International Journal of Communication 12(2018), 629–646 1932–8036/20180005

Self-Tracking Data as Digital Traces of Identity: A Theoretical Analysis of Contextual Factors of Self-Observation Practices

BERNADETTE KNEIDINGER-MÜLLER
University of Bamberg, Germany

Digital traces occur as a consequence of using digital devices or applications, but they can also be produced intentionally, as in the case of self-tracking activities. Self-tracking increases the amount of data that represents users' or communities' identity traces, and individuals, institutions, and companies are interested in analyzing these data, but few consider the framing conditions of the data collection, distribution, and evaluation. This article demonstrates how contextual factors influence self-observation data. Based on approaches of a sociology of quantification and a theoretical discussion of metadata in scientific research, it examines the individual, social, and technological contextual factors that influence the production, analysis, distribution, and interpretation of digital self-tracking data. The article develops systematization of the phenomenon of self-tracking data.

Keywords: self-tracking, digital traces, frames, thick descriptions, metadata

Individual characteristics, performance indicators, activities, and experiences are always rooted in and influenced by specific natural and social contexts. The opposite is also true: Individual characteristics leave traces in the contexts. According to Cheney-Lippold (2011), the data of individual and social practices or characteristics can be treated as "cultural objects" that always remain "embedded and integrated within a social system whose logic, rules, and explicit functioning work to determine the new conditions of possibilities of users' lives" (Cheney-Lippold, 2011, p. 167). This article discusses how contextual factors should be considered for the analysis of digital identity traces. Approaches from a sociology of quantification are used to discuss the reciprocal effects of numbers and data on social life and the influences of individual, social, and technological factors on the data. Additionally, parallels between the metadata's demand for the secondary analysis of empirical data and that for the interpretation of digital traces on the Internet are described. Based on Li, Dey, and Forlizzi's (2010) stage-based model of self-tracking, self-observation data are described as a special form of digital trace, and potential contextual factors that influence self-tracking data at different stages of the self-observation process are discussed.

Bernadette Kneidinger-Müller: bernadette.kneidinger@uni-bamberg.de
Date submitted: 2016–05–17

Digital Traces of Identity

In public and scientific discussions, it is frequently argued that individuals are no longer able to move around in modern society without leaving digital traces (Estrin, 2014; Li, Dey, & Forlizzi, 2011). Most such discussions focus on the handling of such data; less interest is devoted to the act of leaving "traces" (Reigeluth, 2014, p. 248). A physical footprint may be quite ephemeral; the digital one is potentially perpetuated forever. According to Reigeluth (2014), the aggregation of all such digital traces of an individual forms the "digital identity" of the person. Hand (2016) discusses digital traces as a kind of individual and social memory.

In the context of digital traces, the distinction between small and big data appears quite frequently. Estrin (2014) defines the term *small data* as data derived from individual data, and she describes this with the equation "n = me." By contrast, *big data* refers to extensive data sets that, in many cases, combine data from multiple sources or individuals and demand computational extraction and analytical methods. Self-tracking data can represent both data types (Swan, 2013). The providers of self-tracking applications that combine and aggregate multiple users' self-tracking data often store the data that the users collect for themselves (small data) in parallel (big data; Lupton, 2015). This article refers to self-tracking data used in both forms as individual small data and aggregated big data sets.

Digital Self-Tracking as Intentional Trace Production

Self-observation in the sense of data collection regarding individual activities, physical characteristics, and experiences has a long history (Crawford, Lingel, & Karppi, 2015; Swan, 2013). Its only new aspect is the technological simplification of data collection, evaluation, and analysis for the individual user (Li et al., 2010). The terms for digital forms of self-observation are manifold, including "personal informatics systems," "living by numbers," "quantified self," "self-tracking," "personal analytics," "life-logging," and "self-monitoring" (Lupton, 2014b). This article follows Lupton's (2014b) definition of *self-tracking*, which is described as the individual use of a technology to log everyday traits, observe them, and reflect on them.

According to Lomborg and Frandsen (2016), the state of research on self-tracking can be categorized into three main areas: (1) studies related to "health care" (Chiang, Yang, & Tu, 2014; Steele, 2013; Turner-McGrievy et al., 2013; Wang et al., 2014); (2) "interaction design and systems development research" (Ahtinen, Isomursu, Ramiah, & Blom, 2013; Consolvo et al., 2008; Epstein, Cordeiro, Bales, Fogarty, & Munson, 2014; Kim, 2014; Kranza et al., 2013; Li et al., 2010); and (3) studies based on a "critical-sociological lens" that address questions about surveillance, work, and the loss of privacy (Ruckenstein, 2014).

This article uses self-tracking data as examples of digital traces. Self-observation data constitute digital traces of a special type because they result from an intentional and active process of digital trace production (Lupton, 2014b). The user produces, collects, interprets, and distributes data that would otherwise not exist (Li et al., 2010). Self-tracking can be analyzed as a process of "datafication" (Mayer-Schoenberger & Cukier, 2013; van Dijk, 2014) that transforms individual characteristics and activities into digital traces of the individual. This can happen with the simple use of digital self-tracking tools or the intentional sharing of self-observation data.

