collecting the data. For example, they may include advisors, experts, family members, caregivers, clinicians, trainers, or teachers. The particular role of the secondary users may help the user to collect data. More often they will help the user make sense of data. Those in a caregiver role may have more subtle uses such being able to gain peace of mind as they can have assurance that the user is doing well and knowing when to check in with the user [10, 41]. Such roles will typically mean that the relevant data is recent and short-term and the tracking is therefore purposeful. However, this may change if the secondary user discovers a way to make use of long-term data. For example, a caregiver might notice that an elder seems less well and then discover that long-term data can show a steady decline in activity.

In long-term tracking, the user has two roles. On the one hand, by self-tracking the user is the producer of data. The user's tracking behavior determines the availability, amount, and quality of data. Here, a user is likely to want to minimize the burden of tracking, particularly over the long-term, since sustained effort is particularly challenging. This wish for low effort may of course result in less and lower-quality data.

On the other hand, the user is also the consumer of the services delivered by the application such as visualized trends or recommendations for new activities to undertake. Here, the user's priority is to gain the benefits that may only be possible if there is more and higher-quality data.

Balancing the effort of tracking and the quality of data is therefore a key challenge of long-term tracking. There are two general approaches here: improving the data without increasing the effort needed for tracking is one way, e.g., by exploiting secondary sources such as social networks or digital calendars, or by developing better tracking devices that provide more and better data with the same amount of tracking effort. The other way is to motivate the effort needed for tracking, e.g., by providing short-term rewards or long-term benefits. For example, design of applications for long-term monitoring could provide a compelling case in terms of the promise to answer the user's future questions.

3.3 Reflecting on Case Studies

Our refined model of the relationship between different stakeholders and data in long-term tracking processes characterizes some of the dynamics surfaced in our case studies which other models were unable to capture. Jasmin and Joe's experiences demonstrate the relationships between tracking sources, the data they produce, the applications which leverage that data, and the various stakeholders which are involved in long-term tracking processes. IStuckWithIt and designing for lapsing provide example applications for supporting long-term tracking.

Jasmin had primarily purposeful motivations for tracking, trying to maintain a healthy lifestyle. Joe's goals were more incidental, driven to track by curiosity about his health and interest in trying new technologies. Although Jasmin and Joe engaged with secondary stakeholders only minimally, others intersected with their long-term tracking journeys at key points. Joe would occasionally share his activities with others to show his efforts toward maintaining a healthy lifestyle. Jasmin never used an application to share her data with a secondary user, but learned from others what tracking devices supported their needs and adopted them into her own practice.

IStuckWithIt and designing for lapsing are examples of applications which effectively mediate between primary users and their data. For incidental trackers, both applications deliver a service. In the case of IStuckWithIt, helping people make sense of their longitudinal adherence and activity levels. Designing for lapsing points out the opportunity to intervene at a point where users are disengaging. These applications also support the needs of purposeful trackers by helping people understand trends in their data and get value from it. While purposeful trackers might use applications like these to answer specific questions they have about their data (like Joe's use of Excel for his own analysis), they also automatically process the data delivered by tracking devices in ways which incidental trackers might find useful or interesting.

4 Characterizing Long-Term Tracking

To magnify how long-term tracking practices can be used in research, we use our case studies and prior work to explain how people's experiences differ from short-term tracking. Short- and long-term tracking differ in terms of the aspects listed at the left of Fig. 8: the targeted goal, typical duration, method of tracking, and approach to reflection.

As models of personal tracking use have pointed out, tracking can be a cyclic process where people lapse and resume the act of tracking [19]. We see the dichotomy between short-term and long-term tracking similarly. Many people lapse in the goals or methods of tracking we associate with short-term tracking into practices we associate with long-term tracking. Life changes, newfound curiosities, or symptom flare-ups can re-trigger periods of short-term tracking. For example, a person who recently had a child may reflect on their years of passively tracked physical activity data, set a new proximal goal appropriate for their new time demands, and regularly review their activity for a few months. They may then fall into a practice where that data tracking becomes incidental again.

4.1 Targeted Goal

Short-term tracking generally focuses on people's targeted goals, often rooted in their current experiences. Some examples of short-term goals could include, "work

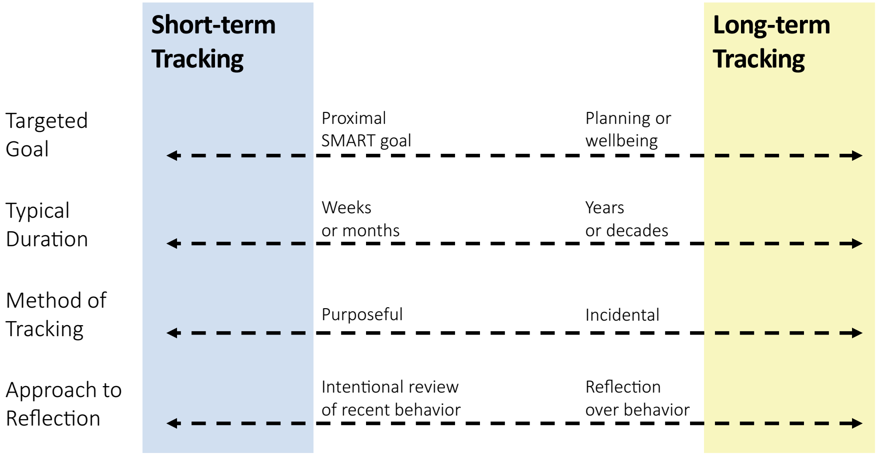

Fig. 8 People's targeted goals, tracking duration, method of tracking, and approach to reflection differ between long-term tracking and short-term tracking

