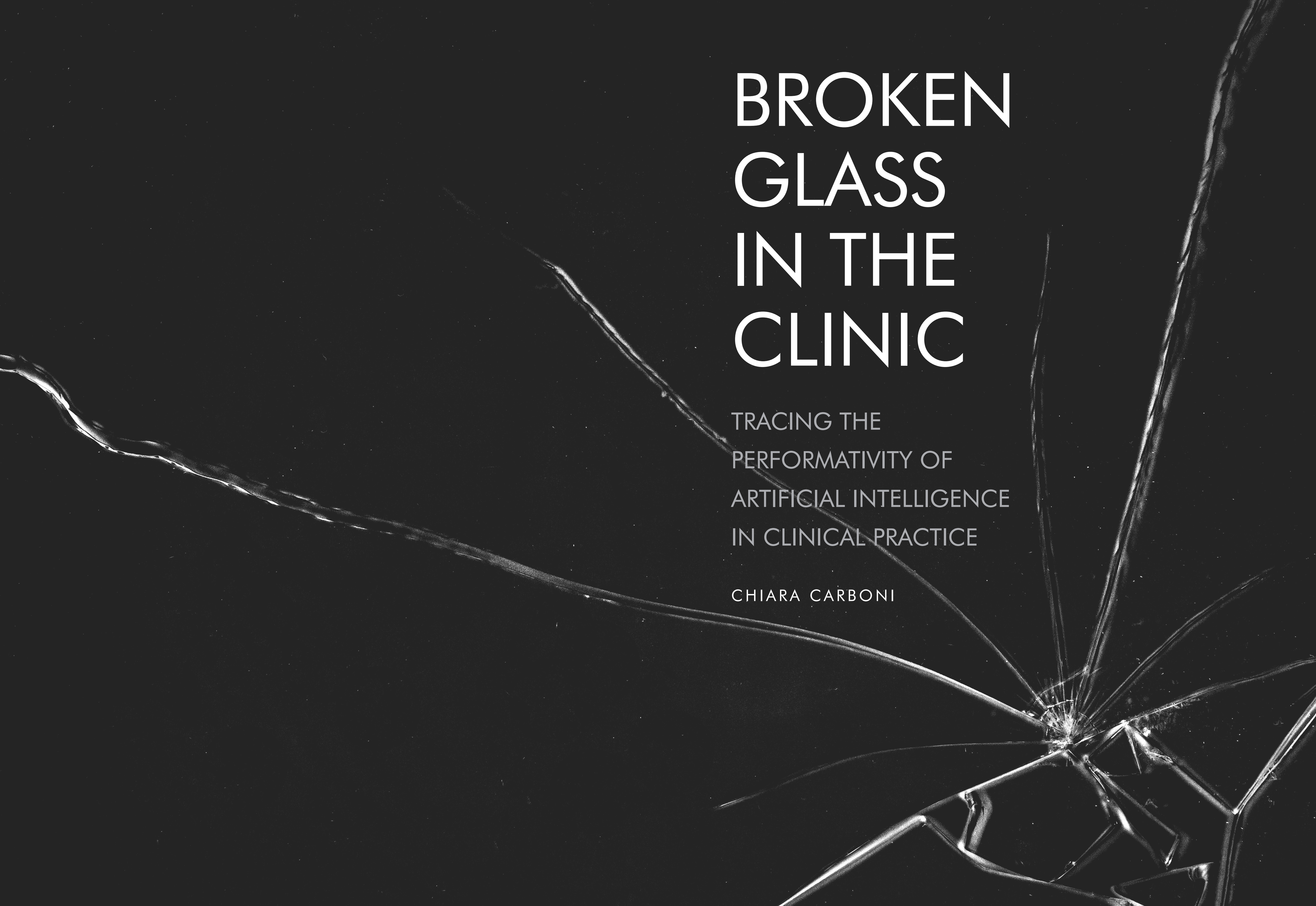


BROKEN GLASS IN THE CLINIC

TRACING THE PERFORMATIVITY OF ARTIFICIAL INTELLIGENCE IN CLINICAL PRACTICE

CHIARA CARBONI

Broken glass in the clinic
Tracing the performativity of artificial intelligence in clinical practice

Chiara Carboni

Produced by Proefschrifterij

www.proefschrifterij.nl

Broken glass in the clinic
Tracing the performativity of artificial intelligence in clinical practice

Gebroken glas in de kliniek
Schetsen van de performativiteit van kunstmatige intelligentie in de klinische praktijk

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr. A.L. Bredenoord

and in accordance with the decision of the Doctorate Board.

The public defence shall be held on

Friday 29 November 2024 at 10:30 hrs

by

Chiara Carboni
born in Sassari, Italy

Erasmus University Rotterdam

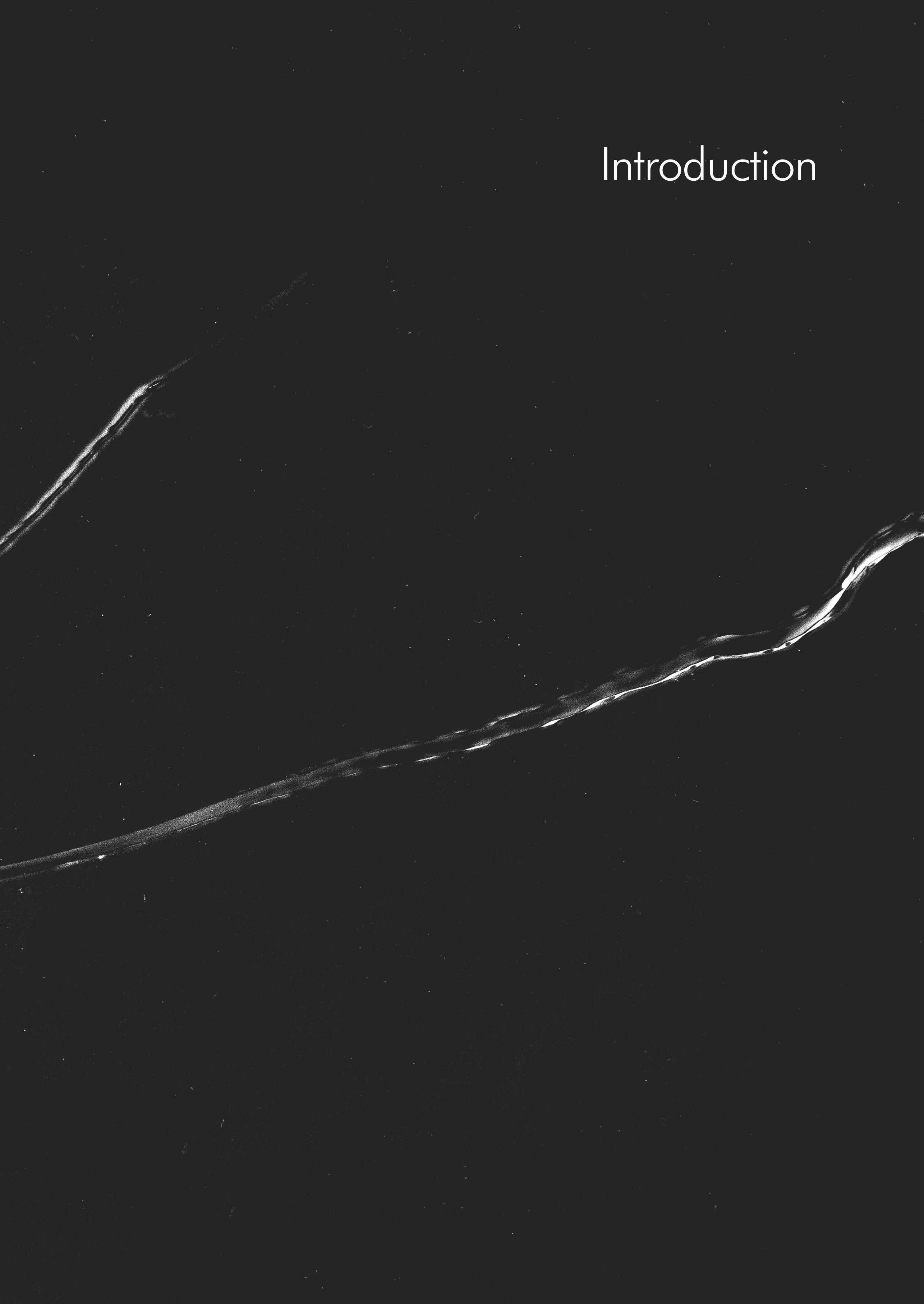
Erasmus

Doctoral Committee:

