

Data Work in Healthcare: An Ethnography of a BI Unit

Asbjørn Malte Pedersen, Claus Bossen
Information Studies, Aarhus University, Denmark
ampedersen@cc.au.dk; clausbossen@cc.au.dk

Abstract. In this paper, we investigate a newly established Business Intelligence unit providing healthcare data and analyses for management and staff at five hospitals in the Region. Created in 2015, the Business Intelligence Unit repurposes data generated by digital healthcare systems and aims to “[...] support the Region [...] in delivering more welfare, better quality, higher impact, and greater sustainability for less money.” Within healthcare – and in other domains – there is a strong push towards becoming (more) data-driven and repurpose data to increase efficiency, quality, and cost-effectiveness. However, whereas there are numerous publications on ‘datafication’ in the abstract, there is a dearth of studies on how the data at the center of such processes is actually produced, and even fewer studies on the people and organizational units engaged in this work. Hence, we are engaged in an ethnographic study on data work at the Business Intelligence Unit in order to gain insights into the work and practices of generating healthcare data for secondary use. In this paper, we present the preliminary findings focusing on three themes: end-user engagement, creating meaningful data reports, and establishing trust. One overall contribution is that whereas Business Intelligence work does require technical competences to work with large-volume data, it also requires competences of engaging with healthcare staff and understanding their work practices.

Introduction

A second wave of healthcare digitization has begun after the first ‘wave broke’ and electronic healthcare records (EHR) were implemented in the first two decades of this century in many countries in Europe and the Americas (Simborg et al. 2013). In the present ‘second wave’, the digital data produced by EHRs and other

healthcare IT systems are repurposed for secondary uses such as managerial overview, quality assessment, optimizing workflows, etc. This repurposing of healthcare data is at the center of numerous initiatives at local and government levels and within commercial companies.

However, whereas there is much attention upon the larger processes and the involved benefits and challenges of healthcare digitization (e.g., Hogle 2016, Hoeyer 2016, Ruckenstein & Dow 2017), there is a dearth of studies on the people working with that data. While the occupation of ‘data scientist’ has emerged and received attention and even led to new university programs, we know little about who is actually doing the work of collecting, filtering, formatting, analyzing and presenting healthcare data. Data work is often invisible, rendered to the background, or abstracted away as numbers and indicators (Star & Strauss 1999), and thus researchers, managers, and policy makers are provided an impoverished understanding of what it means and requires to become data-driven in healthcare and deliver on it. Furthermore, datafication entails shifts in occupations and professions that may change these or lead to the emergence of new ones. Hence, one overall research question to us is what do the actual work practices in healthcare around data look like?

In order to investigate this, we have commenced upon a long-term ethnographic investigation of a data work site in order to understand what ‘data work’ entails. The site is a Business Intelligence unit (BIU) providing data for hospital and ward management as well as regional politicians. In this paper, we present our first, preliminary findings.

Datafication and Data Work in Healthcare

Our research into data work practices is situated in the larger context of datafication and the ambitions for public and private organizations to become data-driven. ‘Data’ has emerged as the new oil of the information economy (Van Es & Schäfer 2017), and as the central notion of Big Data (Mayer-Schönberger & Cukier 2013, Kitchin 2014). Premised on the spread in depth and scope of digital technologies, huge amounts of data are accumulated, algorithmically transformed, and repurposed to deliver purported benefits for citizens, organizations, and society. However, data is not ‘found’ and ‘collected’, but must be generated, shaped and presented – by people. As the emerging field of critical data studies argues, ‘raw data is an oxymoron’ (Gitelman 2013). Further, sharing data creates ‘friction’, since it has to be filtered, formatted and re-contextualized (Edwards 2010). Data requires work. Some of this work involves algorithms and high-level technical skills, but also hermeneutic judgement and situated knowledge is required, since data is entangled with the contexts of its production and use (boyd et al. 2012).