out two times a week," "eat less carbs a day," "or walk three miles every day." For individuals already working out two times a week, for example, increasing the frequency to three times a week could be an manageable short-term goal. Research indicates that people are more likely to succeed if they formulate goals that fit the SMART criteria: Specific and Measurable, Achievable, Relevant, and Time-bound [15]. Alternatively, people's short-term goals may be driven more by curiosity or a desire to build general awareness, such as "understand about how much I walk in a day" [47, 51].

Research and commercial systems have examined a number of packages to support short-term tracking, such as setting a goal based on health standards (e.g., daily step guidelines such as 10,000 [60]) or past activity and recommendations or feedback based on the needs of cohorts [12, 20]. However, tools often provide insufficient guidance about how to identify the right data to collect toward their short-term goal [6] and insufficient support to help people interpret the data they collect [32].

Long-term tracking goals are typically more diverse, exploring broad planning or wellbeing goals versus specific and actionable ones. They sometimes align with more abstract concepts, such as hedonic and eudaimonic wellbeing needs of promoting pleasurable and enjoyable experiences and negative ones [42]. Some people may simply want to collect a long-term record of their activities (e.g., lifelogging) [19]. Other long-term health tracking goals often relate to identifying, enacting, and assessing changes in everyday life that are required to support health outcomes or other goals. These might be broken down into shorter-term goals. For example, losing 10kg in body mass may not be achievable in weeks or even months, and can be broken down into shorter-term tracked goals such as losing a few kilograms at a time. People working to manage chronic conditions may also need to work through an ongoing process of gaining a diagnosis, developing hypotheses about what contributes to symptoms, testing those hypotheses, developing action plans, monitoring, and then repeating and adjusting as circumstances change or symptoms reappear [8, 21, 49, 54]. As contributors to symptoms are understood, people may also use long-term tracking to support predicting—and possibly preventing—more severe symptoms[51, 61].

Tracking goals often change over the long-term as people gain more insight into their habits and limits. For example, people may change desired quantitative outcomes by setting a more aggressive or more realistic weight loss goal, or may switch from a weight loss goal to a maintenance goal. After some time tracking, they may also decide to pursue a different goal altogether, such as deciding to switch from a running goal to a swimming goal to better manage injuries or deal with other physical constraints. For example, some participants using the OmniTrack system re-configured the data they collected after some time to better align with their goals [28]. Jasmin's case study serves as an example of someone whose use of an activity tracker first satisfied a short-term goal of understanding her daily and weekly patterns, extending that to a longer-term goal of maintaining a healthy and active lifestyle.

4.2 Typical Duration

Short-term tracking makes sense for goals that can be achieved in a short time spans (hours, days or weeks). This tracking may have a definitive end, such as when a specific objective has been attained (e.g., a marathon being trained for, a weight loss objective achieved) or when a curiosity has been satisfied [16]. After that point, people may see no benefit on continuing to track [9, 30]. It has frequently been confirmed that many people drop out of self-tracking after short-term goals have been achieved, often within 3 to 6 months (e.g., [52, 53]).

By contrast, the increasing convenience and availability of tracking devices make it ever easier for people to track for years. Long-term tracking may involve multiple short-term goals. It may include multiple phases of changing tracking behavior, including periods of intensive self-tracking followed by months or years where no further data is collected. IStuckWithIt demonstrates how metrics like adherence can effectively represent these phases in designs.

4.3 Method of Tracking

During short-term tracking, data is typically collected purposefully. Depending on the domain and people's preferences, data may be collected manually (e.g., by journaling) or automatically (e.g., by passive sensing). People often find it burdensome to track daily or multiple times per day, though this is often required for domains like diet or weight monitoring. Some self-tracking approaches therefore place explicit time bounds on data collection, such as three-day food diaries [58] and fixed-duration self-experiments [13, 25].

People's long-term tracking methods are often incidental, very low effort or a side effect of using a device like a smartwatch. Joe, for example, intentionally chose tracking tools which would require minimal engagement on his part to best align with his incidental tracking goal. To lower the collection burden, long-term data streams often make use of passively collected data from phones and wearable devices such as location, steps, or heart rate. But people regularly switch what data they are collecting, switch tracking tools, and abandon and resume the same tool [19]. Long-term tracking typically therefore needs to be able to operate with a mix of use of whatever available data streams provide a view into their habits or goals, rather than assuming a single consistent and reliable data source. Jasmin's case study particularly reflects challenges in keeping a single reliable data source.

