A Scientometric Analysis of Self-tracking in Relation to Artificial Intelligence and Big Data

Xue Wu¹, Man Zhao¹, Han-Teng Liao¹* and Si-Pan Sun¹

¹ New Media Research Centre, Sun Yat-sen University Nanfang College, Guangzhou, Guangdong, 510970, China

*Corresponding author’s e-mail: h.liao@oxon.org

Abstract. As an increasing amount of personal data has been gathered by wearable and mobile devices, self-tracking, or the practice that people keep track, has become an important topic in Artificial Intelligence (AI) and big data applications. With the aim to provide a systematic review of the literature on self-tracking, this paper presents a scientometric analysis of 109 articles since 2000 collected from the Web of Science. Based on keyword co-occurrence network analysis, the paper has identified four major clusters: (1) wearables as quantified-self applications; (2) big data and critical theory; (3) data and privacy; (4) personal informatics. The further keywords-in-context (KWIC) analysis of the abstracts of the dataset clarifies the seemingly-interchangeable notions of "self-tracking" and "quantified-self": While "self-tracking" refers to more general activities, practices, technologies, and applications of keeping tracks, "quantified-self" refers to the more conscious efforts and meaning-making outcomes of the self-tracking activities. Such clarification, along with the keyword network analysis, suggests that self-tracking has become a specific and major type of datafication of human conditions or existence and that quantified-self is the construction of self through such datafication. A more integrated conceptual framework is needed for future research to better understand what amounts to meaningful datafication of human conditions and existence, thereby helping researchers and designers to discern the classic notions of health, wellness, and happiness for better research and design outcomes.

1. INTRODUCTION

As the practice of people keeping tracks has been recently amplified and enhanced by mobile and data technologies, a social and cultural phenomenon called “self-tracking” refers to the process of people turning their everyday experiences (especially wellness- and health-related) into data and data applications[1]. A major case of proactive self-tracking is the Quantified Self movement, originated from San Francisco in 2007 by Gary Wolf and Kevin Kelly from Wired magazine. The Quantified Self is defined as “an individual engaged in the self-tracking of any kind of biological, physical, behavioral, or environmental data”[2].

Self-tracking and the Quantified Self thus provide opportunities for meaningful collaboration across disciplines, especially social sciences and precision medicine[3]. The related work not only concerns health researchers and practitioners[4], but also may lead to new discoveries and innovations in health informatics[5]. Indeed, as increasing amount of personal data has been gathered by wearable and mobile devices[6], self-tracking has become an important topic in Big Data and Artificial Intelligence (AI) applications[7], with important discussions such as health-monitoring practices[8], open-access and privacy standards for personal data[9].
Several review articles have been conducted around the topic matter. One research has examined the notion of health self-tracking based on a 26 empirical studies [4]. Another has explored the socioeconomic and political implications of the risks associated with digital health technologies, which include self-tracking technologies [8]. Yet another has conducted a research on self-tracking in education settings using machine learning. However, such review work has not yet systematically examined the self-tracking research in intersections with Big Data and AI. This article therefore aims fill the gap by providing such a systematic review.

2. DATA AND METHODS

Following scientometric research conventions, the paper first describes the query design that covers both self-tracking and quantified self, in relation to AI and Big Data as follows:

("self-tracking" OR "quantified self" OR "self-quantification" OR "quantified-selfers" OR "personal quantification" OR "quantified academic selves" OR "autobiologies" OR "self-diagnosis") AND ("Artificial Intelligence" OR "machine learning" OR "Big Data")

About 109 articles are retrieved in November 2019 from the platform of Web of Science. Major research indexes such as SCI-EXPANDED, SSCI, A&HCI, and ESCI are included in the scope of this research. We used both VOSviewer and AntConc for an exploratory analysis of the dataset.

3. RESEARCH RESULTS

The following research findings first describe how literature are clustered based on keywords, and then discuss the semantic relationship among them, based on VOSviewer visualization outcomes. The findings also allow us to clarify and define the notions of self-tracking and quantified-self.

3.1. Keyword Map and Clusters

Using a constructed thesaurus that combine synonyms, we produced a keyword map (see Figure 1).

