

# Information Infrastructures in healthcare and emergent data work occupations: The case of medical scribes and CDIS

*Bossen, Claus<sup>1</sup>, Chen, Yunan<sup>2</sup>, and Pine, Kathleen H<sup>3</sup>*

<sup>1</sup>Aarhus University, Denmark; <sup>2</sup>University of California, Irvine, USA;

<sup>3</sup>Arizona State University, USA

[clausbossen@cc.au.dk](mailto:clausbossen@cc.au.dk); [yunanc@ics.uci.edu](mailto:yunanc@ics.uci.edu); [kghammon@asu.edu](mailto:kghammon@asu.edu)

**Abstract.** The development of information infrastructures in healthcare is often described in abstract terms as datafication, the conversion of qualitative aspects of life into quantified data, which makes the people and actual work involved invisible. To make visible the actors and the efforts implied in the term 'datafication', in this paper, we describe two emergent data work occupations in healthcare: Medical scribes and clinical documentation improvement specialists (CDIS). These cases provide a starting point for understanding the impacts of digitization of healthcare and the emergence of new kinds of work and new occupations that health organizations adapted to accommodate such impacts. Making data work visible is important in order for these occupations to be acknowledged. If data work remains invisible, healthcare organizations and researchers alike will have an incomplete understanding of how data is actually produced in practice, hindering the organizational design, human resources, and organizational learning that are essential for healthcare organizations to become competent producers and users of data.

## Introduction

Digitization of healthcare has been ongoing in Europe and Northern America over the last two decades, involving electronic health records (EHR) for hospitals, electronic

care records for home care, and most recently patient generated health data via mobile and monitoring devices. Together with demands for increased accountability (Wiener 2000, Power 1997), and ambitions for healthcare to become more data-driven (Madsen 2014, Raghupathi et al. 2014, Mayer-Schönberger et al. 2018), data-centric technological development and implementation has become the norm in healthcare.

Concomitant with the widespread implementation of digital information infrastructures for healthcare and the expanded ability to collect, store, manage, analyze, and deploy data is the phenomena of “datafication.” Datafication is defined as the conversion of qualitative aspects of life into quantified data (Ruckenstein & Dow Schüll, 2017). Discussions have emerged of implications of the datafication of healthcare in a wide range of areas. On the one hand, increased computing power for data collection and management coupled with increasingly complex data analytic tools promise earlier discovery and prevention of diseases; more precise diagnoses; discovery of new correlations between symptoms, treatments, and cures; and more effective management of healthcare (Raghupathi et al. 2014, Mayer-Schönberger et al. 2018). On the other hand, questions have been raised about the assumptions and critical implications of this development. Thus, one overall trend is the increased integration of clinical research and commercial domains. As Hogle (2016) describes: “The data curated from ‘nonmedical’ and conventional medical sources can be combined and repurposed from original contexts and uses, and made available to serve a variety of healthcare, marketing, and governance needs.” (ibid, p387).

Critical scholars and concerned citizens alike point to privacy as an ongoing concern with data-intensive health research, particularly research that utilizes data that have previously not been used for health research, such as social media data. An oft cited example of the dangers of data intensive business intelligence tools for privacy is that of the USA-based Target Corporation identifying the pregnancy of a young woman via her purchasing behavior before her father knew, and disclosing her secret to him via targeted advertising (New York Times, Feb. 16, 2012 “How Companies Learn Your Secrets”). Another concern is the new kind of social sorting of populations into groups based on algorithmically created categories, which might reinforce old beliefs about social differences (Hogle, 2016, p388). Overall, datafication implies that healthcare organisations reorganize around data production and analysis in a process that can be labelled as ‘data-intensive resourcing,’ defined as: “... attempts at getting more data, of better quality, on more people. Sourcing is a dynamic process of creating, collecting, curating and storing data while simultaneously making them available for multiple purposes, including research, governance and economic growth” (Hoeyer, 2016, p74).