A Sociology of Quantification

The quantification of social phenomena has always been a characteristic of social scientific research. Espeland and Stevens (2008) define quantification as "production and communication of numbers" and describe it as "a constitutive feature of modern science and social organization" (p. 402). Furthermore, quantification becomes a main claim in many political, economic, and even cultural contexts. Espeland and Sauder (2007) describe it as a "flood of social measures designed to evaluate the performances of individuals and organizations" (p. 1). Therefore, a systematic analysis of such quantification practices and their consequences in the society is emphasized by a sociology of quantification. The main assumption of this relatively young research approach is that social measures not only lead to an increased amount of data but also influence behaviors, experiences, and self-evaluation of individuals, institutions, or society as a whole (Espeland & Sauder, 2007; Hacking, 1990; Porter, 1995). Therefore, measures should be seen as highly reactive: Individuals alter their behavior in reaction to being observed, measured, and evaluated (Espeland & Sauder, 2007). But—and that is the second assumption of a sociology of quantification—the influences are not unidirectional; the relationship between data and the social is a deeply reciprocal one. Data influence social life, but at the same time, social factors influence the process of data collection, analysis, interpretation, and presentation. Therefore, following the arguments of a sociology of quantification, a systematic and critical analysis of quantified data in society is necessary and not only puts into question the quality of collected data but also considers the multitude of contextual factors that influence the data on different levels. The interdependence of data collection and social measurements on the one side and the influencing factors and behavior changes on the other side are mainly discussed with a focus on quantifications in the political and economic context and how such public measures influence power relations, financial aspects, or the efficiency of an organization or institution (Centemeri, 2012; Hayes, 2011; Miller, 2001; Rose, 1991). A quantification of the individual—the quantification of everyday life—is much less discussed. Therefore, this article discusses digital traces as a result of a conscious self-quantification.

The Social Meaning of Numbers

Following Espeland and Stevens (2008), a quantification of social life appears in two distinct forms: marking and commensuration. Whereas *marking* describes situations where numbers are used like names to identify particular persons, locations, or objects, *commensuration* refers to "the valuation or measuring of different objects with a common metric" (Espeland & Stevens, 2008, p. 408; see also Espeland & Stevens, 1998). This leads to a transformation of "all differences into quantity" (Espeland & Stevens, 2008, p. 408). Commensuration is a more complex process than using numbers simply to mark individuals or objects. The objects of commensuration must be classified first to make them comparable. As a consequence, the underlying data—the individual characteristics of an individual or object—are reduced and simplified (Espeland & Sauder, 2007). That is why all forms of ranking are frequently criticized for being oversimplifying (Espeland & Sauder, 2007). However, simplification often makes information seem more robust, definitive, and authoritative than if it were presented in a more differentiated form (Espeland & Sauder, 2007). Moreover, simplified data produce "decontextualized, depersonalized numbers that are highly portable and easily made public" (Espeland & Sauder, 2007, p. 18). That could explain the high importance of quantified data about political, economic, and social issues that are discussed in the media. However, Espeland and Sauder (2007) also see commensuration as a way to inspire people to scrutinize the meaning of numbers. This leads to the distinction by Desrosières (2001) of four attitudes toward the reality of collected data made. He distinguishes a metrological realism, in which the social relationship that is measured is as real as a physical object; an "accounting realism," in which the meaning of the numbers is inseparable from the trust that is predicated on the "fair" standardized practices that produce the numbers; a "proof in use" realism, in which numbers are judged real to the extent that they produce consistent, plausible results; and constructionism, which understands the reality of the measures shaped by measurement conventions (Espeland & Sauder 2007). The last of these—constructionism—highlights the main assumption of this article that digital data (traces) should always be seen in the context of various influencing factors to understand the process of data production, interpretation, and use.

In sum, what is discussed mainly in the context of public data collection and ranking politics can be transferred to the individual practices of self-monitoring and self-evaluation. Even for individually initiated data collection, we must question which factors stand behind the motivation of self-tracking activities, how data collection and data interpretation take place, and what is done with the collected data and their results.

The Significance of Metadata

The demand for a profound description and interpretation of existing data may sound like a premise of qualitative approaches, but, in fact, it should be a demand for all kinds of analyses. Cliff Geertz's (1973) description of his anthropological approach as "thick description" is worth applying to other scientific disciplines, including the field of big data analysis. Geertz assumes that no "pure data" exist, because all data that are extracted on an individual or social level are influenced by individual and social expectations and experiences that lie behind observable behavioral patterns and expressed attitudes. Additionally, technological aspects can influence behavior and, as a consequence, the digital traces produced. Mayer-Schoenberger and Cukier (2013) express a similar position by assuming that each data set is likely to have intrinsic, hidden, and yet unearthed values. As a consequence of their "multivalent" character, data should always be approached as multi-interpretable texts (Mayer-Schoenberger & Cukier, 2013). Gitelman (2013) states that "'raw data' is an oxymoron" because "data are not facts, they are 'that which is given prior to argument' in order to provide a rhetorical basis" (p. 7). Therefore, such contextual factors should be considered a form of metadata for the interpretation of digital data traces. Metadata should be defined as data that offer further information regarding primary data (e.g., background information about the creation of data). They influence the data that are of primary research interest. Similar to the secondary analysis of empirically collected data sets, the analysis of digital traces requires further information about technological, individual, and social factors that could have influenced the primary data source and offers additional information that is necessary for the interpretation of the digital traces. Such metadata are defined as necessary preconditions for the sustainable use of empirical research data in the form of secondary studies (Jensen, 2011). Even a kind of standardization of such metadata already exists (e.g., data about the project context, the methodology; for more detail, see Gebel & Liebig, 2013). The analysis of digital traces still cannot be based on such criteria of metadata. Therefore, this article identifies relevant forms of metadata that must be considered at different stages of the self-tracking process. Based on these assumptions, I present a first step toward a standardization of metadata in the context of self-observation data.

A Systematization of Self-Tracking Data

This article discusses why metadata in the sense of contextual information should be considered for the interpretation of individual and social traces on the Internet. Using digital self-tracking data as a special field, I identify various contextual factors and discuss their influence on the meaning and outcomes of individual data. This section describes how digital traces are produced during self-tracking activities and why they stand for a special case of digital traces. Subsequently, various contextual factors that have to be considered for the interpretation of self-tracking data traces are mentioned. The argumentation is oriented toward the stage-based model of self-tracking that Li et al. (2010) developed. It allows the differentiation of contextual factors depending on different stages of the self-tracking process. This model is used to illustrate how social factors influence the meaning of digital self-observation data starting from the preparation stage, over the data collection and evaluation stage, and up to the action stage.