Ambitions for data-driven healthcare in many countries (Madsen 2014, Raghupathi et al. 2014, Mayer-Schönberger & Ingelsson 2018) has followed

digitization of the healthcare domain (Ruckenstein & Dow 2017). Those ambitions are also strong in Denmark, where the present study is conducted, and where ‘data-intensive resourcing’ is pursued in government strategies (Hoeyer 2016). Amongst healthcare staff, this has at times led to a feeling of being flooded with requirements to provide data to the extent that ‘meaningless data’ erodes professional judgement (Wadman & Hoeyer 2020). Whether data is actually ‘meaningless’ depends of course on perspective and on whether one understands and agrees with purpose of generating it (ibid).

Surprisingly few studies on data work have been published. ‘Data scientists’ have received attention as a new, emerging profession, but beyond that there are few studies upon who conducts data work. A couple of recent studies show how generating and making data useful is also conducted by professions such as electricians and energy advisors (Foster 2016, Fischer et al. 2017, Kristiansen et al. 2018). Also, within healthcare, an emerging attention to data work can be seen by studies that examine the actual work invested in producing, managing, analyzing, and deploying data (Bjørnstad & Ellingsen 2019, Bonde et al. 2019, Grisot et al. 2019, Islind et al. 2019, Pine 2019). Also, new emergent data work occupations such as ‘medical scribes’ and ‘clinical documentation integrity specialists’ and their role in producing healthcare data have been studied (Bossen et al. 2019, Pine & Bossen 2020). Most recently, Håland & Melby (2021) analyzed the work of coding cancer care pathways and identified five types of ‘accountability work’ which we see as an exciting analysis of what we call data work. Hence, gradually a better understanding of data work practices in healthcare is emerging.

However, we know relatively little about the work going on within BIUs in healthcare. Early frameworks for the links between Business Intelligence (BI) and healthcare were presented by Mettler & Vimarlund (2009) and Wixom & Watson (2010), and Madsen (2012) presents guidelines for healthcare organizations to become data-driven, including how to achieve overview of the availability of data, the need for standardization, how to get the data warehouse in order, the work of Chief Data Officer, etc., but there is little concrete information on the data work and workers themselves. Gaardboe et al. (2017) applied DeLone & Mclean’s model of IT system success to investigate the factors that contribute to BI success and found user satisfaction to be decisive. In a similar study, they identified two groups of users of BI information - system users and information users – and that BI was used for financial reporting, improving patient progress, and enhancing learning in hospitals (Gaardboe et al. 2018). Manchini et al. (2019) provide a review of 50 papers from 2010 to 2017 on BI in healthcare based on a framework of four BI capabilities: organizational memory, information integration, insight creation, and presentation capabilities. However, none of these studies provide a detailed analysis of BI work practices.

In summary, there is a need for ethnographic studies of data work as well as of BIUs in order to better understand what data work and datafication is.

Methods

The present paper builds on participant observation by the first author, who participated in three online staff meetings (2 hours each), conducted four interviews (½ hour each - including three BIU staff members and a manager), one interview with two managers (1 hour), an introductory presentation by a BI representative (2 hours). Further, we have conducted analysis of a wide range of documents: Policy documents by the Region outlining the purpose of the BIU, and various internal BI documents describing work processes, the unit's strategies, etc. Additionally, a significant number of online videos for healthcare staff produced by the BIU explaining how to work with data have been watched.

Notes were taken during participant observation, interviews were transcribed and subsequently analyzed, and documents scrutinized for the purpose of this paper.

Findings

The BIU was established in 2015 as a part of the Region's digitization strategy to make work procedures in healthcare more efficient and data-driven, as well as accommodate the demand for evidence-based decision-making. The BIU is also a part of the vision to proactively utilize data for forecasting with the help of artificial intelligence and machine learning. The BIU's own strategic vision is to better the "[...] organization, services, resource consumption, impact, and reputation by always delivering the best qualified data-information basis for decision-making and development." (internal BI strategy document). In addition, their mission is to support the Region's areas of business "[...] in delivering more welfare, better quality, higher impact, and greater sustainability for less money." (internal BI strategy document).

The BIU has developed rapidly. Whereas it had 16 employees and 290 users in 2015, the BIU has since grown to 45 employees and nearly 4000 users in the beginning of 2021. A central component in their strategy is to have an 'open BI-setup', which means that they actively pursue user participation by healthcare staff and management in their service delivery.