The current discussions around datafication have been centered on the use and integration of various types of digital health data and the impacts of digitizing such data on both clinicians and patients, e.g. the research that focuses on the implications of EHR implementation for clinical practice and patient experience. Another line of scholarly work places a critical spotlight on datafication, attending to the political and ethical implications of measurements and algorithms and the potential downsides of valorizing data-intensive methods over other forms of seeking knowledge for

organizational and institutional reflection and decision making. However, both lines of work fail to provide deep empirical accounts about the work of actually producing data. Scholarly attention to datafication thus far, in both healthcare and in other domains, tends to treat datafication as an abstract process without active subjects when describing the overall process. We find this reflected in phrases like ‘data is the new oil (or gold)’ and ‘datafication will (insert some effect) ...’, and while we acknowledge the need to summarize and describe overall processes, people as actors seem to fall into the background.

But, data does not come from air--nor is it simply a byproduct of other processes (e.g. clinical documentation in EHRs). In this paper, we take a broader approach to include active subjects in the process of digitizing data and focus on the people, occupations and actual work involved in producing data that populates data repositories and is fed into algorithms and complex data analytics. This is a way of making otherwise invisible work that underlies the phenomenon of datafication and the implications of the development information infrastructures in healthcare visible. To understand the dynamics of how the new ecology of healthcare information infrastructures transforms work, we focus on two occupations: medical scribes and clinical documentation improvement specialists (CDIS). Making this data work visible is important in order for these occupations to be acknowledged, and because neglecting it can have detrimental consequences for the achievement of the right skill-mix in healthcare organizations and thus their use of digital information infrastructures.

## Surfacing data work in healthcare

Visions for Big Data herald high-skilled professions build of formal knowledge such as statisticians or data scientists (Mayer-Schönberger and Cukier 2013, Madsen 2014), while far less attention is paid to the low-paid work behind, for example, Amazon Mechanical Turk (Irani 2015), or the moderators removing content in conflict with nudity policies of e.g. Facebook (Roberts 2016). Entrepreneurs and innovators like Bezos and Zuckerberg are highly visible, while the work behind these fronts remains in the background, almost invisible. Data work, however, encompasses a far wider range of present and future ‘data professionals’ (Foster 2016), and indeed a small number of emerging studies does exist that focus on the work of producing and making data meaningful (Foster 2016, Fischer et al. 2017, Kristiansen et al. 2018).

However, data work is often invisible. When infrastructures such as electricity or data-generating e-health systems are in place, they tend to disappear in the sense of being taken for granted and unnoticed (Bowker & Star, 1999). Invisibility of work derives from three main components: The division of social spaces into front- and backstage (e.g. sales counters vs storerooms) (Goffman, 2002); the social construction of what counts as ‘real work’ or not (e.g. paid work vs. domestic work); and how it is represented or not (e.g. statistics abstracting away sweat and exertion) (Star & Strauss, 1999). Invisibility is often closely connected to low status and remuneration, but can

also crucially influence the implementation of ICT systems and make them fail, when designers and other stakeholders fail to attend to and account for crucial work perceived as superfluous or routine in their design (Bowers, Button, & Sharrock, 1995; Suchman, 1995). Hence, ‘making work visible’ is important both for the data workers themselves as well as for the organisation in which they are employed.

## The new data work occupations

In the following, we will focus on two emergent data work occupations: medical scribes and clinical documentation improvements specialists (CDIS). The purpose is to get a better understanding of what kinds of new work and occupations emerge in the wake of datafication of healthcare.