The Five Stages of Self-Tracking

Li et al. (2010) differentiate in their model five stages of self-tracking: At the "preparation stage," the general motivation of self-observation and the decision regarding which information should be collected in which form are important features. At the "collection stage," all relevant data are logged and collected. This is the stage where the digital traces are produced. The "integration stage" forms a middle position between the collection stage and the "reflection stage"; it includes data preparation and the calculation of statistics and other key values. At the reflection stage, the findings based on the data are interpreted and reflected on. Finally, at the "action stage," the users react to the findings of the self-observation and use the newly acquired knowledge about themselves to draw conclusions about success or failure, good or bad habits, and so on. Various individual, social, and technological factors influence the digital traces that are prepared, produced, analyzed, and interpreted at these five stages. To offer detailed insight into relevant contextual factors, they are presented separately for each of the five stages. Figure 1 provides an overview of the five stages and the corresponding contextual factors that are discussed in detail later.

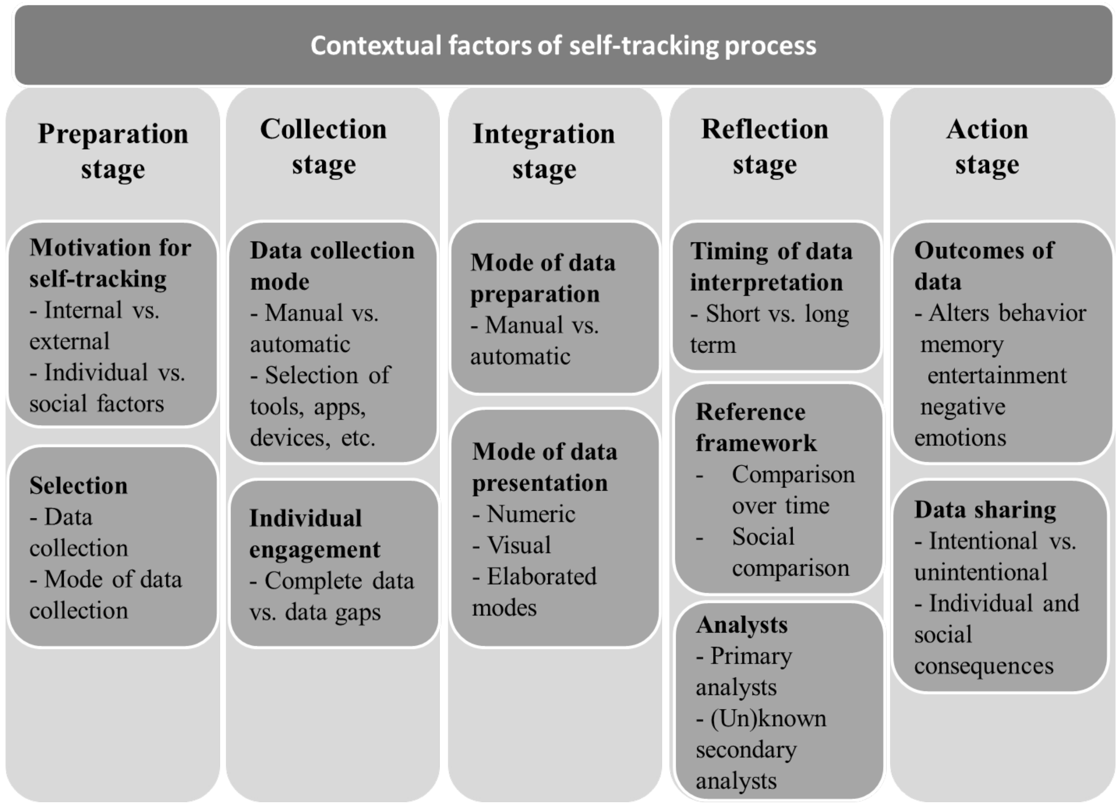

Figure 1. Contextual factors of the self-tracking process.

Contextual Factors at the Preparation Stage

Motivation as a Contextual Factor

At the preparation stage, the questions about why a self-tracking process has been started and what should be tracked should be answered. The individual who is interested in specific self-related data (self-motivation) can intrinsically define the motivation for self-tracking. Alternatively, others, who motivate the individual to collect data to get a predefined advantage (external motivation), can impose the motivation externally. An example could be that of a medical professional who tells a patient to keep records of specific physical conditions to improve his or her health status. Lupton (2014b) differentiates between five modes of self-tracking, depending on whether the self-tracking happens voluntarily or involuntarily and privately or publicly. Three modes can be interpreted as self-motivational forms ("private self-tracking," "pushed self-tracking," "communal self-tracking") and two forms as external motivations ("imposed self-tracking," "exploited self-tracking"). Lupton's categorization makes it clear that motivation as a contextual factor represents not just an individual factor but also a social factor. Rooksby, Rost, Mossiron, and Chalmers (2014) also identify different forms of self-tracking that are strongly influenced by the users' life stories and the social context of the usage.

The wish for improved self-knowledge, the desire to change life habits, and the optimization of one's own life and personal decisions can be summarized as the frequently named internal motivations for self-tracking (Choe, Lee, Lee, Pratt, & Kientz, 2014; Li et al., 2010; Ruckenstein, 2014; Ruckenstein & Pantzar, 2015). Regarding self-optimization, the work of Michel Foucault is frequently cited and forms the theoretical framework for many articles about contemporary selfhood (Elliott, 2013; Rose, 1990, 2007). Foucault (1988) discussed the increasing social demand for continuous self-observation in the 1980s, which should help to optimize the individual self and, consequently, improve social life as a whole. These assumptions indicate that the underlying motivation reflects individual and social factors. The first one describes an individual wish (e.g., the improvement of one's own well-being). The second one refers to social motives (i.e., the monitoring and improvement of physical characteristics to meet anticipated social norms such as ideal weight). In all cases, the specific motivation influences the data that are collected, how engaged the users are in the data collection process, how the data are analyzed, and how the collected data are treated in general (e.g., if the data are shared with others online or if the individual uses the data exclusively).