The BIU staff hold several job titles such as BI developer, BI architect, senior developer, and user experience designer. Most common – including our interviewees – the staff members hold the title of BI developer. In general, BI developers have broad profiles encompassing both technical skills (i.e., front- and back-end programming) and competencies such as working with large-scale datasets, identifying user needs as well as solving problems with data. Formally, their work covers data warehousing, data analysis, report development, support, advice, and education of other occupations in healthcare. Their main product is data

reports, which have a standardized format of presenting data as well as showing which data and which filters were applied.

Open BI: User Participation, BI Portal Ninjas, and Data Heroes

The BIU's notion of an open BI-setup implies a user-centered approach, where they aim to be both inclusive and relevant to healthcare staff, administration, and management. It involves the following elements: User participation when developing new data products; BIU staff members and technology being easily accessible when needed; and training healthcare staff in data processing and use. Hence, the BIU seeks to transform healthcare workers into "smart data workers", as one BIU staff member phrased it. The purpose is three-fold: To ensure that data is valid and data reports meaningful; to ensure implementation of data products in healthcare practices; and, finally, to empower healthcare staff to work with and use data. Thus, the BIU staff does a lot of work to make themselves relevant for healthcare staff.

The open BI approach was applied from the beginning, when healthcare staff was included in the development of the 'BI-portal', the central portal through which healthcare staff and the BIU can access and work with data. Also, BI developers engage in iterative processes with healthcare staff when they develop new data reports in order to identify user needs, receive feedback, develop new ideas, and improve the data products. This enables BI developers to understand healthcare staff's context, and thereby ensure that the reports provide meaningful and valid data as well as being useful to them. In the BIU's own view, this is a crucial part of their work, as it is a way to gain healthcare members' trust in data and the BIU's work – an argument we will elaborate further in later sections. This practice is also a way to ensure the implementation of their data products in local healthcare practices. That data is put into actual use. One way to facilitate this is to design the data reports in the BI-Portal to be user-friendly. As much as possible, they should be self-explanatory and easy to navigate, thus reducing complexity and thereby reduce time spent on learning to read data reports and using the BI-Portal. This can potentially enable both busy as well as technically lower-skilled staff members to utilize the data reports.

The BIU also trains healthcare staff to become BI users and 'smart data workers', through workshops such as BI-Garages, SQL workshops as well as BI-Cafés. At BI-Garages, the BIU invites healthcare staff members into their office and introduces them to the BI-portal and the data reports. At the SQL workshops, healthcare staff learns how to use SQL to manage healthcare data. Furthermore, the BIU has developed two training programs, the 'BI-portal Ninja education' and the 'Data Hero education'. Through the BI-portal ninja education, administrative personnel and healthcare staff learn about the BI-portal and how to operate existing data reports. The participants at these training programs are typically medical secretaries and healthcare staff members who manage data from data reports such

as diagnostic reports or treatment time reports. The Data Hero training program is aimed at medical staff like clinicians and associated medical professions (e.g., biochemists, chief physicians, radiologists, etc.), administrative officers, and data workers at local wards. In this intensive three-week program, participants learn to create their own specifically focused data reports from existing data resources, and how to visualize data to support their own practices and needs. Further, they learn about Tableau, (a data-processing and visualization program), data sets, key figures, and the challenges involved in granularity.

The open and user-involving approach can be seen as an intervention where the BIU distributes data competencies to healthcare staff. This aims to empower healthcare staff to understand data reports, device their own when needed, make them trust data, and see how it can be applied to make their work procedures more efficient or improve quality. However, another point is that this intervention is also a way to consolidate the BIU's own work. When they work to make data valid and meaningful, make themselves visible through workshops as well as distribute data competencies, they work to get healthcare staff to accept data work as important healthcare work.

Technical Skills: From Data to Meaningful Information

“To give data back to those who have fed in data in the first place”. That is how one of the BI developers explained the purpose of his work. Others offered us variations on the same theme: Providing meaningful information in the right place at the right time. More elaborately, their aim is to support healthcare staff - whether it is clinicians, administration personnel, or management - with data for decision-making and thus make it possible to improve work procedures as well as patient treatments and experiences. To accomplish this purpose, BI developers work to transform data into what they call “meaningful information” as well as to present it in a “meaningful way”. To make data meaningful is, we purport, a sociotechnical endeavor that requires technical skills, specific knowledge, and collaboration between different professions. In this section, we will illustrate how BI developers utilize technical skills and technologies to transform data into meaningful information and thus improve work procedures.