Promotors:	Prof.dr. A.A. de Bont
	Prof.dr. R.J. van der Veen
Other members:	Prof.dr. R.A. Bal
	Prof.dr. B.R. Winterheik
	Prof.dr. C.L. ter Hoeven
Co-promotor:	Dr. R.L.E. Wehrens

Contents

Introduction



Breaking glass for AI

In a 2019 lecture entitled *The language of broken glass*, German filmmaker Hito Steyerl introduces some of her recent work, *The city of broken windows*. The video portrays a team of engineers at work in a World War II airplane hangar — a strange location, in which they engage in a strange activity. One by one, we see them taking turns in smashing windowpanes with large hammers, while an array of microphones and wires record the noise they produce (figure 1). As Steyerl's interviews them, we find out that they work for a company that develops sound-recognition technologies for policing. Specifically, these engineers are trying to develop an AI capable of recognising the sound of a window being smashed — the purpose of this being to alert the police immediately when someone tries to break into a house.

Figure 1: Still from Hito Steyerl's 'The city of broken windows.' Image courtesy of the Artist, Andrew Kreps Gallery, New York and Esther Schipper, Berlin.

Engineers alternate between smashing the glass, carefully sweeping the floors, and replacing broken windows with new ones to be broken, while the bang of the hammers gives way to crystalline explosions of shattering glass that progressively bleed into a glockenspiel-like melody. One of the engineers, a man in his 50s, begins to explain how "breaking the glass … can help you *model sounds better and do better artificial intelligence*" (emphasis added). As he goes on to elaborate,

> There's a French zoologist called Buffon, and he established sort of an inventory of animals — some animals have two feet, some animals have four. And we're kind of doing the same with sounds: there's really a huge variety of sounds ... So, breaking a window is a percussive bang followed by some kind of smashes and so on. (from Steyerl 2019; emphasis added)

Breaking windows over and over to attempt to reproduce this "huge variety of sounds" is precisely what is needed to "model sounds" and "do better artificial intelligence." Since AI models need exposure, the engineers' task becomes to provide such exposure: "what we're doing here is really, actual window breaking — we're taking a hammer and breaking a window. *So this is reality*" (emphasis added).

The first way in which Steyerl's work speaks to the small book you are (be)holding right now is in its foregrounding the tension between 'exposure' and "reality" that is inherent to Artificial Intelligence (AI). As we will see emerging across the chapters that follow, in order to learn and function, AI needs to be provided with a very specific version of that reality. Investigating what reality *comes to be* in the wake of AI's introduction, and what this means, specifically, for clinical practice, is one of the tractive forces of this book.

In this sense, this study departs from some of the currently common approaches to the study of clinical AI in the social sciences. So far, a great deal of the literature has focused either on discursive aspects of AI, foregrounding promissory narratives and expectations around these technologies (e.g. Bareis and Katzenbach 2022; Vicsek 2021), or on the organisational dimension, thus looking chiefly at questions of trust, use, and implementation (Glaser, Pollock and D'Adderio 2021; Lebovitz, Lifshitz-Assaf and Levina 2022). In contrast, this book proposes that we need to find ways to merge these two analytical approaches. We need theoretical and methodological frameworks, as well as conceptual vocabularies, to consider organisational *and* discursive aspects of clinical AI in their inextricability. More concretely, we need to account for how the ways in which AI is *speakable* at the current moment (i.e. what it is expected to achieve, how it is expected to work, what it is expected to need) are remaking organisations and modes of care provision.

Especially in this latter sense, my approach here matches Steyerl's. Like her, I gesture towards the often seemingly paradoxical activities professionals and other workers (in healthcare and beyond) are being asked to perform in order to (potentially) produce better technologies and usher them into practice (Gray and Suri 2019). In a way, this aligns with extant efforts, in Science and Technology Studies (STS) and medical sociology, in tracing the (often unintended) effects of medical technologies in use (e.g. Bailey et al. 2020; Timmermans and Berg 2003). While strongly inspired by this literature's attention for the subtleties of technological change in clinical settings, this study tries to make up for what might be seen as this literature's agnosticism to the political implications of the changes it so painstakingly describes.[1] On the one hand, indeed, its strong emphasis on the empirically observable has the undoubted merit of moving away from both technological determinism and social essentialism. This is particularly valuable for scholarship that, as Høyer has shown recently, needs to find ways to tell stories that foreground the fundamentally paradoxical nature of datafication, which has "created both less work and more work; … empower[s] and disempower[s] staff and patients; … both uncover[s] patient concerns and cover[s] up patient concerns; … both tightens organisational control and generates new forms of organisational disintegration" (2023: 12). However, while not denying the value of empirically pinning down the many paradoxes of digitalisation, this study aims to move beyond this descriptivist approach. This might materialise, as I argue, in opening up how such paradoxes may be distributed along fault lines of power imbalances in and beyond organisations. Moreover, it might entail interrogating the paradoxical changes we observe in practice by tracing them back to discursive forces that, in their invisibility, still manage to achieve something in the world.

To do so, I begin here to consider what it might mean to think of contemporary clinical AI as implicated in what, borrowing from Steyerl, I term *glass-breaking*. Glass-breaking, in this introduction, and throughout this book, will refer to *the reconfigurations, observable in contemporary clinical practice, that aim to make clinical practice itself a more suitable ground for training and applying artificial intelligence*. Studying AI as glass-breaking, as I will argue in this introduction, entails being cognizant of the power inherent to this more or

[1] The political agnosticism stemming from STS' foundational descriptivist stances is of course a well-known issue, discussed for instance by Radder (1998).

less accomplished remaking of clinical practice. What I am proposing, in a nutshell, is that the current reconfigurations that clinical settings are undergoing are geared towards making them better settings for AI's needs and affordances, and that we need to attend to the forces driving these changes.

On the one hand, and in the current moment, clinical practice becomes pre-eminently, from a machine learning perspective, a domain of *model training*: highly digitised care provision and concurrent registration practices are namely, at this point, a source of data for model training. That is, clinical practice is being reconfigured in order to enable its appraisal as a deluge of datapoints (Rajpurkar et al. 2022). On the other hand, the whole point of this is to turn clinical practice, in a more or less distant future (and, in a few cases, in the present), into a domain of *model application*, where AI can be implemented to assist with or take over a variety of tasks (e.g. Elemento et al. 2021). This entails another kind of reconfiguration, that would enable, for instance, the capture and analysis of relevant data in real time (Ismail et al. 2020). Clinical practice is thus being remade as both a source of data for models and as an apt space for releasing (or, currently, mostly testing) them. Thinking about this remaking through the lens of glass-breaking emphasises the labour of glass-breaking itself, as well as the political nature of the organisational reconfigurations we are witnessing. Foregrounding glass-breaking, we might begin to ask what kind of work and tasks are made necessary in achieving a version of clinical practice that is suited for AI's needs. In what follows, I offer a quick review of some recent social science literature concerned with current instances of "glass-breaking," particularly with reference to clinical AI, and which the present study builds upon. I start with reviewing literature on digitalisation, datafication, and AI, moving on to studies of data production in clinical contexts, and of the implications of datafication and AI for clinical knowledge-making. I then articulate the performative theoretical framework through which, in this book, I attempt to make sense of and to repoliticise glass-breaking as a matter of work and shifting priorities — that is, political decisions that are often framed as stringent necessities. I condense these reflections in a few questions guiding the study. Finally, I lay out some methodological considerations in the context of a more concrete expositions of the cases constituting this book.

Studying glass-breaking: Data work, professional reconfiguration, and knowledge-making

Things: Digitalisation and datafication

In this first section, attending to the technological aspects of glass-breaking enables me to dwell on how this study is positioned vis à vis some of the terms I have somewhat unproblematically been mobilising so far (AI, machine learning, datafication, digit(al)isation), as well as on the relationships among these different phenomena as this study envisions them.