Selection as a Contextual Factor

The second step entails determining which data should be selected in which form. The decision strongly influences which data will be available at later stages. Only those aspects that are defined as worthy of being logged become data (Lupton, 2015). Therefore, digital traces always represent a highly selective part of individual identity (Rooksby et al., 2014). Thus, this step must reflect the reasons why certain individual or social aspects are monitored and transformed into digital traces and others are not taken into account at all and therefore remain invisible in the digital landscape.

Additional selection happens at the technological level when the individual has to decide which form of data collection to use. Traditional paper-and-pencil solutions are possible, but various computer-based digital self-tracking applications also exist. Depending on the mode of data collection, different forms of data quality and data depth are created. Therefore, the selection of a specific self-tracking mode influences, to a large extent, the forms of digital traces that appear in the digital world.

Contextual Factors at the Collection Stage

Data Collection Mode as a Contextual Factor

At the collection stage, the digital traces are created by converting observations of habits or characteristics into digital data. Various individual, social, and technological factors influence the data collection process. At the technological level, the mode of the data collection and the type of data collection tool form another frame for the analysis of digital traces. A superficial distinction can be made between manual and automated data collection. In the manual mode, the individual notes all the significant data based on individual perception and evaluation. In the automated data collection mode, computer programs or specialized devices (e.g., wearables) log the relevant data (Rooksby et al., 2014). Manually collected data are often considered more subjective and, therefore, less precise. Van Dijck (2014) observes a kind of "dataism," which is the idea that people trust objective quantifications based on the automated logging of human habits more than they do their own subjective perceptions. This goes along with the findings of Espeland and Sauder (2007) about the perception of quantified data about social issues that are seen as more reliable than case studies without concrete statistics. Reigeluth (2014) discusses a "naturalization of data," which highlights "objective" qualities and the ability to "say the truth." He critiques the distinction between data, information, and knowledge as being blurred. Digital traces in their pure form offer quite little information and, therefore, cannot be interpreted as "secured knowledge." Only in combination with the various contextual factors do the "naked data" get their actual meaning (Reigeluth, 2014).

The data collection mode influences the form of digital trace. Therefore, it is important to determine why a certain mode was selected. Some observable dimensions predefine the collection mode because they cannot be logged automatically and have to be estimated and recorded manually (e.g., mood, pain). By contrast, manual logging would be difficult in the case of unconscious activities (e.g., counting the steps that one takes over 24 hours). If no predefined restrictions exist, the individual can choose freely between manual and automated data collection. The mode of data collection in combination with the resulting data quality should be considered another contextual factor for the interpretation of digital traces. Similar conclusions can be drawn regarding the decision of which of the various data collection tools, apps, or devices to use for self-tracking. All of them imply the production of a specific quality of digital data. They influence which data are collected, the form in which they are collected, and the precision with which they are collected. Therefore, the applications and self-tracking tools used should be considered contextual factors of digital self-observation traces.

Individual Engagement as a Contextual Factor

Individual engagement during the data collection process constitutes another framing factor, especially in cases where the self-tracking is not fully automated. Depending on the individual interest in the data collection—a fact that the basic motivation of the whole self-tracking process influences—the relevant data may be collected in their entirety or show missing data points. Only in the first case will the digital traces be able to display the specific characteristics and habits of the individual completely. If there are gaps in data collection, the digital traces could be biased because of the missing information (Elsden & Kirk, 2014). In addition to the problem of the missing data, the lack of knowledge about the existence of such data gaps in the digital traces forms the most problematic aspect for further analyses. The literature about dropout rates in survey data documents the handling of missing data well, but little research discusses the handling of gaps in digital traces. A first step in identifying potentially existing data gaps in self-tracking data involves monitoring and considering individual engagement in the data collection process as another framing factor.

Contextual Factors at the Integration Stage

At the integration stage, when the preparation of the digital traces takes place, technological and individual elements form important metadata. The preparation of the data is necessary for the extraction of significant information from the collected data. Depending on the data collection mode and the intended analytical steps, the preparation of data can be a quick and easy task or a time- and calculation-consuming process. Analogous to the data collection modes, the data preparation can be manual or automated and involve computer-supported methods or paper-and-pencil methods (Choe et al., 2014). The preparation mode used influences the information value that can be extracted from the collected data. Computer-supported methods allow the combined analyses of different logged aspects, the calculation of longtime trends, and the prediction of further performances, experiences, or developments. In many cases, data are not only evaluated in the form of numbers or tables but also in visualized forms. Visualizations are especially helpful in the context of self-tracking (Whitson, 2013). They can take the form of simple line graphs or bar charts, or they can be elaborate infographics, maps, or photo grids (Choe et al., 2014). Visualized data help to identify relations between various digital data. Additionally, visualized data are evaluated as being more reliable and more precise due to individual and subjective perceptions (Ruckenstein, 2014; Ruckenstein & Pantzar, 2015). As a consequence, the mode of visualization that results from technological options and individual decisions frames the perception of the digital traces.

Contextual Factors at the Reflection Stage

At the reflection stage, the user interprets the individual meanings of the collected data. According to Ruckenstein and Pantzar (2015), data "does not have value or meaning in itself, rather it becomes part of the process of sense making." During the reflection process, the individual enters into communication with the self (Lomborg & Frandsen, 2016) by comparing and evaluating the individual self with a "data double" of the self (Ruckenstein, 2014). Thereby, the individual brings together subjective experiences with "objective" data. If the individual experiences do not fit the data, the individual must decide which one he or she trusts more. In many cases, the conclusion is deeply personal, even idiosyncratic (Ruckenstein & Pantzar, 2015). Various individual, social, and technological factors influence this process of sense making.

