# ORCA – Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:https://orca.cardiff.ac.uk/id/eprint/130468/

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Meyer, Jochen, Kay, Judy, Epstein, Daniel A., Eslambolchilar, Parisa and Tang, Lie Ming 2020. A life of data: Characteristics and challenges of very long term self-tracking for health and wellness. ACM Transactions on Computing for Healthcare 1 (2) , -. 10.1145/3373719

Publishers page: http://dx.doi.org/10.1145/3373719

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

# A life of data - Characteristics and challenges of very long-term self-tracking for health and wellness

JOCHEN MEYER, OFFIS Institute for Information Technology, Germany
JUDY KAY, The University of Sydney, Australia
DANIEL A. EPSTEIN, University of California, Irvine, USA
PARISA ESLAMBOLCHILAR, Cardiff University, UK
LIE MING TANG, The University of Sydney, Australia

As self-tracking has evolved from a niche movement to a mass-market phenomenon, it has become possible for people to track a broad range of activities and vital parameters over years, even decades. The associated opportunities, as well as the challenges, have had very little research attention so far. With the phenomenon of long term tracking becoming widespread and important, we have identified its key characteristics, by drawing on work from Ubicomp, HCI, and health informatics. We identify important differences between long- and short-term tracking, and discuss consequences for the tracking process. Going beyond previous models for short-term tracking, we now present a model for long-term tracking, integrating its distinctive characteristics in purposeful and incidental tracking. Finally, we present major topics for future research.

CCS Concepts: • **Human-centered computing** → *Interactive systems and tools*; • **Applied computing** → *Consumer health*; *Health informatics*.

Additional Key Words and Phrases: longterm tracking, self tracking

**ACM Reference Format:**
Jochen Meyer, Judy Kay, Daniel A. Epstein, Parisa Eslambolchilar, and Lie Ming Tang. 2019. A life of data - Characteristics and challenges of very long-term self-tracking for health and wellness. 1, 1 (November 2019), 4 pages. https://doi.org/10.1145/1122445.1122456

## 1 INTRODUCTION

In recent years, self tracking of physical activity, sports, sleep, and much more has become widespread. This has spurred substantial work in the CHI community, demonstrating benefits of short-term tracking (e.g. [1]), over a couple of weeks or months, with health being the most frequent, but not the only application domain. With the increasing pervasiveness of dedicated tracking devices, of smartphones and smartwatches [7] with their multitude of sensors, with digital diaries for e.g. food, mood, or sleep, but also as a by-product of our digital life, people are increasingly accumulating long term tracking data, ranging from short bursts of tracking to multi-year records of activity or health data. We are just starting to understand how this introduces new opportunities, as well as challenges [3]. Building upon a workshop conducted at CHI 2018 [4], we subsequently

Authors' addresses: Jochen Meyer, meyer@offis.de, OFFIS Institute for Information Technology, Oldenburg, Germany; Judy Kay, The University of Sydney, Sydney, Australia, judy.kay@sydney.edu.au; Daniel A. Epstein, University of California, Irvine, Irvine, California, USA, epstein@ics.uci.edu; Parisa Eslambolchilar, Cardiff University, Cardiff, UK, eslambolchilarp@cardiff.ac.uk; Lie Ming Tang, The University of Sydney, Sydney, Australia.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.
XXXX-XXXX/2019/11-ART $15.00
https://doi.org/10.1145/1122445.1122456

identify key features of long-term tracking, present a model for long-term tracking, and suggest a research agenda.

## 2 SHORT-TERM VERSUS LONG-TERM TRACKING

It has frequently been confirmed that many people drop out of self-tracking within 3 to 6 months (e.g. [6]). We define long-term tracking as '*monitoring parameters of one's own life in timescales of years*'. This may include multiple phases of changing tracking behavior, multiple short-term goals, or periods of habitual and even inadvertent tracking e.g. with smartphones. For people to maintain a practice, amidst the realities of everyday life, leads to factors which are also relevant in short-term self-tracking, but have specific characteristics and properties in the long-term and therefore need special consideration:

**Incompleteness of data:** While trackers may be used every day in the short term, this may change in the long term as people will undoubtedly stop tracking for a few hours or days, or for weeks, months, or years, intentionally or accidentally, resulting in incomplete data [8]. Gaps in the data, which in the short term are often considered an exception, are therefore the rule and can even carry information of their own.