As a person's goal and purposes and contexts change over time, tracking behavior may change from purposeful to incidental. For example, a person who initially began tracking for weight loss may achieve their goal, but still continue to observe their weight because they developed a habit of logging it. However, they may later pursue another short-term tracking goal with purposeful intent.

4.4 Approach to Reflection

Reflection on short-term data is often an intentional exercise, such as opening an app with the goal of making note of daily physical activity logged by a phone or watch [23] or sitting down with a clinician to review a recent log of diet data [50]. People's review typically focuses on their recent behavior to understand their habits or experiences over the past hours or days.

Reflection over long-term data tends to occur in two ways, aligning with Schön's principles of reflection-in-action and reflection-on-action [48]. The act of collecting data leads some people to learn about their behavior and make changes to their practices (e.g., reflection-in-action) [19]. Jasmin in particular reflects this practice, learning about her practices and improving her ability to manage them over her first months of tracking.

But long-term data also presents opportunities for people to intentionally reflect over their behaviors to understand how they have changed or how their practices have evolved (e.g., reflection-on-action). Joe's efforts to make sense of his data serve as an example of this practice. The approach of presenting visual cuts to people who have lapsed in tracking points out how this specific moment can be leveraged to support reflection-on-action.

Schön also highlighted the importance of reflection-on-reflection, where the user re-assesses how well they have been using reflection. For example, a person who sees takes time each week to reflect on the physical activity progress may realize that this is not enough to recognize long-term drops in activity from one year to the next.

5 Discussion

We highlight a few recommendations when conducting research involving long-term data or developing a new design which collects or represents long-term data.

5.1 Consider Return of Data—and Actionable Insights—To Participants

We notice that researchers sometimes ask how to increase adherence in studies that require purposeful tracking, e.g., through remembering to use a device or answering experience sampling questions, often without returning data to participants [29]. Such studies might be intended to develop or validate new sensing devices or public health surveillance capabilities, or to better understand various aspects of everyday life, and so the return of data may seem like extra work that is not central to the study's goals. If the returned data influences the behaviors being studied (i.e., if it is an intervention), that return of data might be counter to study goals. But in other cases, returning data or insights to participants could be effective for increasing adherence.

However, when study designs collect, but do not return, tracking data, they must align with participant motivations in some other way, such as through participant motivation to support science or through financial or other incentives. Even when these other motivations are present, participants often join long-term tracking studies expecting the study to offer them some insights about themselves or their context, or at least access to and return of their data [29]. We encourage researchers leveraging tracking technology in their studies to consider whether they can support some return of tracking data, and resulting insights, to participants, both as an approach to participant motivation and as a way of making research less exploitative. Doing so also introduces opportunity for further research contributions leveraging long-term data, such as understanding what data and insights participants do and do not find useful for self-understanding or challenge their perceptions of themselves.

5.2 Anticipate and Respect Holes in the Data

Even when returning data or resulting insights effectively motivate participants to engage in long-term tracking and in studies that require it, the, data will have gaps. Holes can be short breaks in use, such as a few hours or days of missing data (e.g., "micro" holes). Or there can be periods where a user decides not to track for weeks, months, or years (e.g., "macro" holes). Most people do not keep complete records and track consistently. The emergent holes can be intentional, such as choosing not to track weight or physical activity over a busy holiday season, or choosing to focus on other priorities at a time when one's tracking goals are not the top concern. They can also be accidental, such as forgetting or losing the device being used to collect data [19].

Having complete and consistent logs may not be necessary or even desirable. A person can often understand their habits after a couple of days, weeks, or months of use. It is perhaps best for tracking tools to fade in and out of people's lives, supporting them in understanding how changes in their lives have impacted their habits and routines. That said, long-term data is essential for continued health in some chronic conditions. For example, a person with type-II diabetes must continually monitor and react to their blood glucose level. Technological advances aim to automate much of this monitoring (e.g., closed-loop insulin delivery systems), but the need to continually monitor still remains for many.

Expecting consistent and reliable long-term data does users a disservice. For "micro" holes, averages summarizing a period of activity can be skewed by including zero-counts or missing data alongside regular activity. For "macro" holes, a risk is that an application assumes that the user has stopped using the tracking device and begins prompting them to re-engage. When in reality, they may consciously decide to pick it up months or years later. Attempting to fill in missing data, in the short or long term, also risks errant conclusions, as events that affect one's goals could also affect one's tracking behavior, confounding results.