Timing of Data Interpretation as a Contextual Factor

The interpretation of digital traces can happen immediately after the data are collected (short-term) or some days, weeks, or even months afterward (long-term). Short-term interpretations are mainly focused on single data points, whereas long-term evaluations use multiple data collection points in combination (Li et al., 2010). Most self-tracking activities show a short-term orientation (Rooksby et al., 2014). Depending on the timing of the data evaluation, the digital traces fulfill different functions and acquire different meanings for the individual. In the case of short-term reflection, single data points stand for a status quo description (e.g., the current performance). In the case of long-term reflection, single data points lose part of their significance, but relations between data points become more interesting for identifying developments over time. Elsden and Kirk (2014), who invented the term "quantified past" and use it to refer to self-tracking data employed as a kind of long-term memory, discuss an extreme form of data interpretation with time delay. They highlight the fact that even data that are collected primarily for short-term interpretation may be interesting many years afterward, but they will probably go along with different intentions and experiences based on the data (Elsden & Kirk, 2014). Thus, the timing of data interpretation influences how a digital data trace is used and evaluated.

Reference Framework as a Contextual Factor

Due to the fact that single numbers get special meanings depending on their contexts, a comparative approach appears highly relevant for self-tracking. Such comparisons happen at the individual level with the comparison of self-tracking values over time to explore developments and trends (Li et al., 2011). Another form is social comparison, where individual data are compared with the data of other people to determine how the individual values compare to those of others. The selected comparison group influences the evaluation of the individual data. If the athletic performance of an amateur athlete is compared with the results of professional athletes, the evaluation of the data would probably be worse than it would be if the same performance were compared with those of other amateurs. The chosen reference framework influences the meaning attributed to digital data. Mortier, Haddadi, Henderson, McAuley, and Crowcroft (2014) discuss the significance of "human-data interactions" and recommend shifting the focus away from the question of how people handle self-tracking devices or applications to how they interpret and evaluate the collected data.

Analysts as a Contextual Factor

The third important contextual factor at the reflection stage concerns the type of analyst using the self-tracking data. In most cases, the primary analyst is the individual who has collected the data about himself or herself. The individual data interpretation happens based on self-tracking motivations, experiences during data collection, individual characteristics, technical competencies, personal goals, and so on. Secondary analysts of the data can be divided into known and unknown persons or institutions. For instance, known secondary analysts include supervising medical professionals who use the self-tracking data of patients to develop treatment plans. Unknown secondary analysts remain more or less invisible, and, in some cases, individuals are not even aware of the external use of their data. The unknown secondary interpretation of individual data happens quite frequently, because many providers of self-tracking applications collect, store, and analyze the self-tracking data of their users. In many cases, they even store more of the users' data, which they return to the users as the output of their individual self-tracking activities (Estrin, 2014; Estrin & Juels, 2016; Till, 2014).

Without delving into a discussion of data transparency and the protection of privacy, the significance of differentiation between various types of data analysts must be emphasized. Individual and external analysts have different knowledge regarding the contextual factors of self-tracking that are discussed in this article. In most cases, primary and secondary analysts have different intentions for and interests in the collected data. And these exact factors influence the interpretation of the digital traces. In their study concerning health-related data, Fiore-Gartland and Neff (2015) demonstrate how technological designers, medical practitioners, advocates, and patients interpret the same data completely differently.

Whereas the medical practitioners use the data as information for the development of conduct recommendations, the patients interpret the data that they have collected as narrations of their selves. These findings explicitly confirm that digital traces, no matter how objective and quantifiable they may be, have different meanings depending on the intentions and positions of the particular analysts.

Contextual Factors at the Action Stage

Outcomes as Contextual Factors

The action stage is characterized by the results of the whole self-tracking process. At this stage, the individual decides how the newly gained knowledge based on the self-monitoring process will be used for concrete actions or conclusions. Many self-trackers use their self-observation data to document advantages and control goal attainment (Choe et al., 2014). If a deviation from the intended course is observed, a behavior correction is planned based on the self-observation data. Some applications send automated warnings or reminders in such cases (Lie et al., 2010; Lomborg & Frandsen, 2016). Additionally, the digital traces are used to identify triggers for problematic or unwanted situations or conditions. These triggers can be observed systematically and avoided, which can help to improve health conditions or work processes. Conversely, behaviors that are identified as having positive outcomes can be intensified in other situations (Choe et al., 2014; Lie et al., 2011; Ruckenstein, 2014; Ruckenstein & Pantzar, 2015). This goes along with the assumption of a sociology of quantification that collected data influences the behavior of the individuals that are measured. In some cases, the collected self-tracking data are used as "reminiscing of the past" (Peesapti et al., 2010) and for "aiding memory" (Hodges et al., 2006). Additionally, one experiences the simple process of collecting data about oneself as entertainment and pleasure (Lomborg & Frandsen, 2016; Ruckenstein & Pantzar, 2015; Whitson, 2013). It can generally be said that self-generated digital traces support individuals in their self-management by offering them the feeling of greater control over their lives (Choe et al., 2014; Li et al., 2010; Nafus & Sherman, 2014; Ruckenstein, 2014). Even latent fears can be reduced (Selke, 2014)—a fact that is especially important in a world where traditional social structures are fast disappearing (Lupton, 2014a).

Besides positive experiences with self-tracking data, negative consequences or negative emotions may appear at the action stage. Individual traces can evoke dissatisfaction and frustration if the data reveal unpleasant habits or characteristics that the individual normally ignores (Choe et al., 2014). Restrictions of or threats to privacy can be experienced even in situations of voluntary and self-motivated self-tracking. Insufficient transparency about the generation and use of the produced data traces has to be highlighted as a major critique of self-tracking applications (Andrejevic, 2014; Lupton, 2015; Nafus, 2013, 2014).