BI developers work closely with data in a technical sense and engage in data warehousing, data analysis, and data visualization. One central work process is to work with different, complex layers of the data structure that must be applied properly to enable data to function and be put to work. BI developers integrate data into the data warehouse from different internal sources (i.e., EHR, clinical quality databases, salary systems, etc.). They aggregate data into data sets and apply metadata; add different logics such as key figures and definitions; and apply visualizations on top of it and security beneath it, so the data is ready for presentation. A core task of their function is then to further process and visualize

these data in a meaningful way through data reports that are published on the BI-portal where healthcare staff can access them. These data reports are developed with Tableau software that has features enabling filtering and visualization of data through tables, statistics, charts, and graphs.

An example of this process is the ‘Turn-Around Time Report’ (TAT)’. This report was developed to support a biochemist’s practice of monitoring his unit’s turn-around time: How long does it take from the arrival of a sample to an answer is available in the laboratory system? Depending on the priority of a sample, this can be measured in minutes, hours, or perhaps even weeks. Previously, the biochemist would once a week manually enter the unit’s data into a spreadsheet to monitor its turn-around time. After consultation with a BI developer, this practice was transformed into the TAT, where the relevant data is automatically integrated on daily basis and presented in a way that is useful and meaningful to the biochemist. When needed, the report can effectively provide answers to questions that correlates with the unit’s objectives (e.g., how many samples are analyzed within 60 minutes?), and thus saves the biochemist time.

Of course, technical skills are imperative when automating redundant, time-consuming tasks, but this example also demonstrates how mutual collaboration between healthcare staff and BI developers in line with the open BI strategy results in relevant data that is put into use. Thus, the BIU does not only provide meaningful information and make healthcare staff’s tasks more efficient, but also give data back to them, so data work seems more meaningful in the first place.

Collaboration and Domain Knowledge: Data Quality and Trust

As can be seen from the above, a crucial aspect of BI developers’ work is to collaborate closely with healthcare staff to ensure meaningful data of high quality. Data quality is important to both the BIU staff members and healthcare staff: If errors occur in data or data sets, they can result in healthcare staff distrusting data validity. Consequently, this can prompt them to not use the data products and instead solely rely on their professional judgement. To meet this challenge, BI developers strive to achieve specific domain knowledge in order to be able to correct errors in collaboration with healthcare staff, and hence ensuring that data is valid and meaningful.

Data quality and meaningful data often pose a challenge to the BIU, when end-users have not been involved in development from the outset. When “[...] the one who defines key figures is not necessarily the one who must use it in a given situation”, as one BI developer phrased it, the final data product can seem flawed to the end-user or not compatible with the actual work practice it should support. This is evident in situations where there is a difference between the registered data presented in a report and healthcare staff’s experienced reality. For example, a medical secretary was looking up how many of her patients were about to exceed

the limit of the warranted treatment time (In Denmark, per government regulation all patients must have a treatment plan within 30 days). In this case, the report displayed three patients that the medical secretary knew had been treated already. So, the question for her was, how many other errors could there be in the data report? According to the BI developer, who investigated the data, the answer was ‘none’, but it is more complex as such. As the BI developer explained, these kind of ‘errors’ occur because of differences in registration practice; differences in key figures, definitions, or codes; or different understandings of the usage situation. In this case, the three patients were registered with a wrong code at the local ward. The BI developer explained that someone had decided that they should use another code. The medical secretary objected to that code being ‘wrong’, because they had always used that code at her ward. The BI developer further elaborated on the situation: “[...] it is that kind of logic that is hard to communicate, that is, the complexity that is in this logic that we have in our data”.