By emphasizing adherence in wearing behavior or chastising abandonment, applications and devices imply a "correct" and a "wrong" way of tracking. But the longer someone tracks, the more gaps there are likely to be, whether intentional or not. In general, adherence operates at multiple levels. For estimates of daily step counts and physical activity derived from wearables, it is important to take into account the amount of time the user wore tracker when measuring amount of activity. Similarly, to estimate weekly averages, the adherence matters for both how many days the user wore the tracker and how much of the day they wore it. Studies of large collections of tracker data show that adherence is important for interpreting data [57], but should not be used to tell users how to use their tracker. It is therefore important for an interface to respect and communicate the limitations of the data that a person collects. Likewise, some users may wish for interfaces that keep them engaged in tracking, but designers need to respect when users do not want to be engaged. This often comes into tension with the metrics that commercial products are often judged on, such as retention rate and the daily or weekly time spent in the app. However, we believe that holes in the data should be treated as a normal part of data rather than an exception to be avoided.

Researchers leveraging long-term tracking can use techniques like notifications [2], high financial incentives, or personal follow up when lapses are observed to promote adherence. These are valuable tools when they do not affect the behaviors the researchers are studying. However, they also can introduce confounds: they can interfere with studies to evaluate new tracking technologies. Even when the tracking technology is not the focus, they can affect the behavior or other factors being tracked through the demand effect or just increasing the salience of that behavior or type of data.

Researcher looking to leverage long-term tracking data in their research should not expect study participants to fully adhere to tracker use over a long period, even when strong incentives are offered. Analysis plans should be robust to these gaps, and researchers might consider also falling back to secondary data sources. For example, if collecting step counts, can a researcher supplement with data from a phone when someone does not wear another tracker? Gaps in use might also be relevant to researchers' questions, and so investigating lapses during interviews or by triangulating lapses with other data could offer more insight than tracking alone.

5.3 Leverage Implicit Tracking with Secondary Sources

There are numerous ways to track data. Using a dedicated tracking device such as an activity tracker or a sports watch is one of them; logging data by manually entering information in a diary is another one. In both cases, logging is based on the user's active decision for logging, and on the user being actively involved in the logging process, requiring some additional effort for logging, even if unobtrusiveness and ease of use reduce the effort to a minimum.

However, people also track data as a by-product to our daily digital lives: when posting information to our social networks, when communicating by email or instant message, when using digital calendars, when taking digital photos that store time and location, and when interacting with smart and networked buildings at home and at work. Sometimes people may not be aware the technology they use tracks them (e.g., Google Maps tracking their location, and Apple Health recording their steps). This data provides deep insights into our behaviors and daily lives, and it can be available over very long times without either initiative or ongoing effort from individuals. However, they can also violate people's privacy, or put the person (or others tracked) at risk. Challenges emerge in keeping this data accessible and persistent as people change devices, as there are no requirements for interoperability among different tracking platforms.

As one way to making it easier for people to participate in research that leverages tracking, we encourage researchers to ask "can we answer our research questions using data people are already collecting?" Additionally, similar to how we encourage researchers to consider what data and insights can be returned to participants, we also encourage researchers to ask "how can we help people make sense of the data that is tracked about them and available anyway?" When people may not be aware that this tracking is already occurring, can the research also promote their awareness and help them leverage the data better?

Some research work has already exemplified how such secondary sources may be unlocked. For example, De Choudhury et al. used social media posts in combination with logged food data to understand social support around weight loss [14]. Murnane et al. analyzed use of apps on mobile phones to understand biological rhythms [40], while Mehrota et al. leveraged use and duration of different phone features infer emotional state [34].

Such secondary data may be less precise or accurate than data that is purposefully tracked. This imprecision makes such data difficult to use in studies where small changes in short periods of time are important. However, the fact that no additional effort is required for tracking implies that such data can be made available over a very long time. In spite of the fluctuations, broad trends may well be identified with high reliability. Finally, because the data were generated as a routine part of other behaviors, the tracking may be less likely to influence those behaviors—important for studies intended to observe and understand, but not to intervene. Unlocking secondary sources to facilitate implicit tracking is therefore a strong opportunity for studying or supporting long-term tracking.

5.4 Treat Data as Subjective

Collected sensor data may seem perfectly objective: 5000 steps are 5000 steps, and 6 h of sleep are 6 h of sleep, no matter what. However, there is more to data than just the numbers: data has a meaning and a context, and this severely impacts the objectivity. Somewhat trivially, the devices and measurement methodologies influence the data. Using a dedicated activity tracker that can be worn at all times results in different data than using a smartphone that resides on the table a good part of the day.

However, the fact that a person decided to switch devices may be as or more important than the data itself. Many people change or abandon tools in response to changes in life circumstances, or because they achieved their behavior change goals [16]. In this sense, even the lack of data can help surface important information about technical issues faced, changes in health status, or what life events which triggered the outcome.

The meaning of data also changes over the long term. Walking 1,000 steps on a day can be little for most healthy people, but maybe a huge achievement for someone entering rehab after a severe health incident. Sleeping 4 h in a row during the night is not much for many people, but a lot for young parents. Such fluctuations are inevitable when aiming to make sense of trends in long-term data. Without contextual information it is therefore hard, if not impossible, to actually make sense of the data. Context is important for making use of long-term data in research.