In the broadest sense, this study focuses on digitalisation, which I mobilise here as a broad umbrella term encompassing all processes related to the introduction of some type of digital technology (cf. Trittin-Ulbrich et al. 2020; for a definition of digital healthcare technology, cf. chapter one). It thus joins a quite substantial, if relatively recent, social science literature centring on digitalisation in the healthcare sector. This literature, as Ruckenstein and Schüll (2017) note, tends to focus on the Global North, thus discussing digitalisation in relation to the various demographic, financial, and political challenges characteristics of these regions — namely, "an aging population, rising rates of chronic disease, unsustainable health care costs … and the retreat of social welfare" (262).

As a feature more specific of the current historical moment, I zoom in on processes of datafication and digitisation — two terms that, in this study, I consider almost interchangeable. If datafication is generally meant to encompass processes of data production, digitisation has a stronger connotation of turning something previously analogue into a digital form. However, as Høyer (2023) notes, currently, "digitalisation [sic] turns all types of records into data" (18), meaning that the data of datafication tends to be digital. This pre-empts attempts to empirically disentangle the two phenomena. Like many scholars contributing to the emerging field of critical data studies (CDS), in this study I work with a broad notion of data, spanning "numbers, characters, symbols, images, sounds, electromagnetic waves, bits— that constitute the building blocks from which information and knowledge are created" (Kitchin 2014: 1, emphasis added; cf. also Borgman 2015). In an attempt to operationalise such a broad definition, Høyer (2023) has recently suggested we should think of data as "something traceable that can be subjected to computation, decontextualised, and transferred to other users," as well as reused (17-18, emphasis added; cf. also Thylstrup 2019). The relationship between data and knowledge-making, as well as digitalisation's imposition of compulsory computability, are at the heart of this study.

In the last few years, CDS scholarship has drastically sharpened our understanding of the ontological, epistemological, and ethico-political challenges associated with data and processes of datafication. Mainstream, hyperbolic conceptualisations of (big) data as bringing about epistemic revolutions and making theory obsolete (Anderson 2008; Mayer- Schönberger and Cukier 2013) have been criticised on account of data always being made rather than found (Leonelli 2016); always being the result of an interpretation, and thus not "raw," but rather "cooked" (Biruk 2018; Gitelman and Jackson 2013). Data have been conceptualised as presenting a specific view of the world, thus reproducing and importing into datasets epistemologies that tend to be white, male, ableist, universalist, (D'Ignazio and Klein 2018; Denton et al. 2021). Finally, as much as they are decontextualised, data are always nonetheless, and irremediably, shaped by their context of production (Loukissas 2019; Thylstrup et al. 2022).

In clinical contexts, a considerable variety of data types are often part and parcel of care provision: "digital imaging in pixels, lab tests in numbers, chart notes, or treatment plans" (Hogle 2016: 374). So far, most of the literature on the datafication of healthcare (Ruckenstein and Schüll 2017) has looked at either individuals' mobilisation of health data in research adjacent to the Quantified Self movement (Lupton 2016; Pink et al. 2017; Sharon 2017), or, more recently, at the ingression of big tech in the healthcare realm (Sharon 2016; 2021; Heimstädt, Egbert and Esposito 2020; Ozalp et al. 2022). As a consequence, both theorisations and empirical analyses of datafication in clinical settings and its implications for mundane practices care provision have been limited so far. An exception has been the work of Linda Hogle (2016), whose work resonates with the concept of glass-breaking I am articulating in this introduction. Hogle has fruitfully emphasised the 'circularity of rationalities' underpinning current narratives around datafication in healthcare, according to which the belief that "capturing big data will enable the transformation of healthcare' serves to usher in the conviction that it is 'necessary to transform healthcare to capture big data" (380-381). In other words, these narratives manage to present the labour involved in transforming clinical practice into settings suited for data collection as an unquestionable necessity, all the while concealing this very labour by focusing instead on the (yet to be realised) transformative potential of datafication itself.

Processes of datafication and their materialities are what is, in most cases, empirically observable in contemporary clinical practice. In fact, when I refer to glass-breaking, I mostly think about how instances of datafication materially reconfigure clinical practice to make it a space more suited for data extraction and, to a lesser extent, model application. Like the scenes opening this book, these instances can be odd, and only begin to make sense in connection with AI as an end goal, positioned in a yet-undefined future. Thinking of datafication as glass-breaking helps us centre on and make sense of the changes in the practice and organisation of care provision that are aimed at providing more and better clinical data, while keeping an eye on AI's reliance on these processes and thus role in fostering them. If datafication is what we can most often observe empirically, it is increasingly hard to disentangle this phenomenon, and its politics, from data analytics and, consequently, AI (Ruckenstein and Schüll 2017). This study, inspired by the emic perspectives I encountered in my fieldwork, thus offers a reading of datafication as chiefly a necessary predecessor of, and indeed often a somewhat uneasy steppingstone towards, the yet immaterial object of clinical AI.[2] In other words, I propose that datafication is increasingly justified not only in terms of goals such as care provision and continuity, but also, and crucially, in view of the AI models that databases will enable training, and which real-time digital data will enable running.

Things? Algorithms, artificial intelligence and machine learning

Of course, I am not alone in linking datafication to aspirations and developments in the field of AI. In his 2014 seminal book, geographer Rob Kitchin states that "the goal of much [Big Data] research is to develop automated processes that can assess and learn from the data and their analysis." On a similar line, Pasquinelli (2019) has emphasised how the

[2] This despite acknowledging that the multiple ontology of clinical data means that datafication can be mobilised for a variety of aims, as convincingly shown by Høyer (2023).

generation of training datasets, that is, compiling data traces captured in the wild,[3] ought to be considered a foundational component of machine learning systems.

Machine learning is a branch of AI that "seeks to iteratively evolve an understanding of a dataset; to automatically learn to recognise complex patterns and construct models that explain and predict such patterns and optimise outcomes" (Kitchin 2014). On the technical side, machine learning materialises in algorithms, or models, that are developed by tweaking their weightings in relation to a training dataset in which they aim to identify "clusters and relationships between the data." Machine learning can be "supervised," when "a model is trained to match inputs to certain known outputs," or "unsupervised," when "the model seeks to teach itself to spot patterns and find structure in data without the use of training data" (ibid.). After being tested, and if their output "shows an adequate alignment or 'fit' with the training data (as assessed by human operators)" (Suchman 2023), models are released in the wild, that is, "applied to automate the classification of patterns or predict the probability of the recurrence of a pattern in future data" (ibid.).

Alongside machine learning, in this book I repeatedly refer to (clinical) AI. Albeit occupying centre stage in current research, policymaking, and practice in and beyond healthcare, AI is a concept that seems to consistently escape definition (e.g. AI HLEG 2019). Standard definitions, casting AI as a research 'field that aims to create computational systems capable of demonstrating human-like intelligence' have been criticised as excessively circular (Suchman 2023). Although I am not interested in normatively defining AI, in this book I subscribe to the point, recently articulated by Lucy Suchman (2023), that these difficulties in definitional efforts are not without a reason. As Suchman argues, the vagueness surrounding the definition of AI as a "thing" is indeed strategic and serves the interest of the proponents of impending AI revolutions by creating the illusion of a concrete referent. To counter the "over-representation of AI's existence as an autonomous entity and unequivocal fact," Suchman calls for critical AI (Raley and Rhee 2023) scholars to trouble the "thingness of AI," and to develop instead "a keener focus on [the] locations, politics, material-semiotic specificity, and effects" of "the proliferating data and

[3] Although increasingly produced synthetically to circumvent privacy concerns.

compute-intensive techniques and technologies that travel under the sign of AI" (2023).

In this study, I interrogate "AI" as an emic term (cf. also Seaver 2017), in an attempt to take seriously participants' definitions of what should be subsumed under this category. However, by considering AI always in connection with glass-breaking (i.e. the empirically-observable disruptions and reconfigurations it both requires and engenders), I aim to simultaneously trouble this notion in its "enchantment" (Campolo and Crawford 2020; Suchman 2007).[4] Indeed, this book's endeavour to pick apart the glass-breaking unfolding in contemporary clinical practice can be read as a way of engaging with the concrete processes that underpin, and are perhaps concealed by, clinical AI's often uncritically accepted "thingness." A considerable part of such processes comes down not only to data and their materialities, but also to the people involved in data production and mobilisation in the clinic.