If not resolved, situations like these can give rise to distrust in data validity and consequently cause specific data reports not to be used at all. So, in addition to having certain skills and knowledge of data analysis, BIU staff also need to have what one of the staff members characterizes as “domain knowledge”. Some of the BI developers have educational backgrounds in medicine and medical science, which provides for an initial understanding of the context and thus the data needed. However, to achieve the relevant knowledge for a specific work procedure, they must collaborate closely with healthcare staff and acquire experience within the domain. Thus, some BI developers shadow clinicians and take notes of their work procedures as well as sit next to them and go through all their tasks to identify the clinicians’ needs when they develop new data products:

“[We] go out and sit with the physician, when he has time, and talk about it and show him what is accessible. We ask him, do these numbers, these 27 patients we have found with bowel disease in your ward, do they sound right? Because we have no clue. We have no chance of knowing, if these 27 patients are 7 right ones and 20 wrong ones and if perhaps another 100 patients are missing” (Interview, BI developer 2)

As this example illustrates, the registered data might not be wrong, but the applied logics or visualization might differ from the medical staff’s experienced reality. To ensure data validity, and thus trust, BI developers move back and forth between their office and the wards and offices of healthcare staff to constantly test and validate the data quality. This shows how collaboration between BI developers and healthcare staff is crucial to ensure high data quality as it depends on the interdisciplinary knowledge of both data structuring and data context.

Discussion

This paper presents the preliminary findings from an ethnographic study of data work in a healthcare BIU that provides data to five regional hospitals and decision makers. Based on interviews with BIU staff and management, observations at staff meetings, and document research, we identify and analyze what their work entails.

First, we have analyzed how the BIU employs a strategy of user participation, called the ‘open BI setup’ to transform healthcare workers into ‘smart data workers’. This approach implies involving healthcare staff when developing new data products, making data products easier to understand and use, and training healthcare staff into being BI users. We have found that the purposes of this strategy are to ensure data is valid and meaningful; to ensure implementation of data products in healthcare practices; and to distribute data competencies and thus empower healthcare staff to work with and use data. Further, we have argued that the BIU works to get data work accepted as important as well as meaningful and thus to make the BIU relevant to healthcare staff. This part of their practice can also be understood in relation to Håland & Melby’s (2021) notion of ‘legitimation work’ where work is done to legitimize a practice “[...] as being important and make their colleagues accept this as important [...]” (Ibid, p9). When doing this, they seek to consolidate their own work as well as ensure the implementation of data products.

Second, we have demonstrated how BI developers’ data work practices revolve around transforming data into meaningful information presented in an easily understandable way. This involves transforming tedious data tasks into automated and standardized data reports. Furthermore, we have argued that this is a work practice that requires technical skills, specific knowledge of the domain, and collaborative competencies by BIU staff.

Third, we have pointed out that high data quality and meaningful data is crucial to avoid mistrust towards data amongst healthcare workers. When errors occur, it can lead to mistrust of data when alleged errors are due to differences in registration practices; differences in key figures, definitions, or codes; or different understandings of the usage situation. We have found that BI developers counteract these challenges by acquiring domain knowledge and collaborating closely with healthcare staff to account for errors as well as to correct them. We found that this work is important to ensure implementation of data products in healthcare practices.

This article contributes to an empirical understanding and insight into the practices of data work in healthcare and is the first step at an ethnographic analysis of data work practices in a healthcare business intelligence unit, which aims at providing insight into the work of becoming data-driven.

Acknowledgments

Our warm thanks to the management and staff at the BIU for their time and generosity.

This research is supported by the Independent Research Fund Denmark, Grant 0132-00200B “Making data work visible”.