The descriptions are based on data from qualitative methods, and a literature review. In the case of medical scribes, we conducted 14 interviews with medical scribes and their manager amounting to approximately 13 hours of interviews (shortest 39 minutes; longest 68 minutes; mean 53 minutes), and 3 rounds of observations amounting to 9 hours of medical scribes working in an Emergency Room at a hospital in the Western region of the USA. In addition, we conducted a literature review on previous studies of medical scribes (Bossen, Chen & Pine 2019). For the case of CDIS we conducted 12 interviews with CDIS for approximately 4 hours of interviews in total (shortest 14 minutes; longest 29 minutes; average 24 minutes), and 9 rounds of observation of CDIS working on patient records (total of 27 hours). The present paper is based on preliminary analysis of the transcribed interviews and extended field notes.

### Medical scribes

Medical scribes write down information from conversations between patients and doctors, and enter this information into EHRs on behalf of doctors in order to lessen doctors’ workload. One could call them a kind of doctor’s secretary.

Medical scribes have grown in number especially in the wake of digitization of healthcare in connection with the implementation of EHRs in the USA since the 2010, and are especially popular in Emergency Departments where the pressure for speed and seeing many patients is high. Medical scribes have no formal teaching or education and are typically trained for some weeks before starting to scribe for physicians. They are usually paid the minimum wage and generally leave the occupation after a year. Several companies have emerged to provide scribe services to healthcare organizations. These companies recruit, train, and contract with hospitals to provide scribes at a fee. In 2012, medical scribes were acknowledged by the non-profit USA based healthcare accreditation organisation Joint Commission as a job or an occupation within healthcare. In summary, one could say that medical scribes are low-skilled, low-paid, have a large turn-over and that the provision of scribes services has become occasions

for businesses to grow (For an overview, see (Bossen, Chen, & Pine, 2019). We can also say that digitization of healthcare, the precondition of data-intensive research and data-driven decision making in healthcare, has led to the emergence of a new kind of data work. Medical scribes are contributing to the production of the data through providing the raw material for big data, because digitization in the form of EHRs entail more documentation and a higher workload for doctors.

A closer look at medical scribes provides a more nuanced narrative. Being a good medical scribe actually requires several skills. First, they need to be able to extract the relevant information from conversations between doctors and patients in the hectic, noisy, and sometimes intense context of emergency rooms, and enter this information into the proper fields in the EHR. To be able to do this, they learn medical terms and the structure of doctors' medical interviews with patients, including the following items: Reason for the visit; History of Present Illness (HPI); Past Medical History (PMH); Review of Systems (ROS); Social History (including recreational drug use); Family History (FH), Allergies, etc. They also learn the structure in which information is formatted in the EHR, and the idiosyncratic styles in which each particular doctor wants their notes to be written. Further, scribes must also be able to be the doctor's unobtrusive and silent shadow, who is unavoidably visible to the patient, but does not obstruct the conversation, and they must be able to move around with their computers on wheels in a tight, crowded space. Often, they look up medical terms online at idle moments, or when they come home, in order to learn more. Despite the demanding work and low pay most scribes find the work experience rewarding, and becoming a scribe is relatively competitive. One might wonder why?

Our interviews revealed that most medical scribes have a bachelor's degree – often in natural sciences – and use their medical scribe job to get experience within healthcare, and thus improve their chances of being admitted to medical school (See e.g. (Lowry, 2017)). Therefore, another narrative that emerges is that medical scribes are highly competent persons gaining relevant experiences as a step in their educational career.

## Clinical Documentation Improvement Specialists

Another emergent data occupation is that of the Clinical Documentation Improvement Specialists (CDIS). Like the medical scribes, their occupation has emerged and grown along with the implementation of EHRs. Their job is to code medical records with the labels that form the basis for reimbursement from insurance companies for medical treatment and care. A job that requires solid and broad knowledge of medical terms.

