

# The Construction of Meaning through Digital Traces

Joschka A. Hüllmann, Münster University, huellmann@uni-muenster.de

## Changing Nature of Work

Contemporary work has seen a shift towards project-based forms of organizing, as spanning temporal and spatial boundaries becomes a common experience. Employees, dispersed across the globe and time zones, serve on multiple teams, which are rapidly staffed to collaborate for a short period of time (Barley & Kunda, 2001; Wageman, Gardner, & Mortensen, 2012). The variance in schedule, location, tasks, tools, colleagues, and hierarchy, as well as its dynamics, increases the complexity of work, leading to a heterogeneous network of relationships, and a fragmented task and social environment (Espinosa, Cummings, Wilson, & Pearce, 2003). In such a workplace employees work together through extensive use of communication and collaboration tools (Espinosa et al., 2003; Orlikowski & Scott, 2016), which generates digital traces (Barley & Kunda, 2001; Behrendt, Richter, & Trier, 2014). I question how the analysis of these digital traces complements existing methods of inquiring the changing nature of work by presenting results from two pilot studies.

Previous research on the changing nature of work used a variety of methods of data collection, including participant observations, content analyses, semi-structured interviews, surveys (e.g. Espinosa et al., 2003), or video recordings (e.g. Poels, Tucker, & Kielem, 2017). Specifically the rich and context-sensitive approaches, such as observations and ethnographic accounts, show limitations, when reliable observations of people are difficult, who perform computer-based work or are part of dispersed teams (Orlikowski & Scott, 2016). Surveys—and in part—interviews provide snapshots, but lack a rich description of context, because people do not know what they do “*unless they are in the process of doing it*” (Barley & Kunda, 2001) and subjective perceptions may be skewed (Orlikowski & Scott, 2016). In response, Barley & Kunda (2001) call for research about the changing nature of work on a micro-level based on rich longitudinal data that is embedded into the task environment and social context. They suggest digital traces, e.g. login and logoff events, email or chat messages, and editing and saving documents, as a complement to traditional data. “*Most work practices involve digital technology*” (Orlikowski & Scott, 2016) suggests that digital traces can provide fruitful grounds for data collection and yield rich insights into digital work activities. Digital traces depict the actions and interactions an employee performs on devices such as the computer or smartphone. They are generated through routine use of technology and potentially cover the complete history of using a particular software or device. Digital traces enable exploratory and explanatory research. On the one hand, exploratory research into digital traces describes and visualizes the data, and allows to identify patterns (e.g. Riemer, Lee, Kjaer, & Haefner, 2018)—either manually or based on statistical analysis (e.g. clustering). On the other hand, explanatory research derives metrics from the digital traces to systematically test hypotheses related to work activities (e.g.

Riemer, Finke, & Hovorka, 2015). Although digital traces are not the only source for investigating such questions, they demonstrate distinctive benefits and drawbacks. Their collection does not inconvenience the employees, because researchers can extract data from hosted locations without end-user interaction, e.g. software-as-a-service (“cloud”) or on premise solutions. The data is a by-product of using these tools (see also “*persistence*” in Treem & Leonardi, 2012). For local data, the researchers can deploy software sensors for a selected period of time to collect data from the employees’ devices passively at any time or location (Table 1). Software sensors are small applications or services that run in the background on a local device and hook into other applications to collect activity data on a fine-granular level. This activity data can include every single computer action, down to a single click or keystroke, or recordings with the camera or microphone. In combination, the continuous data collection on technology use and the extensive use of technology in performing work, result in a detailed depiction of employees’ overall activities (Begole, Tang, Smith, & Yankelovich, 2002)—contingent on a set-up spanning relevant devices. Such a set-up is highly invasive, potentially yielding deep insights into how the nature of work changes, while simultaneously raising privacy concerns. So far, I only find studies making use of isolated sources such as enterprise social networks, but very few, if any comprehensive studies.

Data Source	Extraction Method	Examples
Hosted Data	Structured Content Export	Enterprise Social Networks, Instant Messaging, Email, Access Logs
Local Data	Software Sensors	Email, Access Logs, Video-/Audio records, Keystrokes

Table 1 – Data Sources.

### Analyses of Digital Traces

While analyzing digital traces is widely adopted in research on relational social structures—which serves me as an example—I hardly find any such papers on temporal structures, or collective media choice. For this reason, our research group piloted two studies on temporal structures and media choice to examine the utility and value of digital traces in these domains. I briefly summarize extant research on the three topics, before I report on our two studies (Table 2).