People: (Care and data) work

Moving on to the people involved in glass-breaking, in this section I string together literature that focuses on two increasingly entangled issues: the implications of datafication for care provision, and the work of data production in clinical setting. In discussing datafied care provision, I am supported by flourishing strands of STS and sociological literature. In fact, particularly in its inception, my work has drawn heavily upon studies at the intersection of STS and (medical) sociology, focusing on technologically-induced professional reconfiguration. This literature builds on classical sociological interests in the reshaping of professional roles and jurisdiction (Abott 1988), as well as investigation of professional values and the organisation of work in healthcare

[4] Here, it is worth making a note of a political shortcoming of my way of mobilising the term AI in this book. Embracing emic uses of "AI" in the end could conceal the deliberate work, and the power, of Big Tech actors that, by owning the infrastructure necessary to develop, run and maintain these technologies, de facto own "AI" and get to dictate the terms of glass-breaking to a very considerable extent. Unfortunately, an analysis of these dynamics goes beyond the scope of this study, which is strongly anchored to the everyday realities of clinical practice. Indeed, although they do sometimes emerge in the late stages of development of the technologies I examine (cf. chapter three), these actors are usually too far to substantially influence the dynamics I focus on here. For a further discussion of this limitation, see the methodological implications discussed in chapter six.

(Strauss 1975).

As chapter one articulates in more depth, when it comes to studying the digitalisation of clinical settings, recent STS and sociological literature overlap considerably in their subscribing to the 'technology-in-practice' approach (Timmermans and Berg 2003). This approach brings anti-deterministic STS sensitivities to bear on classic sociological themes by espousing STS insights into the politics of technology design (Henriksen and Blond 2023) and its practical achievements. In the case of digitalisation in clinical settings, it reminds us, for instance, how 'digital tools are typically designed not only to document what people do, but also to shape what they do and how they do it' (Høyer 2023: 95). Especially in the case of the Electronic Health Record (EHR), this has often been found to translate into reduced spaces for professional judgement, with deleterious (affective) consequences for healthcare professionals, such as "feelings of meaninglessness, whether understood as alienation …, moral disorientation … or powerlessness" (Høyer 2023: 96).

If technology-in-practice approaches to the EHR abound, few empirical studies of clinical AI in practice have been published so far (Jaton and Sormani 2023). This somewhat sparse literature tends to foreground healthcare professionals' selective reliance on algorithmic recommendations and integration into pre-existing diagnostic practices (Maiers 2017), as well as their ability to manipulate algorithms' outputs by selecting what information is fed to algorithms themselves (Russell 2012). These studies also tend to share a concern with the reductionism inherent to algorithms, especially vis à vis the skillset of the professionals in whose practices they are supposed to intervene. Such reductionism, these scholars worry, risks translating into the marginalisation of forms of embodied and idiosyncratic knowledge that various professionals mobilise in their care practices (Henriksen and Bechmann 2020; Maiers 2017; Russell 2012; Schwennesen 2019).

Another line of discussion spans questions of a more organisational nature, asking chiefly whether algorithms can be mobilised to speed up or take over labour currently performed by professionals. Some studies find that new roles attending specifically to algorithmic outputs and coordinating them with care processes are necessary to harness algorithmic processes in healthcare organisations (e.g. Bailey et al. 2020). Others, usually focusing on radiology, emphasise how algorithms for automated image analysis have enabled either strengthening professionals' diagnoses, or have taken over the task of double-checking other colleagues' diagnosis (cf. Høyer 2023 for a review; cf. also Avnoon and Oliver 2023). If a focus on glass-breaking does entail surfacing empirically observable changes to work practices and professional roles, this book does not engage directly with the question whether algorithms introduce more work or save work for professionals. Instead, as is often the case in STS analysis, it attempts to rephrase the terms of the question, by unearthing different locales at which work is done not only by AI, but also, and crucially, *for* AI (cf. Suchman 2006).

Insofar as AI relies on data, we can think of the work done for 'it' through the conceptual lens of data work. The relatively recent strand of scholarship on data work merges an interest in the practical, ongoing efforts in enabling datafication with an attention to the consequences of these efforts for what (clinical) data comes to be (Bossen et al. 2019). This literature imports an attention to empirical detail and orientation from extant literature on the sociology of professions and technology-in-practice. However, it accentuates the political dimensions inherent to these traditions by drawing on feminist STS's endeavour to make the invisible visible in pursuit, mainly, of recognition (Star and Strauss 1999; cf. also Foster et al. 2018; Rothschild et al. 2022). Moreover, it builds on CDS insights into not only the "cooked" nature of data, but also on the situated knowledge involved in generating *good* data (Leonelli 2012).

The genealogy of the concept of data work dates back at least to Berg and Goorman (1999), who developed the so-called law of medical informatics in the context of their analysis of the EHR. This 'law' drew attention not only to how, the farther away it needs to travel, the more information needs to be simplified and abstracted from local condition of production — but also to the work this process of abstraction entails, and to whom needs to perform it (cf. also Loukissas 2019; Høyer 2023). Indeed, even before them, Suchman (2007 [1987]) showed how much of the "enchantment" surrounding supposedly autonomous machines relies on the expunction of the human labour that is essential to their very functioning. This emphasis on human labour proves a particularly valuable insight in a time when clinical AI is routinely invested with revolutionary potential as the ultimate technofix for the multifaceted crises affecting healthcare systems worldwide.

This study's focus on glass-breaking in contemporary clinical practice centres on data work as the human investment required to sustain glass-breaking itself. I adopt here a broad view of data work, and this is for two reasons. First, I consider 'data' in the diversity of forms it can take in clinical practice, ranging, as articulated above, from numbers, to words, to images. Because of this, as we shall see in chapter two, I try to unpack what might be gained by considering as data work the work involved, for instance, in producing 'good' digital images. Second, it might be worth broadening the scope of glass-breaking to consider not only *new* instances of data work emerging in connection with clinical AI. Indeed, insofar as glass-breaking relates to the aims of making clinical settings more suited for AI, the *repurposing* of already existing data for new, AI-related aims, should arguably also be included as a crucial part of glass-breaking. We might thus wonder how, in this context, pre-existing tasks, such as data registration in the EHR, are being reframed, and perhaps reconfigured, to also address these emerging aims. Simply put, who does the work of making data suitable for both care provision and model training, and what does this work entail? Attending to these perhaps more subtle reconfigurations helps us gain a more complete picture of glass-breaking in contemporary clinical practice.

People and things: Datafied knowledge practices

As mentioned earlier, in this study I consider data not only in their material dimension and reliance on organisational processes, but also, and crucially, in their connection with information and knowledge-making. This means that the processes of datafication that I group under the banner of glass-breaking are considered in relation to, and in their implications for, knowledge-making in clinical practice. Under the rubric of knowledge-making, I consider here a number of processes related to the generation, registration, mobilisation and interpretation of various types of data, which are routinely carried out in clinical practice — processes commonly referred to with notions such as diagnosis, monitoring and treatment.

Professional knowledge has traditionally been considered, especially in economic and innovation literature, as somewhat antithetical to forms of digitalisation. Especially when pondering questions around automation-driven unemployment, this literature tends to emphasise how aspects of professional knowledge that resist codification and thus automation hinder the possibility of machine-driven substitution of professionals (e.g. Autor 2015). These aspects, traditionally considered uniquely “human,” have to do, for instance, with professionals’ adaptive expertise and ability to respond to emerging situations in environments that are by nature unruly and unpredictable (Holford 2022).

Other scholarship pitting professional knowledge practices against data emphasises, for instance, the reductionism of forms of knowing rooted in data by noting how datafication necessarily flattens the complexity of lived experience (Islam 2022), and how some aspects of the world plainly resist datafication (Kitchin 2014). In the medical context, Høyer underlines that “there are types of knowledge needed in healthcare that do not derive from data” (2023: 152). Indeed, a considerable part of knowing in clinical practice is embodied, and it does not stem from, nor is it sufficiently captured through, data analysis (cf. also Goodwin 2010; Gardner and Williams 2015; Moreira 2019; Friis 2021). In this sense, research on sensory work in healthcare has foregrounded the embodied skills that professionals need to acquire and master in order to provide care (e.g., Bijsterveld 2018; Harris 2021). In an optic fundamentally opposing professional knowledge practices to data, the increased datafication of clinical practice is often interpreted in terms of a progressive loss of clinical competence (Nettleton, Burrows and Watt 2008; Hunt et al. 2017).