Unlike the medical scribes that have recently finished their undergraduate degrees,, CDIS are typically nurses with 10 or more years of experience who for varying reasons want to make a change in their career. Also, while scribes work at the minimum pay, CDIS earn on average US\$ 68,000 a year, approximately the same as the salary of registered nurses working in clinical practice, but significantly more than the

approximately. US\$ 29,000 a year for medical scribes. Indeed, more than 25% of all CDIS earn more than 100,000 a year (See: Association of Clinical Documentation Improvement Specialists (ACDIS) (2019 “2018. CDI Salary Survey). Thus, one CDIS we met found nursing work physically too hard, while another had worked with cancer patients for many years and found this to be too emotionally demanding. Also, unlike the medical scribes who work surrounded by patients, nurses and doctors, CDIS work in quiet offices away from patients on desktop computers through which they have access to two systems: One is the hospital’s EHR in which they look at particular patient’s medical notes, examinations, test results, etc, while another is a system based on Natural language processing combined with the classification codes embedded in Diagnosis-Related Groups (DRG) used as the basis for reimbursement of healthcare services in the USA, and many other countries, including Denmark (On computer aided coding, see for example systems such as 3M’s ‘3M 360 Encompass’ system; for DRG See: (Busse, Geissler, Quentin, & Wiley, 2011)).

The work for CDIS consists in making sure that clinical documentation by doctors is precise, correct and comprehensive. Seen from their perspectives, doctors are – for good reasons, CDIS acknowledge – more focused on keeping up with treating incoming patients than on producing complete and accurate records for post hoc purposes (e.g. billing and calculating quality measurements). However, “bad” records (from the perspective of these post hoc usages) mean that hospitals provide services for which they are not paid, resulting in economic deficits and hence potentially lower quality of treatment and care. For that reason, it is important for hospital administration that patient records are accurate and comprehensive, and it is the job of CDIS to read records and code them accurately in order for the hospital to be reimbursed. This job includes occasionally writing to doctors asking them to be more specific. When the patient was admitted, did he or she have ‘mild sepsis’ or ‘severe sepsis’? What kind of pneumonia did the patient have? These inquiries called ‘Clinical clarifications’ are conducted via special forms in which CDIS may ask for precision or reconsideration of a diagnosis. Particularly in these instances is it important CDIS to have and display medical knowledge, since doctors would otherwise disregard such clarifications. Notably, clinical clarifications cannot point in any specific direction in order to ensure that CDIS do not pressure or lead doctors to choose diagnoses or terms that result in higher reimbursement.

Unlike medical scribes, CDIS often remain in their positions long term and consider being a CDIS a career choice, where scribe work is seen as short term, temporary employment on the road to a medical career. The data CDIS produce are those of DRG-codes and numbers indicating the severity of the diagnosis recorded as CCs (Complications and Comorbidities) and MCCs (Major Complications and Comorbidities). Their narratives are checked by another data occupation, Coders, who are experts on the legal specification and coding systems, and go through the claims report before it is sent to the insurance company. As in the case of medical scribes, CDIS work to make medical records more precise and accurate, but they do this in specific ways, since, for example only the medical records of patients admitted to the

hospital and staying more than 2 days are reviewed by CDIs. The rationale for this being that return of investment of CDIS time is low, and records for outpatient patients or those staying for less than two days are simple, relatively low on reimbursement, and hence the opportunity for making high-dollar value improvements to clinical documentation is small.

## Discussion

Medical scribes and CDIS are two examples of occupations that have emerged in connection with datafication of healthcare. Both are relatively new, which goes some way to explain why relatively little research has been conducted with and about them. We found 60 papers published over 43 years - most after the turn of the century - on medical scribes and most are concerned with return of investment on hiring these workers (Bossen, Chen and Pine 2019). Research papers on CDIS are even more scarce. However, we also suspect that their relative invisibility is related to the mundane kind of data work that these occupational groups do. They do not head successful IT companies or digital platforms, nor do they have the lure of new high-skilled occupations such as 'data scientists'. However, acknowledging their contributions to the achievement of datafied healthcare is vital for understanding what strategies data-driven healthcare require, and how this might change the existing skill-mix of occupations.