Finally, even the purpose of the data may change over time. A person may originally collect activity data to monitor their personal fitness, but may later use that data to identify periods of depression. Heart rate data collected during sports may initially be used to optimize a workout, but may later provide valuable insights into changing cardiac health. Researchers and designers therefore need to consider how the same data can be used to answer the different questions people have in the long term.

We currently have few tools to make sense of long-term data. It is necessary to understand the story behind the data, which requires much more contextual knowledge than available today. Manual annotations or diaries may be a short-term approach, such as of moments of reduced air quality [39]. But even these measures have a higher cost than most people are able to maintain in the long term. Implicit tracking and secondary sources, whether intended for tracking or not, can help provide these annotations. Calendars, messages, social media posts, photos may provide the contextual knowledge that is necessary to really make sense of the collected data.

5.5 Ethical, Legal, and Social Implications of Long-Term Tracking

As a technology that goes straight into the highly personal life, long-term monitoring raises numerous ethical, legal and social implications. The specifics of these implications vary according to study goals, domain, involved populations, and locale. However, we offer some observations based on our experiences with long-term tracking.

Privacy of data is probably the most salient issue. Collected data are a valuable—often in ways that we may not even fully understand at the time we collect and first analyze them—many stakeholders may be interested in accessing the data or resulting analyses. Depending on the orientation and affiliation of the researcher, people may feel coerced into participating. People who desire tracking tools or the insight they provide, but are financially burdened by the cost of such devices or insights may feel coerced into providing their data, while people with means are free to ignore those incentives. Employers may give their employees a tracking device for free as part of research initiatives, but might want to observe their practices. Life insurance companies may similarly introduce research efforts which reduce customer premiums if their activity trackers record them achieving behavioral goals. This essentially disadvantages those without trackers or who choose not to use them and creates first and second class customers. This can further exacerbate inequities between people who are interested and able to do activities which the tracker does record (e.g., walk around) versus those who cannot or do not want to (e.g., if they live somewhere without sidewalks or good walking paths).

The ability to access one's own data is a topic that is becoming more pressing. Companies happily claim that "your data belongs to you", but at the same time build barriers to access and process the data outside the company's closed ecosystem. For example, many wearable devices only enable fine-grained data export for the past 30 days, making it challenging to provide long-term data exports. Other companies may not offer an API or an easy to process export at all. Policies such as the European GDPR provide a theoretical right to access one's own data. However, processes may be complicated and take a lot of time, and non-technical users may be overwhelmed and unable to understand and process their own data collection.

Data ownership also goes further. Parents may collect their children's data, but at some point need to hand over not just the responsibility, but also the data itself. However, the parents may want to retain some ownership over those data, as they also represent their memories and experiences as well as their children's. And what happens with my digital heritage, my data, after I die? Some of these questions have been discussed in related areas such as data stewardship (e.g., [5]). Self-tracked data introduces new kinds of records to consider preserving, sharing, or archiving, many of which were assumed to remain private.

When conducting research on and with long-term tracking data, we therefore need to be careful in our policies and practices around privacy, ownership, and stewardship. Using commercial self-tracking apps for research purposes can lower the design and deployment burden, but often means participants must consent to share their tracked data with the device manufacturer as well as the researcher. It can be highly ambiguous what about an individual that data might reveal when thoroughly analyzed, such as their habits or demographics. At minimum, it is important to enable research participants to delete or filter any of their data from study inclusion, whether prior to consenting researchers access or long after. Moving forward, it is worth considering how we as researchers can effectively communicate the risks (and benefits) which come from disclosing long-term self-tracked data.

Research studies requiring participants to collect long-term data should further consider what negative feelings or practices that data could evoke. Literature has pointed out how the act of self-tracking can lead to unhealthy changes in behavior, such as eating prepackaged foods because they are easier to journal [11] or trigger negative emotions, such as obsession with data collection to increase the likelihood of becoming pregnant [21]. Long-term tracking exacerbates these risks because the practices get further intertwined with the challenges of everyday life. It is therefore important to enable and support participants in disengaging from tracking, like they might naturally do if long-term tracking outside of a research context.

5.6 Making These Recommendations Work Together

To illustrate how many of these recommendations can work together, we note a study conducted by Propeller Health, the Institute for Healthy Air Water and Soil, and the Department of Civic Innovation at Louisville Metro, Kentucky, USA [1].

This study, AIR Louisville, enrolled 497 people with asthma to use connected rescue inhalers. Every time they used their inhaler, the use and location were automatically logged, and participants were also asked why they used it. This combined incidental data collection (use of the inhalers) with active data collection (asking about why). Data were collected, transmitted, and used consistent with the Health Insurance Portability and Accountability Act (thus following the relevant legal framework), and participants could choose whether to authorize their health provider to view the data (this protecting privacy and also participant ownership of the data). To prevent this study from exacerbating health disparities, researchers provided syncing hubs so that people could participate without a smartphone. These data were then also aligned with environmental data about nitrogen dioxide, particulate matter, ozone, sulfur dioxide, pollen, temperature, humidity, and wind (a secondary data source).