The research on **informal relational structures** has shifted from manually collected network data to the extensive use of digital traces, which researchers extract from communication and collaboration tools such as email, instant messaging, or enterprise social networks and that are prevalent in organizations nowadays (Wehner, Ritter, & Leist, 2017). In this case, the digital traces consist of interaction data between employees, informing various strands of research including—amongst many others—knowledge sharing (Mäntymäki & Riemer, 2016), social capital (Riemer, Finke, et al., 2015), social onboarding (Hüllmann & Kroll, 2018), or informal hierarchy (Riemer, Stieglitz, & Meske, 2015). Typical methods of inquiry include social network analysis, content analysis, and natural language processing (Behrendt et al., 2014). Begole et al. (2002) study **temporal rhythms** by looking

at the activity levels of computer workers, recording activity and calendar logs. Fisher & Dourish (2004) examine email trace data to identify recurring temporal patterns. Others extract, visualize and cluster temporal rhythms of email usage (Perer, Shneiderman, & Oard, 2006; Viégas, Boyd, Nguyen, Potter, & Donath, 2004). I was unable to identify studies on **collective media choice** based on digital traces. Yet, insights may result from the quantification of media collections, longitudinal inquiries on changes in heterogeneous media use, identification of media co-use, or investigations of cause and effect for varying types of media collections.

Our pilot studies use a combination of digital traces and semi-structured interviews to evaluate and illustrate the analysis of digital traces. Table 2 briefly summarizes respective project sites, data sources and research questions. In “small case”, we identify start, end, and break times of the day, as well as holidays and regular weekend work using digital traces. In “big case”, we identify two types of media collections: social focus, and file sharing focus. In subsequent tests, we find that the organizational factors “assigned manager”, and “functional department” significantly influence employees’ media choice (tested with  $\chi^2$ ).

	<b>Small Case Temporal Rhythms</b>	<b>Big Case Media Choice</b>
Site	Co-located scholars and product managers (N = 5)	Individuals in dispersed virtual teams of IT service provider (N = 1,400)
Source	Local: 13,062 sent emails (2.5 years), and semi-structured interviews	Hosted: monthly usage frequencies of Office365 (8 months), and Active Directory data
RQ	How do employees structure their day and week?	Which media collections are in use? What are drivers of media choice?

Table 2 – Pilot Studies.

## Lessons Learned

First, the data collection is potentially subject to concerns on technical and organizational levels. We had issues in “big case”, where responsibilities are widely shared across departments and data collection must be coordinated with legal and technical administration. Especially in large organizations, the research sponsor’s leverage over these departments varies, as they are centralized and serve the entire organization. In the “big case”, it took seven months from legal approval to provision of access to data provided by the technical administration. An alternative to getting institutional approval is to go for individual consent to deploy non-intrusive software sensors on relevant devices. Technical obstacles occur, when the information system at the organization does not support any data extraction out of the box, or multiple data sources must be extracted and integrated. As a result, we had to implement a manual data extraction, transformation, and loading process for the “small case”.

Second, digital trace data only shows activities performed at the computer and thus always is agnostic about the non-digital activities of the employees. Depending on the role and task, the extent of digital work varies considerably across employees, possibly incurring bias in subsequent analyses. The absence of data is as relevant as the captured data and

researchers must cautiously consider and interpret any lack of digital traces. Complementary data collection may be necessary, e.g. via surveys, interviews, recordings, or socio-metric badges to explain observations (cf. Venkatesh, Brown, & Bala, 2013). For example, in the “small case”, we used interviews to make sense of the gaps in the digital trace data.

Third, the quality of data may vary in terms of breadth (“number of features”), depth (“level of aggregation within features”), or number of observations. In the “big case”, we worked with a high number of observations, but the breadth and depth were too low (monthly aggregates) to investigate the micro level and how work is performed. The data can be extended by means of surveys or interviews, which, however, can only cover small samples or a subset of the data. Conversely, breadth and depth can be high with a small sample size, in which case deep insights are possible, but the generalizability of the findings is limited.

Fourth, invasive data collection may reveal rich insights, but comes at the cost of increased privacy concerns. As digital traces capture all electronic actions of employees, automated and large-scale collection is under scrutiny for privacy and surveillance (Ball, 2010; Begole et al., 2002), and legal policies can prevent the extraction of data. Researchers can mitigate the privacy concerns by being transparent about the purpose and extent of data collection (Pachidi, Huysman, & Berends, 2016). Because the interest lies in generalizable findings, anonymization and limiting the period of data collection may reduce privacy concerns. In the “big case”, we worked with anonymous data, and in the “small case”, we anonymized the recipients and content of the emails, but sought personal contact for interviews.

Fifth, transparency about the extent of data collection induces behavioral change. While this is less an issue for the extraction of (historical) traces from hosted sources, it is relevant for deploying sensors for short periods of time. Impression management and attempts to game the system skew the patterns in the digital traces (Pachidi et al., 2016).

Sixth, the social and task context when analyzing digital traces is limited. However, researchers can infer the social context by analyzing interaction partners and the content of interactions, manually, or automatically via a combination of social network analysis and natural language processing (e.g. Cetto, Klier, Richter, & Zolitschka, 2018). The task context can be inferred via meta data of the digital traces, for example, associated project or type of action. In “big case”, we inferred part of the context by retrieving the associated department, team members, and manager. In “small case”, we conducted interviews to address open questions.