The dynamics of digitization and the changing tasks, emergence of new tasks, and ongoing negotiation of the scope of work of different professions and occupational groups is part of the dynamic interrelations between professions. Professions are seen as integral to the division of work in modern bureaucracies such as healthcare (Abbott, 1988; Bourgeault, Dent, Denis, & Kuhlmann, 2016; Freidson, 2001), and the boundaries between them change as part of power struggles, technological developments and emergence of new government policies (Dent, Bourgeault, Denis, & Kuhlmann, 2016). Within healthcare, new occupations emerge as routine tasks are delegated to new occupations (e.g. physician assistants), or when new technology is implemented (e.g. radiology technicians). The drivers of change to the overall skill-mix in healthcare can be attributed to technological innovation along with new expectations towards healthcare services, and changes in inter-professional and profession-state relations (Cooper, 2001; de Bont et al., 2016; Tsiachristas et al., 2015).

The dynamics of changes in professions' roles and work tasks can be described as entailing four processes: 1) *diversification* (adding tasks to the existing portfolio. E.g. physicians taking ownership of the technology of anesthetics); 2) *specialization* (acquiring increased level of expertise. E.g. anesthetic nurses), 3) *vertical substitution* (task adoption across hierarchical boundaries. E.g. nurses doing prescriptions); and 4) *horizontal substitution* (task adoption across same-level professions. E.g. ward secretaries taking over nurses' tasks) (Nancarrow & Borthwick, 2005). A central factor

in the change dynamics is the ability of high-ranking professions to delegate routine or unpleasant tasks to professions with lower rank.

Medical scribes and CDIS are two examples of how these dynamics can unfold. Both occupations work to produce advanced, structured, and precise narratives of healthcare data. The former at the point of entry to the hospital, the latter whenever admittance and more complex treatment and care is required. Both occupations have grown as an outcome of digitization of healthcare and the push for more documentation and data-intensive forms of management, accountability, and regulation of healthcare and. The emergence of the occupation of medical scribes can be characterised as the result of a process of vertical substitution, where doctors delegate the routine (but still skilled) task of documentation to people with (relatively) less formal education (a BA degree is an achievement). For medical scribes, working for a short time at a minimum pay is balanced out by the acquirement of knowledge and experience within the healthcare domain, and thus the improvement of the chance of getting admitted to medical school. It should be noted, however, that this balance is only temporary, since medical scribes usually stay for a year maximum after which they move on either to go into medical school, or pursue another career. As for the CDIS, the emergence of this occupation can be seen as a similar process of delegating the routine tasks of coding documentation that doctors otherwise would have to do, to persons with less formal education (and with lower pay). However, at the same time, this process can also from the perspective of nurses be characterised as a case of diversification, where the tasks of improving and coding clinical documentation becomes an addition to the portfolio of their occupation (though nurses do not have a monopoly on this task). For nurses (and related professions), working as a CDIS is physically and emotionally less stressful, and potentially offers the gain of a higher salary.

Based on the examples from these two occupations, datafication of healthcare entails both the delegation of routine tasks of creating data to less skilled people, as well as specialisation of nursing work with similar or higher salary prospects. In a wider perspective, the use of medical scribes is a way to lessen the increased documentation and work load that datafication of healthcare has entailed for doctors. CDIS are a result of that same process of increased documentation and data work for doctors, and both increase and decrease the pressure of more data work for doctors: They do the required coding work that doctors otherwise would have to do, but at the same time through for example 'clinical clarification' inquiries increase data work by requiring more specific and detailed information than doctors sometimes deem necessary.

Getting a full picture of the implications for data work in healthcare is challenging and there we have focused on only two emergent occupations. We suspect that several others can be found. However, one overall point is that datafication is a process that does not unfold by itself, but requires new kinds of tasks and work, and which we need to put into the equation when considering the overall gains and costs of the new era of datafied healthcare.