Sharing of Self-tracking Data

Digital self-tracking data can be used exclusively by the individual who logs the data but also can be shared actively with others. With the sharing of self-tracking data, communication between different social contacts and networks takes place (Lomborg & Frandsen, 2016). The scenario mentioned earlier, where self-tracking data are shared with medical practitioners, is just one example of intentional data sharing (Li et al., 2010). Self-tracking data are also shared in private social networks, for example, during the comparison of individual athletic performances with those of training partners and the coordination of joint training. In the age of digital social media, self-tracking data are easily shared online in broader social contexts. But what moves people to share personal data with others and to create digital data traces with their sharing activity? The distribution of self-tracking data has mainly been analyzed in the context of the sharing of sports results. Lupton (2015) identifies the elevated importance of the competitive factor or the wish to get support and encouragement as the main reasons for this. Additionally, users feel pleasure when they share their sports successes on social media. Lomborg and Frandsen (2016) highlight the significance of the social recognition of individual performances. Moreover, Li et al. (2010) name the option of receiving advice from people on the extended social network as a reason for the sharing of self-tracking data in general. In many cases, the self-tracking data are used in a strongly performative way with the sharing of successes and positive elements in particular and the leaving out of failures and negative elements (Lupton, 2015). Based on Goffman's (1959/1984) "self-presentation theory," sharing can be interpreted as a process of identity construction and self-presentation. The sharing of self-tracking data is stimulated by strong identification with the group with which the data are shared (Stragier & Mechant, 2013). Similarly, the sharing of data intensifies the feeling of group belonging (Lomborg & Frandsen, 2016; Whitson, 2013). A special form of sharing occurs when defined individual goals are shared. The sharing is experienced as an act that facilitates the achievement of these goals and motivates others to set targets for themselves (Lomborg & Frandsen, 2016). The sharing of individual data can contribute to problem solving at a more general social level—for example, the development of alternative medical treatments (Estrin & Juels, 2016).

Independently of individuals' intentions, the sharing of personal data allows others to observe and eventually use them. The voluntary public distribution of personal data is discussed as "social surveillance" (Marwick, 2012), "participatory surveillance" (Albrechtslund & Lauritsen, 2013), and "reflexive self-monitoring" (Lupton, 2014b, p. 12).

To summarize, the digital traces that individuals leave intentionally must always be interpreted, because, in many cases, the selection of the shared data happens strategically and selectively. Ellison, Heino, and Gibbs (2006) call it "managing impressions online" and Goffman's (1959/1984) theoretical concepts regarding the construction of identity depending on specific social roles are still valid in the context of digital self-tracking habits. This implies that shared digital data traces always represent nothing more than a selective part of the collected data. The data selection does not happen randomly, but consciously; thus, a systematic bias regarding the transmitted data can be assumed. Consequently, all interpretations of publicly shared self-tracking data have to consider the motivations behind the sharing of the data; these motivations should be treated as another framing factor of digital traces of the self.

Discussion

Theoretical considerations and a review of actual self-tracking research are used to identify contextual factors that influence the production, preparation, analyses, and use of self-observation data. These contextual factors are discussed as important metadata for the adequate evaluation and interpretation of this special form of digital identity trace. The description of relevant contextual factors reveals the many influences on the creation, preparation, and distribution of personal data. At different stages of the self-tracking process, different metadata have to be considered during the analysis of the data. The relevant contextual factors discussed here should not be seen as complete and are mainly focused on the special field of self-tracking data. Nevertheless, this discussion supports the assumption that publicly available digital data traces should not be used as "pure data" if the individual or social significance of the values is of interest. The assumptions and approaches of a sociology of quantification, which is still mainly focused on political, economic, or scientific data collection, must be transferred to individualized data collection as a form of self-initiated production of digital traces. The increasing interest in monitoring various aspects of one's own body and everyday life may be interpreted as an extension of a quantification of the social to a quantification of the individual. Numbers become more and more important on the micro-, meso-, and macrolevel. They are used to measure, compare, rank, and evaluate oneself, an institution, or a society as a whole. Data become an important value in modern society. Consequently and more than ever before, the analysis of individual or social habits, behavioral patterns, and attitudes should always go far beyond the pure observation and use of the available data traces. Individual, social, and technological framing conditions have to be used more frequently as metadata and contextual factors for the analysis of small and big data sets. Similar to trackers of animal prints, who need to have knowledge about the behavioral patterns and living habits of different animals to understand their traces, social scientists need profound background information for the analysis of digital human or social traces. Digital traces can only become information and, ultimately, knowledge if they are analyzed while they are embedded in their contexts and in combination with each other. The research challenge in the next few years will not be the collection of more extensive data sets of digital traces, but the development and application of methods that allow the integration of contextual factors into the analysis of digital traces.

References

Ahtinen, A., Isomursu, M., Ramiah, S., & Blom, J. (2013). Advise, acknowledge, grow and engage: Design principles for a mobile wellness application to support physical activity. *International Journal of Mobile Human Computer Interaction*, *5*(4), 20–55. doi:10.4018/ijmhci.2013100102

Albrechtslund, A., & Lauritsen, P. (2013). Spaces of everyday surveillance: Unfolding an analytical concept of participation. *Geoforum*, *49*, 310–316.

Andrejevic, M. (2014). The big data divide. *International Journal of Communication*, *8*, 1673–1689.

Centemeri, L. (2012). The contribution of the sociology of quantification to a discussion of objectivity in economics. In J. Castro Caldas & V. Neves (Eds.), *Facts, values and objectivity in economics* (pp. 110–125). New York, NY: Routledge.

Cheney-Lippold, J. (2011). A new algorithmic identity: Soft biopolitics and the modulation of control. *Theory, Culture & Society*, *28*(6), 164–181.