In this book, I follow Høyer (2023) in his observation that, although an exclusive reliance on data would amount not only to epistemic loss, but to a safety hazard, the concept of healthcare professionals’ knowledge practices needs to be broadened to span their expertise in making sense of data. Indeed, technologies, data, sensors need be, and are increasingly being, incorporated in sensing practices (Maslen and Harris 2021) — to an extent that makes it unproductive to consider them separately from, or even as opposed to, professional knowledge-making practices. Literature at the intersection of anthropology, STS and human-computer interaction (HCI) has for a long time been making the case for a posthumanist account to knowledge practices. Habitually, we conceive of cognition and knowledge as somehow internal to humans, and usually located in their brains. In contrast, this strand of scholarship calls for studies of knowledge-making to move beyond the idea of a liberal subject fundamentally other to and untethered from its (or, disproportionately, his) surrounding nonhuman environment. These scholars have convincingly pointed out how cognitive skills (Hutchins 1995) and even scientific knowledge (Knorr Cetina 1999) are accomplishments not only necessarily embodied, but realised at the *intersection between human bodies and their and sociomaterial surroundings*. Researching how knowledge-making practices are reshaped in contexts of glass-breaking, then, requires broadening our unit of analysis to include not only human professionals, but also machines, artifacts, and the relations amongst them.

In the following section, I delve deeper into the theoretical foundations of these studies of knowledge-making as distributed beyond the boundaries of human bodies. Building on this, I work my way towards the overarching theoretical framework for this book, which, as I argue, enables me to put these insights into fruitful conversation with CDS' treatment of datafication.

Towards a performative account of knowing, being, and taking responsibility

In her seminal examination of human-computer interfaces, Lucy Suchman (2007) famously picks apart situated interactions between people and machines in an attempt to trouble claims of autonomous machine agency. If I am not concerned here with questions of agency per se, Suchman's work is foundational for this study in its aim to provide "an empirical investigation of the concrete practices through which categories of human and nonhuman are mobilised and become salient within particular fields of action" (1). That is, Suchman begins to shift the analysis of the knowledge and action achieved at the human-machine interface towards an ontological plane. According to her, when studying agency (or, for our purposes, knowledge-making) at the human-machine interface, we are better off suspending our received categories of what is a human and what is a machine. This makes the posthumanist approach to knowledge introduced above even more radical: it is not simply the case that processes of knowledge-making should be considered to extend beyond human or machinic entities. Indeed, these very entities we call 'humans' and 'machines' should be considered to emerge as part of situated practices.

The posthumanist framework of agential realism, articulated by feminist science studies scholar Karen Barad (2007) supports us in making sense of this shift towards ontology. While I offer a more thorough examination of agential realism in chapter two, in this section it seems worthwhile to articulate some of its foundational principles, as well as the way in which they relate to other theoretical developments in fields such as STS, sociology, and CDS. Indeed, although not always explicitly mobilised, sensitivities stemming from the agential realist framework underpin my analysis of objects and subjects, as well as epistemic and ethical questions, emerging in the wake of glass-breaking all throughout this book.

Not unlike Suchman, Barad starts from acknowledging that "[t]he "knower" cannot be assumed to be a self-contained rational human subject, nor even its prosthetically-enhanced variant" (379). Their posthumanism manifests in a troubling of the essence of subjects and objects of knowledge practices, claiming that "[t]he subjects … constituted [in knowledge practices] may range across some of the presumed boundaries (such as those between human and nonhuman and self and other) that get taken for granted" (379). Like the studies discussed above, this entails a radical decentring of subjects that are commonly regarded as "humans" in favour of the "larger material configuration of the world" (ibid.) as part of which subjects emerge.

If, up to this point, Barad's approach simply entails broadening our analytical units beyond human subjects, it is in its approach to knowledge that agential realism starts to showcase its innovativeness. Barad proposes to consider knowledge as "a direct material engagement" (378). This goes beyond the point about knowledge practices being embodied and sensorial: it is a way of moving away from representationalism, the philosophical position according to which knowledge constitutes a representation of reality. Conversely, Barad proposes that knowledge practices are inextricable from the semiotic and material reality in which they are situated, and are crucially implicated in the very emergence of that reality. This is a *performative*[5] position

[5] It is worth making a note, here, of how theory is rooted in queer theory (Butler 1990; 1993; Kosofsky Sedgwic 1993). In this study, I engage with the side of performativity theory that is concerned with material performativity, rather than its consequences for subjectivation. The former can be traced back, at least, to Butler's Bodies that matter, where they propose "a return to the notion of matter, not as site or surface, but as a process of materialisation that stabilises over time to produce the effect of boundary, fixity, and surface we call matter . . . Crucially,

(or in Suchman's term, a materially-constructive one): objects are not to be considered as entities autonomously existing in the world, waiting for a knower to represent them by mobilising some technical medium. Rather, subjects, objects and instruments are all *enacted* (i.e., come into being) within situated epistemic practices. In Barad's terms, "relata do not precede their relating" (2007: 334). This interrelating is, in their words, "ontologically primitive" (2003: 815).

Barad discusses this relating in terms of "intra-action" — a term that moves away from the construct of *inter*action, which they reject as presupposing two entities that precede their interrelating. In agential realism there are no primitive subjects or objects: they emerge and constitute one another in the course of intra-actions. As chapter two articulates, agential realism considers technoscientific practices as crucially implicated in the emergence of new subjects and objects. This study thinks as instances of intra-action the numerous and distributed moments in clinical settings in which some kind of knowledge is created and mobilised: providing care to a body, writing up data in the EHR, formulating a diagnosis. This entails that who or what is considered a knowing subject in a specific locale, or who or what does the caring in a specific instance, should not be considered a given, but instead approached empirically.

Intra-actions congeal phenomena in flux into entities with defined boundaries, properties and meanings. Barad borrows from physicist Niels Bohr the concept of apparatus, which he used to describe how laboratory settings and machinery constrained, and participated in the enactment of, the phenomena observed during experiments. In Barad's sense, apparatuses are close to Foucauldian discursive practices (cf. e.g. Foucault 1980) insofar as they *participate in*, and indeed constrain, the 'material production of bodies and meanings' (Barad 2003). Given the entanglement of matter and meaning at the heart of Barad's metaphysics, apparatuses are simultaneously material and discursive, and they "produce, rather than merely describe, the subjects and objects in knowledge practices" (2007: 147). A specific apparatus is culturally and historically contingent, and it constrains reality enacted through it insofar as it supports specific "agential cuts." That is, it presupposes specific

---

then, [the construction of bodies] is neither a single act nor a causal process initiated by a subject and culminating in a set of fixed effects" (1993, pp. 9-10).

boundaries around and between subjects and objects (i.e. in our case, specific notions of what is a knowing subject, and specific boundaries for humans and machines). Building on agential realism, then, we might begin to consider what kind of apparatuses emerge in the wake of glass-breaking, and which cuts they enable. In other words, which processes, subjects and objects are made possible in contemporary clinical settings?

Before moving on to the central questions I address in this study, it is worth dwelling a little longer on the continuities and discontinuities this theoretical approach presents with analytical programs in STS. Indeed, agential realism is a new-materialist framework, and neo-materialist approaches are widespread in STS. For instance, Actor-Network Theory (ANT), a theoretical cornerstone of this interdisciplinary field, has long foregrounded ontological issues, by positing the ontological performativity of the network it centres on. Indeed, ANT does not presuppose connections among pre-existing, stable entities; rather, agents' characteristics and their very agency emerges from their reciprocal relations (e.g. Latour 2005). On a similar note, Woolgar and Lezaun (2013), in discussing the so-called "ontological turn" in STS, point out STS's foundational interest in the performativity of language, emerging in the deflationary investigations of the "situated use and accomplishment of grandiose theoretical concepts," as well as its foregrounding of "the instrumental, performative and material dimensions implied in the making of facts and artefacts" (322).