**Implicit tracking with secondary sources:** With the increasing digitization of daily life, we collect tremendous amounts of data about ourselves in secondary sources such as social networks, chats, workplace productivity software, online services etc. While such data may be unstructured, heterogeneous and fragmentary, the longitudinal coverage makes it a particularly interesting supplementary source to understand e.g. context and connectedness over the long-term.

**Subjectivity of data:** Data is not just an objective measurement. It may be amenable to multiple interpretations [8]. It may also tell a story of a person's context, situation, setting, and memories at the time it was collected [5], making it a memento of e.g. happy changes in life, stressful times, or diseases.

**Applications for long-term tracking:** There are numerous new potential applications for long-term tracking beyond behavior support, such as early-warning systems identifying slow changes, or supporting decisions in situations of change, based on past experiences [5].

**Secondary user:** Parents of small children, medical experts, formal or informal caregivers and other secondary users may have a legitimate reason to access the tracked data, but are not involved in the challenges of the tracking itself. In the long-term this raises additional issues such as changing data ownership and responsibility from the parents to an adolescent.

**Ethical, legal, and social implications:** With long-term tracking, data becomes a virtual representation of the person, capable of surfacing trends in their health and relationships. Long-term data therefore becomes a highly personal and life-long asset that requires even more attention to questions about data protection, access rights, security, privacy, or data ownership.

## 3 THE LONG-TERM TRACKING FEEDBACK LOOPS

As existing models of short-term tracking such as [2] do not account for these aspect, we propose the *long-term tracking feedback loops* in Figure 1. Here, the user is both producer of data and consumer of services, with potentially conflicting demands. As a *producer of data*, the user may decide to trade-off data quality because they want to reduce the effort they need to make for data collection. This may result in less data or lower-quality data. As a *consumer of the services* delivered by applications, there is a need for enough high-quality data. Consequently, we distinguish two different types of tracking:

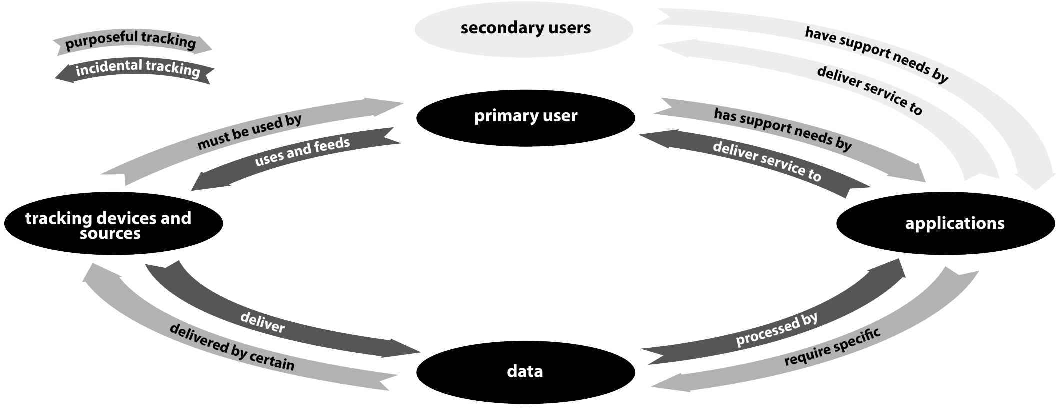


Fig. 1. Long-term self-tracking feedback loop. The outer, clockwise loop represents purpose-driven tracking, starting with the user's need for a specific service. The inner, counter-clockwise loop reflects incidental tracking with the user as a producer of data. These two loops are affected by a potential conflict between minimizing the demands upon the user and the need for data that makes applications effective.

*Purposeful tracking* is driven by a user's need for a certain form of support or service. The application delivering this therefore requires certain data, which must come from relevant tracking devices and sources. Ultimately, the purpose drives the tracking behavior required of the user. Changes in the user's tracking goals and their commitment to tracking have often compromised adherence [8], resulting in incomplete or missing data.

By contrast, in *incidental tracking* the user does not have a specific need. Tracking happens as a side-effect of routinely using devices or feeding secondary sources. Type, amount, and quality of data are determined by the user's tracking routine, not by potential future needs. These potential limitations of the data may in turn limit the types of services that an application can offer.

In practice tracking happens on a continuum from purely purposeful to solely incidental. Monitoring key parameters of diabetes over years is highly purposeful, but still may include incidental tracking e.g. of physical activity using a smart watch initially purchased to help stay on top of email. A key challenge when designing applications for this continuum is to match the required effort and anticipated benefits.