Chiang, J.-H., Yang, P.-C., & Tu, H. (2014, August). Pattern analysis in daily physical activity data for personal health management. *Pervasive and Mobile Computing*, *13*, 13–25. doi:10.1016/j.pmcj.2013.12.003

Choe, E. K., Lee, N. B., Lee, B., Pratt, W., & Kientz, J. A. (2014). Understanding quantified-selfers' practices in collecting and exploring personal data. In *CHI '14: Proceedings of the Conference on Human Factors in Computing Systems* (pp. 1143–1152). New York, NY: Association for Computing Machinery.

Consolvo, S., Klasnja, P., McDonald, D. W., Avrahami, D., Froehlich, J., LeGrand, L., . . . Landay, J. A. (2008). Flowers or a robot army? Encouraging awareness and activity with personal, mobile displays. In *UbiComp '08: Proceedings of the 10th International Conference on Ubiquitous Computing* (pp. 54–63). New York, NY: Association for Computing Machinery.

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. *European Journal of Cultural Studies*, *18*(4/5), 479–496.

Desrosières, A. (2001). How "real" are statistics? Four possible attitudes. *Social Research*, *68*(2), 339–355.

Dijck, J. van. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society*, *12*(2), 197–208.

Elliott, A. (2013). The theory of new individualism. In R. Tafarodi (Ed.), *Subjectivity in the twenty-first century: Psychological, sociological, and political perspectives* (pp. 190–209). Cambridge, UK: Cambridge University Press.

Ellison, N., Heino, R., & Gibbs, J. (2006). Managing impressions online: Self-presentation processes in the online dating environment. *Journal of Computer-Mediated Communication*, *11*(2), 415–441. doi:10.1111/j.1083-6101.2006.00020.x

Elsden, C., & Kirk, D. (2014). A quantified past: Remembering with personal informatics. In *DIS '14: Proceedings of the 2014 Conference on Designing Interactive Systems* (pp. 45–48). New York, NY: Association for Computing Machinery.

Epstein, D. A., Cordeiro, F., Bales, E., Fogarty, J., & Munson, S. A. (2014). Taming data complexity in lifelogs: Exploring visual cuts of personal informatics data. In *DIS '14: Proceedings of the 2014 Conference on Designing Interactive Systems* (pp. 667–676). New York, NY: Association for Computing Machinery.

Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology*, *113*(1), 1–40.

Espeland, W. N., & Stevens, M. L. (1998). Commensuration as a social process. *Annual Review of Sociology*, *24*, 313–343.

Espeland, W. N., & Stevens, M. L. (2008). A sociology of quantification. *Historical Social Research*, *41*(2), 401–436.

Estrin, D. (2014). Small data, where n = me: Seeking personalized data-derived insights from analysis of our digital traces. *Communications of the ACM*, *57*, 32–34.

Estrin, D., & Juels, A. (2016). Reassembling our digital selves. *Dædalus, 145*(1), 43–53.

Fiore-Gartland, B., & Neff, G. (2015). Communication, mediation, and the expectations of data: Data valences across health and wellness communities. *International Journal of Communication*, *9,* 1466–1484.

Foucault, M. (1988). *Technologies of the self*. Amherst, MA: University of Massachusetts Press.

Gebel, T., & Liebig, S. (2013). Die Dokumentation qualitativer Interviews im Rahmen von Organisationsstudien [Documentation of qualitative interviews in the context of organization studies]. In D. Huschka, H. Knoblauch, C. Oellers, & H. Solga (Eds.), *Forschungsinfrastrukturen für die qualitative Sozialforschung* [Research infrastructures for qualitative social research] (pp. 73–112). Berlin, Germany: Scivero.

Geertz, C. (1973). *The interpretation of cultures: Selected essays*. New York, NY: Basic Books.

Gitelman, L. (2013). *"Raw data" is an oxymoron*. Cambridge, MA: MIT Press.

Goffman, E. (1984). *The presentation of self in everyday life*. Harmondsworth, UK: Penguin Books. (Original work published 1959)

Hacking, I. (1990). *The taming of chance*. Cambridge, UK: Cambridge University Press.

Hand, M. (2016). Persistent traces, potential memories: Smartphone and the negotiation of visual, locative, and textual data in personal life. *Convergence*, *22*(3), 1–18.

Hayes, M. (2011). The social history of quantifying inflation: A sociological critique. *Journal of Economic Issues*, *45*(1), 97–112.

Hodges, S., Williams, L., Berry, E., Izadi, S., Srinivasan, J., Butler, A., . . . Wood, K. (2006). SenseCam: A retrospective memory aid. In *UbiComp '06: Proceedings of the 8th International Conference on Ubiquitous Computing* (pp. 177–193). Berlin, Germany: Springer.

Jensen, U. (2011). Datenmanagementpläne [Concepts of data management]. In S. Büttner, H.-C. Hobohm, & L. Müller (Eds.), *Handbuch Forschungsdatenmanagement* [Handbook of research data management] (pp. 71–82). Bad Honnef, Germany: Bock und Herchen.

Kim, J. (2014). Analysis of health consumers' behavior using self-tracker for activity, sleep, and diet. *Telemedicine and e-Health*, *20*(6), 552–558. doi:10.1089/tmj.2013.0282

Kranza, M., Möller, A., Hammerlac, N., Diewald, S., Plötzc, T., Olivierc, P., . . . Roalter, L. (2013). The mobile fitness coach: Towards individualized skill assessment using personalized mobile devices. *Pervasive and Mobile Computing*, *9*(2), 203–215. doi:10.1016/j.pmcj.2012.06.002

Li, I., Dey, A., & Forlizzi, J. (2010). A stage-based model of personal informatics systems. In *CHI '10: Proceedings of the Conference on Human Factors in Computing Systems* (pp. 557–566). New York, NY: Association for Computing Machinery.