I agree with Woolgar and Lezaun (2013) when they point out that various turns to ontology in STS and feminist science studies have the merit of repoliticising STS analyses, which have often been accused of excessive descriptivism (e.g. Radder 1998). For them, moving analyses from the epistemic to the ontological can "draw renewed critical attention to objects that might otherwise appear 'finished' or 'ready-made', to scrutinise those entities that a conventional STS analysis would often consider 'blackboxed' and no longer controversial" (323). Although here I do not pursue the interest in multiple ontologies (e.g. famously, Mol 2002) that has driven much of the ontological turn in STS, by turning to Barad's articulation of onto-epistemology I do share its political commitment to showing that, also on the ontological plane, "things could be otherwise."

I mobilise Barad's agential realism not only because of its emphasis on the ultimate contingency of realities and entities that are enacted in practice, but also because of its normative acknowledgement of the ethico-political import of being implicated in this very enactment. Indeed, as many scholars have noted, this analytical position turns practices, including epistemic practices, into cosmopolitical matters (Latour 2004; Papadopoulos 2011; Stengers 2011). In other words, showing the contingency of any ontology implies the necessity of asking questions about how we participate in the enactment of present worlds, and about the nature of the worlds we want to enact. For Barad, if our knowledge practices are inseparable from ontological questions, our analyses should take the form of an *ethico*-onto-epistemology, in which we do not only consider knowledge practices and their enacted realities, but also the ways in which both we and the people whose practices we study can take *responsibility* for those very realities. In the next sections, I attempt to condense this reflection into questions scaffolding this study.

Research question

So far, I have proposed that data-related reconfigurations empirically observable in contemporary clinical practice amount to a process of glass-breaking, that is, an attempt to make clinical practice itself a more suitable ground for training and, eventually, applying clinical AI. Indeed, I have argued that such glass-breaking is an essential way in which, aside from ever-abounding promissory narratives, AI manifests in clinical practice in the present moment. Mobilising the vocabulary laid out in the previous section, we could think of the relationship between clinical AI and instances of glass-breaking as one between apparatus and intra-actions. That is, clinical AI should not be thought of as determining the reconfigurations I group here under the rubric of glass-breaking. However, as I propose, it needs to be considered as part of the conglomeration of materialities and meanings that shapes dynamics of glass-breaking, thus participating in constraining possible configurations for the worlds glass-breaking creates and for the meanings associated with it.

By trying to delineate contemporary clinical AI as it manifests in practice, this study aims to engage critically with both discursive and material dimensions of clinical AI. AI is, as we have seen, as slippery object, which can be addressed only athwart, by piecing together its scattered manifestations.

Attempting to say something about it entails remaining open to its performative potential, letting "the field" tell us more about what clinical AI, in the present moment, *is*. Throughout this study, I am thus guided by an overarching question:

How does AI manifest in contemporary clinical practice, and what implications does this bear for the organisation and practice of care provision?

To address the first part of the question, in the impossibility to look at AI 'directly' from clinical settings, I resort to attending to its glass-breaking. In other words, I attempt to provide empirically-anchored accounts of contemporary processes of glass-breaking in clinical practice, spanning data and the work it takes to produce it, as well as the knowledge and care practices it is part of and supports. This entails addressing the following sub-questions:

1. How are data produced in clinical settings?
2. What objects for organisational and professional intervention are created or reconfigured by clinical AI (and its attendant glass-breaking), and with what implications for professional knowledge practices?

These first sub-questions address glass-breaking in its core components: data, the work of data production, and the knowledge-making practices in which it participates. All the chapters in this book speak to these sub-questions. The former, eminently empirical, gestures towards a tracing of the subtleties of data work, and its implications for data quality. Chapters two and three, specifically, foreground the material and sensorial dimensions of data work, focusing on two forms of datafication of bodies (present or absent), and foregrounding the complex intra-actions amongst what comes to be defined as human and nonhuman bodies, and as the result of which data crystallises. Chapter four turns to the datafication not of bodies, but of behavioural cues — a question that emerges as crucial in datafied psychiatry.

Given the link between data and knowledge-making, the second sub-question also runs across the whole study. Indeed, processes of glass-breaking, and the increasing mobilisation of various types of data in practices of care provision, are likely to reconfigure what kind of information becomes available to professionals, and where it is to be found. This is bound to be crucial to the performance of knowledge-making tasks, such as diagnosis (chapter two), monitoring and treatment (chapter three and four). Drawing on agential realism, I consider reconfigurations in knowledge-making practices as entangled with newly emerging epistemic subjects and objects. In chapter two, I inflect this insight to address the question of how subjects that are either human or nonhuman are supported by specific material-semiotic apparatuses of which AI is also part. Chapters three and four focus instead on various kinds of objects enacted through the apparatus of specific clinical AIs, either for resource allocation or for safety reasons.

3. What does this tell us about the mechanisms underlying how professionals embed AI in their work?

This final sub-question revisits a central concern of STS literature with how users (in our case, healthcare professionals) appropriate, or domesticate, newly introduced technologies in their daily practices (Lie and Sørensen 1996). Bringing users into the picture helps me avoid deterministic stances toward clinical AI, while simultaneously addressing a major emic and practical concern, widespread amongst those involved in innovation projects: the possibility of resistance and non-use. Needless to say, all the reconfigurations discussed up to this moment would inevitably fail to realise were healthcare professionals to simply refuse to engage with clinical AI and to perform the glass-breaking necessary to bring it about. Across my chapters, I try to take seriously instances perceived as resistance to innovation (chapter two) or temporary non-use (chapter four) precisely by considering them in relation to the reconfigurations in knowledge-making, as well as the subjects and objects emerging in the wake of glass-breaking.

Clinical AI-not-really-in-practice: Chancing upon cases, and other methodological considerations

This study builds on three ethnographic case studies, which were conducted between 2021 and 2022. In this context, discussing case sampling matter-of-factly, as a question of free selection of the cases most fitting in an aprioristically rationally-devised matrix of variables, would be hardly believable. This book aims to uphold the spirit of feminist scholarship emphasising not only the serendipity of any work of research, especially empirical, but also the importance of shedding teleological conventions of academic writing that strive to hide difficulties and failures rather than probing their generativity. Thus, I will not pretend that my case selection was informed by a strict set of inclusion and exclusion criteria. Indeed, there was hardly any selection at all. Although, obviously, I was interested in observing some kind of digital healthcare technology in practice (or about to be introduced), the main inclusion criterion for my case studies was the possibility of obtaining access to a healthcare organisation during (or in the wake of) a pandemic.

As a second preamble, it is worth mentioning briefly that the "field" itself worked, from the very beginning, to unsettle the definition of "technology-in-practice" that I set out with when negotiating access. Indeed, in some cases, gate keepers or other points of contact informed me, prior to the start of my fieldwork, that a specific technology (usually related to clinical AI) was *about to be introduced* in their department. Invariably, hearing this, I felt as if I had struck gold: being able to observe the early days of a technology supposed to be majorly disruptive, as professionals worked to make sense of and generally domesticate it, was certain to translate into incredibly rich fieldnotes detailing the complex tensions emerging in clinical practice. Invariably, I would get my ethics board's clearance and data management plans in order as soon as I could. Invariably, once fieldwork had started, I would find out that the technology they had discussed was not really *there* — nor quite *about to* be, at least measured against the duration of my PhD. Rather, some professionals (and usually my fieldwork's gatekeepers) were, at that point, discussing possibilities for starting a tendering process to purchase such a technology, or else discussing how to develop one themselves, and trying to attract partners to support them in this process.

Once I recovered from the disappointment of not being able to observe clinical AI "in practice" after all, I started to recognise the ever-impending stretchiness of this "*about to*" as a constitutive feature of the field of medical technology innovation. This led me to question what it takes for a technology such as AI to *be* in clinical practice, and what this technology still manages to achieve even when not *tangibly there*. Having to deal with a version of reality that actively worked against my inclusion criteria — *chancing upon* my cases, as it were, rather than selecting them — ended up affording what I believe are some of the most generative insights this study provides.

Nonetheless, it is possible to describe, *a posteriori*, the dimensions along which my chanced-upon cases variated, and what they had in common, as summarised in table 1 below.