Study participants remained active in the study (defined as continuing to have their data sync) for a mean of 297 days—or about nine months. Results about participant's exposure levels were returned to individuals through Propeller Health's platform. Participants reported that this helped them understand the triggers for asthma in their lives. Collectively, participants achieved 78 percent reduction in rescue inhaler use and a 48 percent improvement in symptom-free days.

The results also informed local policy initiatives, such as where and how to enhance tree cover, recommended truck routes, zoning that creates air pollution buffers, and development of a community warning system for asthma. They also informed federal policy recommendations, lowering the ozone standard for healthy air from 70 to 65ppb.

This study illustrates how researchers can combine purposeful tracking with incidental tracking to answer research questions while providing data—and actionable insights—back to participants. This was achieved with a design that was resilient to lapses in tracking, and within a framework that protected participant privacy and supported their agency in how to share and use the resulting data. Following such a model led to better data, better outcomes for the participants, and societal impact.

6 Conclusion

Long-term tracking presents opportunities for observing people's practices by analyzing years or decades of their data, as well as designing technology to help promote longitudinal reflection over behavior to support planning or self-improvement goals. Compared to short-term tracking, the volume and duration of data generated in long-term tracking result in new considerations in the design of tools. Gaps in data must be expected, passively collected data should be leveraged over more burdensome journals, data must be contextualized in people's lived experience, and the data should be leveraged for personal benefit over surveillance. The use of long-term self-tracking in research is still nascent. There are many open challenges for further design, as well as important considerations when leveraging the practice in research.

Acknowledgements We thank our coauthors on our original research described in the case studies in this chapter: James Fogarty, Jennifer Kang, Laura Pina, and Lie Ming Tang.

References

1. Barrett M, Combs V, Su JG, Henderson K, Tuffli M, Louisville AIR (2018) Collaborative. Air louisville: addressing asthma with technology, crowdsourcing, cross-sector collaboration, and policy. Health Affairs 37(4):525–534
2. Bentley F, Tollmar K (2013) The power of mobile notifications to increase wellbeing logging behavior. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 1095–1098

3. Bonander J, Gates S (2010) Public health in an era of personal health records: opportunities for innovation and new partnerships. J Med Int Res 12(3):
4. Bot BM, Suver C, Neto EC, Kellen M, Klein A, Bare C, Doerr M, Pratap A, Wilbanks J, Ray Dorsey E et al (2016) The mpower study, Parkinson disease mobile data collected using researchkit. Sci Data 3(1):1–9
5. Brubaker JR, Dombrowski LS, Gilbert AM, Kusumakaulika N, Hayes GR (2014) Stewarding a legacy: responsibilities and relationships in the management of post-mortem data. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI '14, New York, NY, USA. Association for Computing Machinery, pp 4157–4166
6. Choe EK, Lee NB, Lee B, Pratt W, Kientz JA (2014) Understanding quantified-selfers' practices in collecting and exploring personal data. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 1143–1152
7. Chun KS, Bhattacharya S, Thomaz E (2018) Detecting eating episodes by tracking jawbone movements with a non-contact wearable sensor. In: Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies, vol 2(1)
8. Chung C-F, Dew K, Cole A, Zia J, Fogarty J, Kientz JA, Munson SA (2016) Boundary negotiating artifacts in personal informatics: patient-provider collaboration with patient-generated data. In: Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing, pp 770–786
9. Clawson J, Pater JA, Miller AD, Mynatt ED, Mamykina L (2015) No longer wearing: investigating the abandonment of personal health-tracking technologies on craigslist. In: Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing, UbiComp '15. New York, NY, USA. Association for Computing Machinery, pp 647–658
10. Consolvo S, Roessler P, Shelton BE (2004) The CareNet display: lessons learned from an in home evaluation of an ambient display. UbiComp 2004: Ubiquitous Computing. volume 3205 of Lecture Notes in Computer Science. Springer, Berlin/Heidelberg, pp 1–17
11. Cordeiro F, Epstein DA, Thomaz E, Bales E, Jagannathan AK, Abowd GD, Fogarty J (2015) Barriers and negative nudges: exploring challenges in food journaling. In: Proceedings of the 33rd annual ACM conference on human factors in computing systems, pp 1159–1162
12. Daskalova N, Lee B, Huang J, Ni C, Lundin J (2018) Investigating the effectiveness of cohort-based sleep recommendations. In: Proceedings of the ACM interactive mobile wearable ubiquitous technologies, vol 2(3)
13. Daskalova N, Metaxa-Kakavouli D, Tran A, Nugent N, Boergers J, McGeary J, Huang J (2016) Sleepcoacher: a personalized automated self-experimentation system for sleep recommendations. In: Proceedings of the 29th annual symposium on user interface software and technology, pp 347–358
14. De Choudhury M, Kumar M, Weber I (2017) Computational approaches toward integrating quantified self sensing and social media. In: Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing, pp 1334–1349
15. Doran GT (1981) There'sa smart way to write management's goals and objectives. Manage Rev 70(11):35–36
16. Epstein DA, Caraway M, Johnston C, Ping A, Fogarty J, Munson SA (2016) Beyond abandonment to next steps: understanding and designing for life after personal informatics tool use. In: Proceedings of the 2016 CHI conference on human factors in computing systems, pp 1109–1113
17. Epstein DA, Cordeiro F, Bales E, Fogarty J, Munson SA (2014) Taming data complexity in lifelogs: exploring visual cuts of personal informatics data. In: Proceedings of the 2014 conference on designing interactive systems, pp 667–676
18. Epstein DA, Kang JH, Pina LR, Fogarty J, & Munson SA (2016) Reconsidering the device in the drawer: lapses as a design opportunity in personal informatics. In Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing, pp 829–840
19. Epstein DA, Ping, Fogarty J, Munson SA (2015) A lived informatics model of personal informatics. In: Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing, pp 731–742