References

- Bjørnstad, Camilla and Ellingsen, Gunnar (2019): ‘Data work: A condition for integrations in health care’, *Health informatics journal*, vol. 25, no. 3, pp. 526-535.
- Bonde, Morten, Bossen, Claus, and Danholt, Peter (2019): ‘Data-work and friction: Investigating the practices of repurposing healthcare data’, *Health informatics journal*, vol. 25, no. 3, pp. 558-566.
- Bossen, Claus, Chen, Yunan, and Pine, Kathleen H. (2019): ‘The emergence of new data work occupations in healthcare: The case of medical scribes’, *International journal of medical informatics*, vol. 123, pp. 76-83.
- boyd, danah, and Crawford, Kate (2012): ‘Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon’, *Information, communication & society*, vol. 15, no. 5, pp. 662-679.
- Edwards, Paul N. (2010): *A Vast Machine. Computer Models, Climate Data, and the Politics of Global Warming*, The MIT Press.
- Gaardboe, Rikke, Sandalgaard, Niels, and Nyvang, Tom (2017): ‘An assessment of business intelligence in public hospitals’, *International Journal of Information Systems and Project Management*, vol. 5, no. 4, pp. 5-18.
- Gaardboe, Rikke and Svejvig, Per (2018): ‘Better and more efficient treatment: The individual and organizational impacts of business intelligence use in health care organizations’, *Selected papers of the IRIS*, vol. 9, pp. 1-7.
- Gitelman, Lisa (2013): *Raw data is an oxymoron*, MIT press.
- Hoeyer, Klaus (2016): ‘Denmark at a crossroad? Intensified data sourcing in a research radical country’, in Mittelstadt, Brent D. & Floridi, Luciano (eds): *The Ethics of Biomedical Big Data*, Dordrecht, Springer, pp. 73–93.
- Hoeyer, Klaus and Wadmann, Sarah (2020): ‘“Meaningless work”: how the datafication of health reconfigures knowledge about work and erodes professional judgement’, *Economy and Society*, vol. 49, no. 3, pp. 433-454. <https://doi.org/10.1080/03085147.2020.1733842>
- Hogle, Linda F. (2016): ‘Data-intensive resourcing in healthcare’, *BioSocieties*, vol. 11, no. 3, pp. 372-393.
- Isazad Mashinchi, Mona, Ojo, Adegboyega and Sullivan, Francis J. (2019): ‘Analysis of Business Intelligence Applications in Healthcare Organizations’, *Proceedings of the 52nd Hawaii International Conference on System Sciences*.
- Islind, Anna Sigridur, Lindroth, Tomas, Lundin, Johanm, and Steineck, Gunnar (2019): ‘Shift in translations: Data work with patient-generated health data in clinical practice’, *Health informatics journal*, vol. 25, no. 3, pp.577-586.
- Kitchin, Rob (2014): *The data revolution: Big data, open data, data infrastructures and their consequences*, Sage.

- Madsen, Laura. B. (2014): *Data-Driven healthcare: How analytics and BI are transforming the industry*, John Wiley & Sons.
- Mayer-Schönberger, Victor and Cukier, Kenneth (2013): *Big Data: A Revolution that Will Transform how We Live, Work, and Think*, Houghton Mifflin Harcourt.
- Mayer-Schönberger, Viktor and Ingelsson, Erik (2018): Big Data and medicine: a big deal?, in: *Wiley Online Library*.
- Mettler, Tobias and Vimarlund, Vivian (2009): 'Understanding Business Intelligence in the Context of Healthcare', *Health informatics journal*, vol. 15, no. 3, pp.254-264.
- Pine, Kathleen H. (2019): 'The qualitative dimension of healthcare data interoperability', *Health informatics journal*, vol. 25, no. 3, pp. 536-548.
- Pine, Kathleen H. and Bossen, Claus (2020): 'Good organizational reasons for better medical records: The data work of clinical documentation integrity specialists', *Big Data & Society*, vol. 7, no. 2, doi.org/10.1177/2053951720965616.
- Raghupathi, Wullianallur, & Raghupathi, Viju (2014): 'Big data analytics in healthcare: promise and potential', *Health information science and systems*, vol. 2, no. 1, pp.1-10.
- Ruckenstein, Minna and Schüll, Natasha Dow (2017): 'The Datafication of Health', *Annual Review of Anthropology*, vol. 46, no. 1, pp. 261-278. doi:10.1146/annurev-anthro-102116-041244
- Simborg, Donald W, Detmer, Don Eugene, and Berner, Eta S. (2013): 'The wave has finally broken: now what?', *Journal of the American Medical Informatics Association*, vol. 20, no. e1, pp. e21-e25
- Star, Susan Leigh and Strauss, Anselm (1999): 'Layers of Silence, Arenas of Voice: The ecology of visible and invisible work', *Computer supported cooperative work (CSCW)*, vol. 8, no. 1, pp9-30.
- Van Es, Karin and Schäfer, Mirko Tobias (2017): *The datafied society. Studying culture through data*, Amsterdam University Press.
- Wixom, Barb and Watson, Hugh (2010): "The BI-Based Organisation," *International Journal of Business Intelligence Research*, vol. 1, no. 1, pp.13-28