All three cases were essentially concerned with centring various forms of glass-breaking enacted by clinical AI in a highly digitalised healthcare system, the Dutch one. Since cases are described at length in their respective chapters, for the moment being it will suffice to say that they are situated in largely different settings: namely, a pathology department (chapter two), two adult ICUs (chapter three), and two acute psychiatric clinics (chapter four). Interestingly, this diversity means that, whereas chapter two is centred on laboratory medicine, which is arguably amongst the least acute types of care provided in clinical settings, chapters three and four both offer examples of acute care, where knowledge- and decision-making are bound to be particularly charged, especially in terms of patient (and, at least in chapter four, provider) safety.

Diversity in settings translates into diversity in the type of data that is central to each case. In chapter two, I examine digital *images* as an emerging type of data for pathologists to (not) accommodate in their daily work. Chapter three focuses on the mostly *numerical data*, sometimes translated into waves, that are produced and manipulated on an ongoing basis in ICU care practices. Finally, chapter four turns to *text-based data* and its quantification, looking at EHR reports and risk scores based on them. Thus, these cases manage to cover most of the extreme variety of forms in which data manifest in clinical settings.

Stemming from the variety in data types, cases also engage with three distinct types of clinical AI technologies: chapter two discusses (a yet-to come) AI for *automated image analysis*; chapter three presents a case of what could be described *algorithmic management* in clinical settings, that is, mobilising clinical data to allocate resources efficiently; finally, in chapter four I look at a *predictive algorithm*. It should be noted here that it is not only the case that the "clinical AIs" I consider in the following chapters have different aims. Rather, they differ also, and crucially, in their degree of distance from clinical practice — at least, at the time of my observations. In this sense, the chapters follow clinical AI as it gets *progressively closer* to the clinic: in chapter two it is but a possibility situated in the future; in chapter three it is in its final stages of development, potentially ready to be introduced to the clinic. The technology in chapter four is, technically speaking, the only one that could be defined as "in practice," since it is being tested in a pilot.

	Pathology (chapter two)	ICU (chapter three)	Acute psychiatry (chapter four)
Care settings	Non-acute; diagnostics	Acute	Acute
Type of data	Images	Numbers	Text
Type of AI	Image analysis	Patient monitoring and resource management	Prediction
AI's distance from the clinic	No concrete technology (+)	Final (?) stages of development (+-)	Pilot (-)
Who performs bulk of data work?	Secretaries; lab technicians	Nurses	Nurses

Table 1: Relevant dimensions of chanced-upon cases

Across cases, I have had the opportunity to interface with (and to observe how clinical AI materialises for) different professionals: pathologists, lab technicians and secretaries (chapter two), ICU nurses and intensivists (chapter three), and psychiatric nurses and psychiatrists (chapter four). However, given my focus on glass-breaking, I decided to privilege the standpoint of (and thus, methodologically speaking, to focus my observations on) the people performing data work in each case. Thus, I chiefly discuss nurses in chapters three and four; since there are no nurses in a pathology department, in chapter two I look at lab technicians and, crucially, secretaries.

Bringing the agential realist framework discussed above to bear upon my methodological choices entails taking as a starting point its articulation of ontology, epistemology and ethics as undivorceable from one another. That means foregrounding knowledge-making practices as ontologically-generative sites, sites at which realities are enacted and, simultaneously, sites at which knowing subjects are required to take responsibility for the realities in the enactment of which they participate. I follow here Mol (2002) in her felicitous attempt to come to grips with ontologies in medical practice through engagement with the subtleties of work as is done. Mol's praxiography entails suspending received notions of what the objects of medical work are, and rather observing as they "come into being — and disappear — with the practices in which they are manipulated" (2002: 5). In this sense, an ethnographic focus on concrete practices, that are often, at once, work, knowledge, data and care practices, is warranted in light of my general aim to foreground glass-breaking and its implications for various aspects of clinical practice.

Across my cases, my ethnographic approach has been one of zooming in on moments of knowledge-making, and indeed retracing them in surprising locales that exceed the more visible forms of professional work. Through ongoing ethnographic engagement with professionals and non-professionals responsible for data work in each setting, I have attempted to excavate the epistemic practices hidden in and crucial for data production. I have followed their consequences down the workflow line, where datafied knowledges are mobilised either by other professionals or by various technologies. Finally, my approach has also been characterised by an ongoing attention for the specificity of the technologies I (more or less directly) engaged in my case studies. Indeed, understanding as well as possible the workings of the technologies the (non-)professionals I observed were engaging with was fundamental for me to be able to say something about the politics of their design, and the realities enacted through them. This entailed carefully observing as these technologies were operated in practice (in the case of chapter two's scanners being taught to operate them myself), or, in case of technologies-not-in-practice, picking apart the accounts my research participants would give of them.[6]

Moving closer to clinical AI: Overview of the chapters

This book opens with a review that assembles a vocabulary and an empirically-anchored sensitivity to examine the reconfigurations that clinical AI expects of (and sometimes engenders in) clinical practice. Given the dearth of empirical studies of clinical AI in practice (cf. Jaton and Sormani 2023), *chapter one* broadens the scope of examined technologies to think through what we can learn from past instances of digitalisation of healthcare work. This way, I am able to situate the glass-breaking I describe in the following chapters within a longer history of other technologically-driven instances of professional and organisational reconfiguration. Through a metaphor-based version of a Critical Interpretive Synthesis (CIS), my co-authors and I conceptualise the digitalisation of healthcare work as a phenomenon spanning, at once, the open-endedness of situated changes in work practices which I have referred to as glass-breaking, and the directionality of technological innovation trajectories, a dimension that goes beyond what is empirically observable. We propose focusing on technological scripts, and various forms of invisible work as entry points into the study of the digitalisation of healthcare work.

In the following chapters, my co-authors and I take these insights with us as we conduct ethnographies of datafied clinical practices. *Chapter two* takes us to a pathology department where diagnostic processes are being changed in the name of clinical AI technologies that are yet to come. This is a case of digitisation, where the glass slides pathologists would normally use as part of their diagnostic practices are being replaced with digital images. We lay out here Barad's agential realist framework to examine the epistemic disruptions AI manages to generate, even in its absence. We enrich agential realism with the concept of fauxtomation, to sharpen its political import and applicability to organisational questions. By foregrounding changes in work processes for pathologists, lab technicians, and secretaries, we find here a prime example of glass-breaking: the introduction of new technologies and epistemic objects that seem to bring little benefit to workers, and yet are considered necessary as a steppingstone towards the promises of clinical AI.

[6] I articulate these reflections more at length in the last chapter of this book.

Chapter three moves one step closer to clinical AI, and into the realm of acute care. This chapter centres on the study of an innovation project aimed at making ICU nurses' work more "data-driven" through the implementation of an AI-powered dashboard. Although, strictly speaking, no glass-breaking was observable (yet) in these settings, we pick apart plans for an AI-driven dashboard, and try to foreground its (potential) implications for ICU nurses' practices of embodied knowing. In so doing, we harness the insights provided by agential realism insofar as we consider AI's performativity in terms of the objects it enacts. We argue that AI technologies increasingly proposed to tackle personnel shortages in healthcare constitute nurses as sources of attention, and their attention itself into a scarce resource to be made efficient through data. Attempting to move away from these efficiency-driven views, we assemble a theory of attunement as something crucial in ensuring the reliability of, and thus the possibility of acting upon, real-time data in the ICU.

Chapter four takes us to two acute psychiatry clinics, where we encounter the first (and, in this book, only) example of clinical AI in practice. In this chapter, we follow the pilot of an algorithm aimed at pre-empting inpatient violence, thus flagging patients as dangerous before they would show explicit signs of aggressivity. Even though risk scores were produced and circulated daily for each patient in the two clinics, local staff did not mobilise them as legitimate sources of knowledge in their decision-making. Rather, they considered any prediction that deviated from their own judgement as simply wrong. This chapter offers an empirically-grounded reflection on this case of non-use by dwelling on the practices and ethics of dealing with violence as articulated by local nurses. Juxtaposing them to the reductionism and pre-emptive mode of operation of the algorithm, it argues that introducing algorithmic risk scores to ethically-laden decision-making might enforce a more punitive logic in acute psychiatry.