20. Feustel C, Aggarwal S, Lee B, Wilcox L (2018) People like me: designing for reflection on aggregate cohort data in personal informatics systems. In: Proceedings of the ACM interactive mobile wearable ubiquitous technologies, vol 2(3)
21. Figueiredo M, Caldeira C, Chen Y, Zheng K (2018) Routine self-tracking of health: reasons, facilitating factors, and the potential impact on health management practices. AMIA ... Annual symposium proceedings. AMIA symposium, pp 706–714
22. Fogg BJ (2002) Persuasive technology: using computers to change what we think and do. Ubiquity 2002(December):2
23. Gouveia R, Karapanos E, Hassenzahl M (2015) How do we engage with activity trackers? A longitudinal study of habito. In: Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing, UbiComp '15, New York, NY, USA. Association for Computing Machinery, pp 1305–1316
24. Gouveia R, Pereira F, Karapanos E, Munson SA, Hassenzahl M (2016) Exploring the design space of glanceable feedback for physical activity trackers. In: Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing, UbiComp '16, New York, NY, USA. Association for Computing Machinery, pp 144–155
25. Karkar R, Schroeder J, Epstein DA, Pina LR, Scofield J, Fogarty J, Kientz JA, Munson SA, Vilardaga R, Zia J (2017) Tummytrials: a feasibility study of using self-experimentation to detect individualized food triggers. In: Proceedings of the 2017 CHI conference on human factors in computing systems, pp 6850–6863
26. Kim D, Lee Y, Rho S, Lim Y (2016) Design opportunities in three stages of relationship development between users and self-tracking devices. In: Proceedings of the 2016 CHI conference on human factors in computing systems, CHI '16, New York, NY, USA. Association for Computing Machinery, pp 699–703
27. Kim Y-H, Jeon JH, Choe EK, Lee B, Kim K, Seo J (2016) Timeaware: leveraging framing effects to enhance personal productivity. In: Proceedings of the 2016 CHI conference on human factors in computing systems, CHI '16, New York, NY, USA. Association for Computing Machinery, pp 272–283
28. Kim Y-H, Jeon JH, Lee B, Choe EK, Seo J (2017) Omnitrack: a flexible self-tracking approach leveraging semi-automated tracking. Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies 1(3):1–28
29. Kolovson S, Pratap A, Duffy J, Allred R, Munson SA, Areán PA, Understanding participant needs for engagement and attitudes towards passive sensing in remote digital health studies
30. Lazar A, Koehler C, Tanenbaum TJ, Nguyen DH (2015) Why we use and abandon smart devices. In: Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing, pp 635–646
31. Lee MK, Kim J, Forlizzi J, Kiesler (2015) Personalization revisited: a reflective approach helps people better personalize health services and motivates them to increase physical activity. In: Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing, UbiComp '15, New York, NY, USA. Association for Computing Machinery, pp 743–754
32. Li I, Dey A, Forlizzi J (2010) A stage-based model of personal informatics systems. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI '10, New York, NY, USA. Association for Computing Machinery, pp 557–566
33. Mantua J, Gravel N, Spencer R (2016) Reliability of sleep measures from four personal health monitoring devices compared to research-based actigraphy and polysomnography. Sensors 16(5):646
34. Mehrotra A, Tsapeli F, Hendley R, Musolesi M (2017) Mytraces: investigating correlation and causation between users' emotional states and mobile phone interaction. In: Proceedings of the ACM interactive mobile wearable ubiquitous technologies, vol 1(3)
35. Meyer J, Beck E, Wasmann M, Boll S (2017) Making sense in the long run: long-term health monitoring in real lives. 2017 IEEE international conference on healthcare informatics (ICHI). IEEE, New York, pp 285–294