![Figure 1. A keyword map based on co-occurrence data of literature: clustered outcomes.](https://example.com/figure1.png)

Four main clusters have been identified based on the network of keyword co-occurrences. Each cluster reveals research focus in relation to other clusters/foci. The cluster #1 (wearables as quantified-self applications) includes AI, quantified self, biomedical, health data, wearables, IoT, social media, surveillance, and TORMES methodology. It contains research from benefits wearable devices and big data bring to AIED (Artificial Intelligence in Education) [10], to user motives of habitual integration of wearables in everyday [11]. Note that the topic “wearables” in the cluster #2 is different from that of “wearable technology” in the cluster #2. It appears to us that the term “wearable” refers to the general notion of wearable devices and artefacts, whereas the term “wearable technology” means technologies...
such as embedded sensors and algorithms that track, analysis and guide wearers’ behavior. The cluster #2 (big data and critical theory) includes topics such as big data, self-tracking, self-care, wearable technology, datafication, digital health, critical theory, algorithms, biopolitics, neoliberalism. Based mostly on critical theories, these studies explore social, cultural and political background of self-tracking. For example, research on datafication of clinical and self-care practices[12], and the Taylorist influence on wearable technologies within neoliberal workplaces [13]. The cluster #3 (data and privacy) consists of topics such as personal data, open data, privacy, data mining and analytics. It includes work such as the barriers to the use of personal health data [14], self-tracking data as posterity [15], etc. The cluster #4 (personal informatics) contains topics such as machine learning, personal informatics, stress-behavior pathway, and ecological momentary assessment. For example, how do we capture stress-behavior pathway with mobile technology and machine learning analytics [16].

3.2. The Relation of Self-tracking, Quantified-self, AI and Big Data
As shown in Figure 1, quantified-self has been discussed along with AI, IoT (internet of things) technologies and applications[17]. Also, wearables are considered as a novel interaction technology from “IOme” (internet of me), a concept that highlights personalized self-tracking experience compared to IoT(internet of things) [10]. Both advocate that AI make quantified-self applications smarter.

Quantified-self wearables are often regarded as an outcomes of active and positive engagement with Big Data, which is different from data fetishists[18]. Also, self-tracking wearables can be digital compasses that help consumers navigate everyday choices[19]. The findings indicate that self-tracking constitutes the everyday life practices and applications to health and wellness for constructing quantified-self.

3.3. The Definition of Self-tracking and Quantified-self
However, what makes “quantified-self” differ from “self-tracking”? Table 1 present the findings based on the keyword in context (KWIC) analysis of all abstracts of the dataset, using AntConc.

| Table 1. Keyword in Context Comparison: (a) self-tracking vs. (b) quantified-self |
|---------------------------------------------------------------|
| **Keyword in Context (KWIC)**                       | **source** |
| being restricted to an individualized form,               | self-tracking practices are also becoming part | AjanaB_2017 |
| to the data they generate in                              | self-tracking practices. We argue that the | SharonTZand_2017 |
| intentionally, as in the case of                           | self-tracking activities. Self-tracking | Kneidinger- |
| to follow from the availability of                         | self-tracking technologies and new volumes | McFallLMoor_2018 |
| as physical energy. In this context,                      | self-tracking technologies and related | TilIC_2019 |
| corporate wellness programmes using                       | self-tracking technologies shows that their | TilIC_2019 |
| (a) source on data about themselves,                      | self-tracking technologies reflect the | CrawfordKLi_2015 |
| and fitness companies. But can this                       | self-tracking technology sink its teeth into | BucherTArty_2016 |
| for rich, personalized collections of digital             | self-tracking records. Using qualitative | TraceCBZhan_2019 |
| which escape this data fetishist critique:                | self-tracking as a practice of mindfulness, | SharonTZand_2017 |
| of the ideologies and rationalities                       | self-tracking culture. I argue that such | BomanMSanch_2015 |
| of sensor input and output when                            | self-tracking wellness. The conclusion is that | GrillonAPer_2017 |
| intensity levels in order to allow                         | self-tracking while providing motivation. In | |
| users with higher mastery-orientation with positive      | quantified-self design more important. Users | HamariJHass_2018 |
| evaluations of gamification for users with high goal      | quantified-self design classes. Users with | HamariJHass_2018 |
| specificity fall under the classes of (1) gamification,    | quantified-self features were important. The | HamariJHass_2018 |
| (b) Self information in the workplace. The apps are part  | quantified-self and (3) social networking | LavaliereM_2015 |
| of a wider "era of new digital tools, the                | Quantified-Self movement has enabled | WeberlAchan_2016 |
| accompanying mindset of 'datafication,'                   | quantified self" movement and many opt-in | TorousSSum_2016 |
| by the do-it-yourself and                                  | quantified self-movement, and precision | vandeVenl_2018 |
| ---------------------------------------------------------- | quantified-self movements, it is now possible | SantosOC_2016 |
The research is funded by the Application-oriented Curriculum Development Project of "Information Visualization Design" (NFU 02-40250), under the Guangdong Province Department of Education 2018 "Innovation-Strengthening Higher Education Program" Grants, and partly by a project of HCI for Education Development (2018WQNCX283), under the 2018 Major Projects of Guangdong Province Department of Education "Young Innovative Talents" Grants.
REFERENCE