## 4 IMPLICATIONS AND FUTURE DIRECTIONS

Long- and short-term tracking share many properties and challenges. However, for many reasons, research so far has primarily focused on people's short-term goals and needs. We suggest that the research community is missing an understanding of the specific opportunities of long-term tracking. This requires understanding how long-term tracking is different from short-term tracking that we identified above. It is more incidental rather than purposeful, and less likely to be for a specific, proximal goal. It may involve repurposing data giving it value the user did not anticipate. It can change as people's long term goals and understanding of them evolve. The data and its quality is defined by the user's willingness to track, and applications must be satisfied with whatever data is there. The main challenge is, therefore, to maximize the value of existing data. This calls for future research, including, but not limited, to:

**Systems issues for implicit tracking:** This requires infrastructure, specifically designed to ensure security, user control of privacy and data provenance and methods to manage and analyze diverse data sources with heterogeneous data.

**Interfaces to visualize long-term data:** Visualizing and exploring one's own data is the initial step to self-understanding. However, short-term approaches do not scale well when it comes to unstructured, heterogeneous, and large long-term data. New visualization techniques are needed to make long-term data accessible and support sense-making by the layperson.

**Interfaces for long-term analyses:** Beyond visualization, we also need systems that can analyze and understand the heterogeneous, incomplete, and subjective data. We foresee systems that use rich analyses to identify changes in life and in the user's context, using the same, available data to answer different questions as they evolve, with a person's changing context over a life-time and the impact of new medical knowledge.

**Practice for long-term interventions:** Long-term data offers tremendous opportunities for life-long support. Concepts from short-term interventions cannot directly be applied in the long term. We need generic concepts for long-term interventions that can easily be applied in concrete use cases by designers and developers.

Long-term tracking is still in its infancy. While the boundaries to short-term tracking are blurry, characteristics of long-term tracking require special consideration. It is even more multifaceted and takes places in real life, with all its intricacies, making it both exciting and challenging, with tremendous opportunities for research and practice. We believe that future research should and will reveal exciting opportunities to deliver real benefits from long-term self-tracking to the user.

## REFERENCES

[1] Young-Ho Kim, Jae Ho Jeon, Bongshin Lee, Eun Kyoung Choe, and Jinwook Seo. 2017. OmniTrack: A Flexible Self-Tracking Approach Leveraging Semi-Automated Tracking. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3 (sep 2017), 67:1—-67:28. https://doi.org/10.1145/3130930

[2] Ian Li, Anind Dey, and Jodi Forlizzi. 2010. A stage-based model of personal informatics systems. In *CHI '10 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM Press, New York, New York, USA, 557. https://doi.org/10.1145/1753326.1753409

[3] Jochen Meyer, Elke Beck, Merlin Wasmann, and Susanne Boll. 2017. Making Sense in the Long Run: Long-Term Health Monitoring in Real Lives. In *2017 IEEE International Conference on Healthcare Informatics (ICHI).* 285–294. https://doi.org/10.1109/ICHI.2017.11

[4] Jochen Meyer, Daniel Epstein, Parisa Eslambolchilar, Judy Kay, and Lie Ming Tang. 2018. A Short Workshop on Next Steps Towards Long Term Self Tracking. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18).* ACM, New York, NY, USA, W05:1—-W05:8. https://doi.org/10.1145/3170427.3170605

[5] Tamar Sharon. 2017. Self-Tracking for Health and the Quantified Self: Re-Articulating Autonomy, Solidarity, and Authenticity in an Age of Personalized Healthcare. *Philosophy & Technology* 30, 1 (2017), 93–121. https://doi.org/10.1007/s13347-016-0215-5

[6] Grace Shin, Yuanyuan Feng, Mohammad Hossein Jarrahi, and Nicci Gafinowitz. 2018. Beyond novelty effect: a mixed-methods exploration into the motivation for long-term activity tracker use. *JAMIA Open* 2, 1 (2018), 62–72. https://doi.org/10.1093/jamiaopen/ooy048

[7] Statista. 2019. Wearables unit shipments worldwide by vendor from 2014 to 2018. https://www.statista.com/statistics/515634/wearables-shipments-worldwide-by-vendor/

[8] Lie Ming Tang, Jochen Meyer, Daniel A Epstein, Kevin Bragg, Lina Engelen, Adrian Bauman, and Judy Kay. 2018. Defining Adherence: Making Sense of Physical Activity Tracker Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1 (2018), 37:1—-37:22. https://doi.org/10.1145/3191769