36. Meyer J, Kay J, Epstein DA, Eslambolchilar P, Tang LM (2020) A life of data: characteristics and challenges of very long term self-tracking for health and wellness. ACM Trans Comput Healthcare 1(2):1–4
37. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP (2011) A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the calo-re taxonomy. Psychol health 26(11):1479–1498
38. Montgomery-Downs HE, Insana SP, Bond JA (2012) Movement toward a novel activity monitoring device. Sleep Breath 16(3):913–917
39. Moore J, Goffin P, Meyer M, Lundrigan P, Patwari N, Sward K, Wiese J (2018) Managing in-home environments through sensing, annotating, and visualizing air quality data. Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies 2(3):1–28
40. Murnane EL, Abdullah S, Matthews M, Kay M, Kientz JA, Choudhury T, Gay G, Cosley D (2016) Mobile manifestations of alertness: connecting biological rhythms with patterns of smartphone app use. In: Proceedings of the 18th international conference on human-computer interaction with mobile devices and services, MobileHCI '16, New York, NY, USA. Association for Computing Machinery, pp 465–477
41. Mynatt ED, Rowan J, Craighill S, Jacobs A (2001) Digital family portraits: supporting peace of mind for extended family members. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp 333–340
42. Niess J, Woźniak PW (2018) Supporting meaningful personal fitness: the tracker goal evolution model. In: Proceedings of the 2018 CHI conference on human factors in computing systems, pp 1–12
43. Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska George SM, Olson RD (2018) The physical activity guidelines for americans. JAMA 320(19):2020–2028
44. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, Mohebbi MH, Mooney S, Suver C, Wilbanks J et al (2020) Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digital Med 3(1):1–10
45. Prochaska JO, Velicer WF (1997) The transtheoretical model of health behavior change. Am J Health Promot 12(1):38–48
46. Richardson CR, Buis LR, Janney, Goodrich DE, Sen A, Hess ML, Mehari KS, Fortlage LA, Resnick PJ, Zikmund-Fisher BJ et al (2010) An online community improves adherence in an internet-mediated walking program. Part 1: results of a randomized controlled trial. J Med Int Res 12(4):e71
47. Rooksby J, Rost M, Morrison A, Chalmers M (2014) Personal tracking as lived informatics. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 1163–1172
48. Schon DA (1984) The reflective practitioner: how professionals think in action, vo 5126. Basic Books
49. Schroeder J, Chung C-F, Epstein DA, Karkar R, Parsons A, Murinova N, Fogarty J, Munson SA (2018) Examining self-tracking by people with migraine: goals, needs, and opportunities in a chronic health condition. In: Proceedings of the 2018 ACM conference on designing interactive systems conference, pp 135–148
50. Schroeder J, Hoffswell J, Chung C-F, Fogarty J, Munson S, Zia J (2017) Supporting patient-provider collaboration to identify individual triggers using food and symptom journals. In: Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing, pp 1726–1739
51. Schroeder J, Karkar R, Murinova N, Fogarty J, Munson SA (2019) Examining opportunities for goal-directed self-tracking to support chronic condition management. In: Proceedings of the ACM interactive mobile wearable ubiquitous technologies, vol 3(4)
52. Shih PC, Han K, Poole ES, Rosson MB, Carroll JM (2015) Use and adoption challenges of wearable activity trackers. In: Conference 2015 proceedings
53. Shin G, Feng Y, Jarrahi MH, Gafinowitz N (2018) Beyond novelty effect: a mixed-methods exploration into the motivation for long-term activity tracker use. JAMIA Open 2(1):62–72
54. Sun S, Belkin NJ (2016) Managing personal information over the long-term, or not? experiences by type 1 diabetes patients. In: Proceedings of the 79th ASIS&T annual meeting: creating

knowledge, enhancing lives through information & technology, ASIST '16, USA. American Society for Information Science

55. Tang LM, Kay J (2017) Harnessing long term physical activity data-how long-term trackers use data and how an adherence-based interface supports new insights. Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies 1(2):1–28
56. Tang LM, Kay J (2018) Scaffolding for an olm for long-term physical activity goals. In: Proceedings of the 26th conference on user modeling, adaptation and personalization, pp 147–156
57. Tang LM, Meyer J, Epstein DA, Bragg K, Engelen L, Bauman A, Kay J (2018) Defining adherence: making sense of physical activity tracker data. Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies 2(1):1–22
58. Tremblay A, Sevigny J, Leblanc C, Bouchard C (1983) The reproducibility of a three-day dietary record. Nutrit Res 3(6):819–830
59. Tsai CC, Lee G, Raab F, Norman GJ, Sohn T, Griswold WG, Patrick K (2007) Usability and feasibility of PMEB: a mobile phone application for monitoring real time caloric balance. Mob Networks Appl 12(2–3):173–184
60. Tudor-Locke C, Hatano Y, Pangrazi RP, Kang M (2008) Revisiting how many steps are enough? Med Sci Sports Exercise 40(7):S537–S543
61. Woldaregay AZ, Årsand E, Walderhaug S, Albers D, Mamykina L, Botsis T, Hartvigsen G (2019) Data-driven modeling and prediction of blood glucose dynamics: machine learning applications in type 1 diabetes. Artif Intell Med
62. Wolf G (2009) Know thyself: tracking every facet of life, from sleep to mood to pain, 24/7/365