## Acknowledgments

We would like to extend our sincere thanks and appreciation to the people who helped us get in contact with medical scribes and CDIS at the two sites, as well as the managers, medical scribes and CDIS that participated in interviews. For the sake of anonymity, we do not mention you by name, but you have been most helpful and forthcoming for which we are grateful.

## References

- Abbott, A. (1988). *The system of professions*. Chicago and London: The University of Chicago Press.
- Bossen, C., Chen, Y., & Pine, K. H. (2019). The emergence of new data work occupations in healthcare: The case of medical scribes. *International Journal of Medical Informatics*, 123, 76-83. doi:<https://doi.org/10.1016/j.ijmedinf.2019.01.001>
- Bourgeault, I. L., Dent, M., Denis, J.-L., & Kuhlmann, E. (2016). General introduction: The changing world of professions and professionalism. In *The Routledge Companion to the Professions and Professionalism* (pp. 19-28): Routledge.
- Bowers, J., Button, G., & Sharrock, W. (1995). *Workflow from Within and Without: Technology and Cooperative Work on the Print Industry Shopfloor*. Paper presented at the Proceedings of the Fourth European Conference on Computer-Supported Cooperative Work, Dordrecht.
- Bowker, G. C., & Star, S. L. (1999). *Sorting Things Out. Classification and Its Consequences*. Cambridge (Mass) & London: MIT Press.
- Busse, R., Geissler, A., Quentin, W., & Wiley, M. (Eds.). (2011). *Diagnosis-related groups in Europe*. New York,: Open University Press.
- Cooper, R. A. (2001). Health Care Workforce for the Twenty-First Century: The Impact of Nonphysician Clinicians. *Annual Review of Medicine*, 52(1), 51-61. doi:10.1146/annurev.med.52.1.51
- de Bont, A., van Exel, J., Coretti, S., Ökem, Z. G., Janssen, M., Hope, K. L., . . . Wallenburg, I. (2016). Reconfiguring health workforce: a case-based comparative study explaining the increasingly diverse professional roles in Europe. *BMC Health Services Research*, 16(1), 637. doi:10.1186/s12913-016-1898-0
- Dent, M., Bourgeault, I. L., Denis, J.-L., & Kuhlmann, E. (2016). *The Routledge companion to the professions and professionalism*: Routledge.
- Freidson, E. (2001). *Professionalism, the Third Logic*. Chicago: Chicago University Press.
- Goffman, E. (2002). The presentation of self in everyday life. 1959. *Garden City, NY*.
- Hoeyer, K. (2016). Denmark at a Crossroad? Intensified Data Sourcing in a Research Radical Country. In B. D. Mittelstadt & L. Floridi (Eds.), *The Ethics of Biomedical Big Data* (pp. 73-93). Cham: Springer International Publishing.

- Hogle, L. F. (2016). Data-intensive resourcing in healthcare. *BioSocieties*, 11(3), 372-393. doi:10.1057/s41292-016-0004-5
- Lowry, J. E. (2017). Medical Students' Perspectives on Their Experiences as Medical Scribes. *MA Thesis. Michigan State University.*
- Nancarrow, S. A., & Borthwick, A. M. (2005). Dynamic professional boundaries in the healthcare workforce. *Sociology of Health and Illness*, 27. doi:10.1111/j.1467-9566.2005.00463.x
- Star, S. L., & Strauss, A. (1999). Layers of Silence, Arenas of Voice: the Ecology of Visible and Invisible Work. *Computer Supported Cooperative Work*, 8, 9-30.
- Suchman, L. (1995). Making Work Visible. *Communications of the ACM*, 38(9), 56-63.
- Tsiachristas, A., Wallenburg, I., Bond, C. M., Elliot, R. F., Busse, R., van Exel, J., Rutten-van Mólken M.P., de Bont A., and MUNROS team.(2015). Costs and effects of new professional roles: Evidence from a literature review. *Health Policy*, 119(9), 1176-1187. doi:10.1016/j.healthpol.2015.04.001