In summary, the analysis of digital traces has seen wide application in investigating informal relational structures, with prospects for the analysis of temporal structures and collective media choice. Working with digital traces is minimally intrusive and maximally invasive. Researchers can collect data that details work activities on the micro level and enables investigations across the boundaries of the digital workplace, without inconveniencing employees. Despite the outlined challenges, I believe that the analysis of digital traces will prove valuable to construct meaning for the changing nature of work.

## References

- Ball, K. (2010). Workplace surveillance: an overview. *Labor History*, 51(1), 87–106.
- Barley, S. R., & Kunda, G. (2001). Bringing Work Back In. *Organization Science*, 12(1), 76–95.
- Begole, J. “Bo,” Tang, J. C., Smith, R. B., & Yankelovich, N. (2002). Work rhythms. *Proc. of the ACM Conference on Computer Supported Cooperative Work*.
- Behrendt, S., Richter, A., & Trier, M. (2014). Mixed methods analysis of enterprise social networks. *Computer Networks*, 75, 560–577.
- Cetto, A., Klier, M., Richter, A., & Zolitschka, J. F. (2018). “Thanks for sharing”—Identifying users’ roles based on knowledge contribution in Enterprise Social Networks. *Computer Networks*, 135, 275–288.
- Espinosa, J. A., Cummings, J. N., Wilson, J. M., & Pearce, B. M. (2003). Team Boundary Issues Across Multiple Global Firms. *Journal of Management Information Systems*, 19(4), 157–190.
- Fisher, D., & Dourish, P. (2004). Social and temporal structures in everyday collaboration. *Proc. of the Conference on Human Factors in Computing Systems*, 6(1), 551–558.
- Hüllmann, J. A., & Kroll, T. (2018). The Impact of User Behaviours on the Socialisation Process in Enterprise Social Networks. *Proc. of the Australasian Conference on Information Systems*.
- Mäntymäki, M., & Riemer, K. (2016). International Journal of Information Management Enterprise social networking: A knowledge management perspective. *International Journal of Information Management*, 36(6), 1042–1052.
- Maznevski, M. L., & Chudoba, K. M. (2000). Bridging Space over Time: Global Virtual Team Dynamics and Effectiveness. *Organization Science*, 11(5), 473–492.
- Orlikowski, W. J., & Scott, S. V. (2016). Digital Work: A Research Agenda. In *A Research Agenda for Management and Organization Studies* (pp. 88–96).
- Pachidi, S., Huysman, M., & Berends, H. (2016). Playing the numbers game: Dealing with transparency. *Proc. of the 37th International Conference on Information Systems*, 1–18.
- Perer, A., Shneiderman, B., & Oard, D. W. (2006). Using Rhythms of Relationships to Understand E-Mail Archives. *Journal of the American Society for Information Science and Technology*, 57(14), 1936–1948.
- Poels, T., Tucker, D. A., & Kielema, J. (2017). The development of a theoretical framework of organisational rhythm. *Journal of Organizational Change Management*, 30(6), 888–902.
- Riemer, K., Finke, J., & Hovorka, D. S. (2015). Bridging or Bonding: Do Individuals gain Social Capital from Participation in Enterprise Social Networks? *Proc. of the 36th International Conference on Information Systems*, 1–20.
- Riemer, K., Lee, L. L., Kjaer, C., & Haeffner, A. (2018). Metrics selection for group type identification in Enterprise Social Network (ESN) analytics. *Proceedings of the 29th Australasian Conference on Information Systems (ACIS)*, 1–12.
- Riemer, K., Stieglitz, S., & Meske, C. (2015). From Top to Bottom: Investigating the Changing Role of Hierarchy in Enterprise Social Networks. *Business and Information Systems Engineering*, 57(3), 197–212.
- Treem, J. W., & Leonardi, P. M. (2012). Social Media use in organizations. *Communication Yearbook*, 143–189.
- Venkatesh, V., Brown, S. A., & Bala, H. (2013). Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems. *MIS Quarterly*, 37(1), 21–54.
- Viégas, F. B., Boyd, D., Nguyen, D. H., Potter, J., & Donath, J. (2004). Digital artifacts for remembering and storytelling: PostHistory and social network fragments. *Proc. of the Hawaii International Conference on System Sciences*, 37, 1753–1762.
- Wageman, R., Gardner, H., & Mortensen, M. (2012). The changing ecology of teams: New directions for teams research. *Journal of Organizational Behavior*, 33(3), 301–315.
- Wehner, B., Ritter, C., & Leist, S. (2017). Enterprise social networks: A literature review and research agenda. *Computer Networks*, 114, 125–142.