Li, I., Dey, A., & Forlizzi, J. (2011). Understanding my data, myself: Supporting self-reflection with ubicomp technologies. In *UbiComp '11: Proceedings of the 13th International Conference on Ubiquitous Computing* (pp. 405–414). New York, NY: Association for Computing Machinery.

Lomborg, S., & Frandsen, K. (2016). Self-tracking as communication. *Information, Communication & Society*, *19*(7), 1015–1027. doi:10.1080/1369118X.2015.1067710

Lupton, D. (2014a). Self-tracking cultures: Towards a sociology of personal informatics. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The future of design* (pp. 77–86). New York, NY: Association for Computing Machinery.

Lupton, D. (2014b, August). *Self-tracking modes: Reflexive self-monitoring and data practices*. Paper presented at Imminent Citizenships: Personhood and Identity Politics in the Informatic Age, Canberra, Australia.

Lupton, D. (2015, July 28). *Personal data practices in the age of lively data*. Retrieved from http://ssrn.com/abstract=2636709

Marwick, A. (2012). The public domain: Social surveillance in everyday life. *Surveillance & Society, 9*(4), 378–393.

Mayer-Schoenberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. London, UK: John Murray.

Miller, P. (2001). Governing by numbers: Why calculative practices matter. *Social Research*, *68*(2), 379–396.

Mortier, R., Haddadi, H., Henderson, T., McAuley, D., & Crowcroft, J. (2014, October). Human-data interaction: The human face of the data-driven society. *Social Science Research Network*. Retrieved from http://ssrn.com/abstract=2508051

Nafus, D. (2013). The data economy of biosensors. In M. McGrath & C. N. Scanaill (Eds.), *Sensor technologies: Healthcare, wellness and environmental applications* (pp. 137–156). New York, NY: Apress.

Nafus, D. (2014). Stuck data, dead data, and disloyal data: The stops and starts in making numbers into social practices. *Distinktion: Journal of Social Theory*, *15*, 208–222.

Nafus, D., & Sherman, J. (2014). This one does not go up to 11: The quantified self movement as an alternative big data practice. *International Journal of Communication*, *8*, 1785–1794.

Pantzar, M., & Ruckenstein, M. (2015). The heart of everyday analytics: Emotional, material and practical extensions in self-tracking market. *Consumption Markets & Culture*, *18*(1), 92–109.

Peesapati, T.S., Schwanda, V., Schultz, J., & Cosley, D. (2010): Triggering memories with online maps. *Proceedings of the Association for Information and Technology*, *47*(1), 1-4.

Porter, T. M. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton, NJ: Princeton University Press.

Reigeluth, T. (2014). Why data is not enough: Digital traces as control of self and self-control. *Surveillance & Society*, *12*(2), 243–254.

Rooksby, J., Rost, M., Mossiron, A., & Chalmers, M. (2014). Personal tracking as lived informatics. In *CHI '14: Proceedings of the Conference on Human Factors in Computing Systems* (pp. 1163–1172). New York, NY: Association for Computing Machinery.

Rose, N. (1990). *Governing the soul: The shaping of the private self*. London, UK: Routledge.

Rose, N. (1991). Governing by numbers: Figuring out democracy. *Accounting, Organizations and Society*, *16*(7), 673–692.

Rose, N. (2007). *The politics of life itself: Biomedicine, power, and subjectivity in the twenty-first century*. Princeton, NJ: Princeton University Press.

Ruckenstein, M. (2014). Visualized and interacted life: Personal analytics and engagements with data doubles. *Societies*, *4*(1), 68–84.

Ruckenstein, M., & Pantzar, M. (2015). Datafied life: Techno-anthropology as a site for exploration and experimentation. *Techné: Research in Philosophy and Technology*, *19*(2), 193–212.

Selke, S. (2014). Lifelogging als soziales Medium? Selbstsorge, Selbstvermessung und Selbstthematisierung im Zeitalter der Digitalität [Lifelogging as a social media? Self concerns, self measurement and self thematization in the digital age]. In J. Jähnert & C. Förster (Eds.), *Technologien für digitale Innovationen. Interdisziplinäre Beiträge zur Informationsverarbeitung* (pp. 173–200). Wiesbaden, Germany: Springer.

Steele, R. (2013). Utilizing social media, mobile devices and sensors for consumer health communication: A framework for categorizing emerging technologies and techniques. In S. C. Mukhopadhyay & O. A. Postolache (Eds.), *Pervasive and mobile sensing and computing for healthcare* (Vol. 2, pp. 233–249). Berlin, Germany: Springer.

Stragier, J., & Mechant, P. (2013, February). Mobile fitness apps for promoting physical activity on Twitter: The #RunKeeper case. Paper presented at Etmaal van de Communicatiewetenschappen, Rotterdam, Netherlands.

Swan, M. (2013). The quantified self: Fundamental disruption in big data science and biological discovery. *Big Data*, *1*(2), 85–99.

Till, C. (2014). Exercise as labour: Quantified self and the transformation of exercise into labour. *Societies*, *4*(3), 446–462. doi:10.3390/soc4030446

Turner-McGrievy, G. M., Beets, M. W., Moore, J. B., Kaczynski, A. T., Barr-Anderson, D. J., & Tate, D. (2013). Comparison of traditional versus mobile app self-monitoring of physical activity and dietary intake among overweight adults participating in a health weight loss program. *Journal of American Medical Information Association*, *20*(3), 513–518.

Wang, J., Wang, Y., Wei, C., Yao, N., Yuan, A., Shan, Y., & Yuan, C. (2014). Smartphone interventions for long-term health management of chronic diseases: An integrative review. *Telemedicine and e-Health*, *20*(6), 570–583. doi:10.1089/tmj.2013.0243

Whitson, J. R. (2013). Gaming the quantified self. *Surveillance & Society*, *11*(1/2), 163–176.