[1] Neff, G., Nafus, D. (2016) Self-tracking. (Cambridge, Massachusetts: The MIT Press)
[2] Swan, M. (2018) Quantified Self. Encyclopedia of Behavioral Medicine ed M Gellman (New York, NY: Springer) pp 1–4
[3] Tzovaras, B.G., Angrist, M., Arvai, K.J., Dulaney, M., Estrada-Galíñanes, V., Gunderson, B., Head, T., Lewis, D., Nov, O., Shaer, O., Tzovara, A., Bobe, J.R., Ball, M.P. (2019) Open Humans: A platform for participant-centered research and personal data exploration. GigaScience
[4] Almalki, M., Gray, K., Martin-Sanchez, F. (2016) Activity Theory as a Theoretical Framework for Health Self-Quantification: A Systematic Review of Empirical Studies. J. Med. Internet Res. 18:
[5] Saheb, T., Saheb, M. (2019) Analyzing and Visualizing Knowledge Structures of Health Informatics from 1974 to 2018: A Bibliometric and Social Network Analysis. Healthcare informatics research
[6] Rehman, M.H. ur, Liew, C.S., Wah, T.Y., Shuja, J., Daghighi, B. (2015) Mining Personal Data Using Smartphones and Wearable Devices: A Survey. Sensors 15: 4430.
[7] Swan, M. (2013) The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery. Big Data 1: 85–99.
[8] del Rio Carral, M., Roux, P., Bruchez, C., Santiago-Delefosse, M. (2017) Digital health: Promises, challenges, and fears. A literature review. Prat. Psychol. 23: 61–77.
[9] Crawford, K., Lingel, J., Karppi, T. (2015) Our metrics, ourselves: A hundred years of self-tracking from the weight scale to the wrist wearable device. Eur. J. Cult. Stud. 18: 479–96.
[10] Santos, O.C. (2016) Training the Body: The Potential of AIED to Support Personalized Motor Skills Learning. Int. J. Artif. Intell. Educ. 26: 730–55.
[11] Stragier, J., Vanden Abeele, M., Mechant, P., De Marez, L. (2016) Understanding persistence in the use of Online Fitness Communities: Comparing novice and experienced users. Comput Hum Behav 64:0: 34.
[12] Ruckenstein, M., Schüll, N.D. (2017) The Datafication of Health. Annu. Rev. Anthropol. 46: 261–78.
[13] Moore, P., Robinson, A. (2016) The quantified self: What counts in the neoliberal workplace. New Media Soc. 18: 2774–92.
[14] Bietz, M.J., Bloss, C.S., Calvert, S., Godino, J.G., Gregory, J., Claffey, M.P., Sheehan, J., Patrick, K. (2016) Opportunities and challenges in the use of personal health data for health research. J. Am. Med. Inform. Assoc. 23: e42–8.
[15] Trace, C.B., Zhang, Y. (2019) Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 CHI Conference (Glasgow, Scotland Uk: ACM Press) pp 1–6
[16] Cheung, Y., Hsueh, P.-Y., Qian, M., Yoon, S., Meli, L., Diaz, K., Schwartz, J., Kronish, I., Davidson, K. (2017) Are Nomothetic or Ideographic Approaches Superior in Predicting Daily Exercise Behaviors: Analyzing N-of-1 mHealth Data. Methods Inf. Med. 56: 452–60.
[17] Gruyer, D., Magnier, V., Hamdi, K. (2017) Perception, information processing and modeling: Critical stages for autonomous driving applications. Annu. Rev. Control 44: 323–41.
[18] Sharon, T., Zandbergen, D. (2017) From data fetishism to quantifying selves: Self-tracking practices and the other values of data. New Media Soc. 19: 1695–709.
[19] Schüll, N.D. (2016) Data for life: Wearable technology and the design of self-care. BioSocieties 11: 317–33.
[20] Till, C. (2019) Creating ‘automatic subjects’: Corporate wellness and self-tracking. Health Interdiscip. J. Soc. Study Health Illn. Med. 23: 418–35.
[21] Nafus, D., Sherman, J. (2014) This One Does Not Go Up To 11: The Quantified Self Movement as an Alternative Big Data Practice. Int. J. Commun. 8: 11.

[22] Woo, W. (2018) Augmented Human: Augmented Reality and Beyond. Proceedings of the 3rd International Workshop on Multimedia Alternate Realities - AltMM’18 the 3rd International Workshop (Seoul, Republic of Korea: ACM Press) pp 1–1

[23] Banos, O., Bilal Amin, M., Ali Khan, W., Afzal, M., Hussain, M., Kang, B.H., Lee, S. (2016) The Mining Minds digital health and wellness framework. Biomed. Eng. OnLine 15: 76.