Finally, *chapter five* continues the present chapter's reflection on clinical AI by returning to Steyerl's work. In looking back to the analyses presented across chapters, I answer the questions guiding this study and assemble a number of theoretical, methodological and practical implications stemming from my (field)work.

Chapter one
Digitalising healthcare work[7]

[7] An earlier version of this chapter was published as: Carboni C, Wehrens R, van der Veen R, and de Bont A (2021). Conceptualizing the digitalization of healthcare work: A metaphor-based Critical Interpretive Synthesis. Social Science & Medicine 292: 114572.

Introduction

Digitalisation, "the simultaneous collection, analysis, and manipulation of digital data in real-time," (Trittin-Ulbrich et al. 2020: 10), has been discussed in the last few years as the inevitable and promising future of healthcare. This conviction spans medical literature (Blease et al. 2019; Bourla et al. 2018; Smith Glasgow et al. 2018), consultancy reports (EIT Health and McKinsey 2020; Topol Review 2019), as well as national and international governance strategies (Høyer and Wadmann 2020). The coupling of digital technologies and healthcare is posited as benefitting healthcare systems through increased efficiency, improved access, better allocation of scarce economic and human resources, and more resilience in the face of emerging demographic challenges (Andreassen et al. 2018; Greenhalgh et al. 2012; May 2001; Stevens et al. 2018). Especially in the medical literature, however, there is a general recognition that many of these promises have so far failed to materialise. Although many digital technologies have entered the market, issues such as non-use, resistance and workarounds continue to plague their embedding in the healthcare work practices (Callen et al. 2006; Li et al. 2019). Recently, the NHS-commissioned Topol Review (2019) has identified the broad category of digital healthcare technologies, spanning genomics (genome reading and genome editing), digital medicine (ranging from telemedicine to VR), AI (from natural language processing to predictive analytics), and robotics, as potentially bearing the greatest impact for the practices of the healthcare workforce.

Building on this, it is our assumption here that issues pivoting on embedding of new technologies in work practices are not simply a question of individual unwillingness to engage with innovation, but have to do with the ongoing reconfiguration of the healthcare workforce spurred by digital healthcare technologies. Previous work building on the sociology of profession testifies to the far-ranging implications that technological innovation bears for work-related dynamics (De Bont et al. 2016; Meyer and Paré 2018; Zetka 2001). This urges us to move away from individualised solutions aimed at improving education of and communication with the healthcare workforce, often proposed in the medical literature (Schuster et al. 2018; Smith Glasgow et al. 2016), and to take tensions and workarounds seriously.

In what follows we look at the digitalisation of the healthcare field in its professional and organisational dimensions. Dynamics of technologically-driven professional change have been analysed across many academic disciplines, yielding varying conceptual approaches and empirical findings. Yet, despite their apparent diversity, it is our contention here that different approaches and findings can be reconciled into a nuanced but coherent framework. We thus conduct a Critical Interpretive Synthesis (CIS; Dixon-Woods et al. 2006) bringing together theorisations and insights from (medical) sociology, (digital) medicine and Science and Technology Studies (STS). Our main research question is: *How have the implications of digital technologies for healthcare professionals and organisations been conceptualised and described in the medicine, sociology, and STS literature, and what lessons can we learn by bringing together these insights?* If sociology can offer insights into the professional dynamics of digitalisation, and the medical literature grants us access to emic conceptualisations and first-hand experiences of working in digitalised healthcare, the STS literature provides us with valuable tools to foreground the specific role of technologies in this sociotechnical transformation. As we discuss below, the CIS methodology enables us to foster conceptual and empirical cross-pollinations between different academic fields.

Methodology

Formalised by Dixon-Woods et al. (2006), Critical Interpretive Synthesis (CIS) is a review methodology geared towards theory production. Unlike aggregative syntheses, CIS looks at the literature not so much as a source of data, but rather as a repository of concepts. This allows to bring together studies rooted in different disciplinary and methodological traditions (Flemming 2009) — which, in turn, makes it possible to establish "cross-disciplinary knowledge translation[s]" (Abrishami et al. 2017: 14). CIS's focus on theory-making translates into quality appraisal criteria that centre the relevance and insightfulness of the concepts produced by the literature examined, rather than its methodological rigor (Dixon-Woods et al. 2006; Flemming 2009). The "critical" part of CIS is thus directed to problematising the way issues are framed in the literature, and the assumptions underpinning this framing (Dixon-Woods et al. 2006), with the overarching goal of generating new ways of looking at the issue at hand, and new possibilities for tackling it (Abrishami et al. 2017).

CIS enabled us to synthetise insights from medicine, sociology and STS — bodies of literature rooted in different methodological and epistemological traditions, but that have demonstrated a keen interest in the question of the digitalisation of healthcare work.

Literature search

We conducted our literature search among articles published after 2000 in the top 10 journals for each field considered. To identify the relevant journals, we relied on the 2018 Impact Factor ranking, as listed on Web of Science's *Journal Citation Report*. We added to it thematic journals of particular interest, such as journals focusing on medical sociology, critical data studies, or digital medicine (cf. table 1 in appendix). Two of the included journals, *Digital Health* and *Social Science & Medicine*, despite consistently publishing STS research, were intrinsically interdisciplinary, and thus did not fit easily within disciplinary boundaries. For these journals, we categorised articles on a case-by-case basis, based on how author(s) framed each article (for more details, cf. appendix).

Our search strategy, which combined manual and database searches, is described in detail in our protocol (cf. appendix). We first selected articles based on their title, and determined further inclusion based on the criteria listed in box 1.

An article's abstract is included if its title meets the following criteria:

a. Refers to medicine and/or health or some variant thereof AND
b. Refers to data and information technologies (either in general or to a specific one).

An article will be included if its abstract meets all the following criteria:

- Primarily focuses on health care professionals;
- Primarily focuses on digital healthcare technologies (either in general or on a specific one);

- Focuses on technologies used by professionals (also jointly with patients or relatives);
- Is based on empirical research or on a review of relevant literature (i.e. no opinion pieces or commentaries);
- Establishes a link between technology use and (changes in) work or professional practices;
- If focused on doctor-patient communication, clearly discusses implications for doctor’s role.

Box 1. Inclusion criteria

Our search retrieved a total of 126 articles. For 27 of these, the application of the specified criteria was not uncontroversial. These cases were discussed jointly by the first two authors until consensus was reached. As a result, 18 of these articles were excluded, bringing the total number of included articles for the three disciplines to 108 (cf. fig. 1-3 in appendix).

Data analysis

CIS builds on an abductive approach (Timmermans and Tavory 2012), which entailed progressively refining our “tentative, fuzzy and contested” review question through the encounter with the literature (Dixon-Woods et al. 2006).

Our aim was to create a multidisciplinary conceptual framework to articulate and investigate the digitalisation of healthcare work. Based on our review question, and on a sample of 15 articles (5 per discipline, selected based on the relevance of their title and abstract), the first two authors jointly defined some preliminary variables for the analysis. Like the review question, these variables were also abductively refined as the literature was coded (table 2).

We generated 33 codes for sociology, 31 for STS, and 32 for medicine. Most of these codes overlapped somewhat across disciplines, and fitted into the previously specified variables. As detailed in table 2, after several rounds of consultation between the first two authors, we decided to split the variables “What are the consequences for professionals?” into several sub-variables traced along different axes of professional work: nature of work, social relations with patients, social relations with other professionals, and emotional and psychological implications. This process allowed us to both obtain further analytical sophistication, and to do justice to previously unexpected points of interest consistently present in the literature. The following step in the CIS method required us to generate synthetic constructs that would interpret empirical evidence and transform it "into a new conceptual form" (Dixon-Woods et al. 2006). Since our final aim was to bring three different disciplinary fields into conversation with one another, synthetic constructs needed to select and tie together the main themes in of each corpus, while also articulating their assumptions and main insights.

Variables	How were they obtained?
• Which technology is considered? • Which professionals are considered? • In which country is the analysis conducted? • How is the relationship between the technology and professionals conceptualised?	Deduced from review question and literature sample

• Which implications are described?
    • Implications for individual professionals (nature of work/practices);
    • Interprofessional implications;
    • Implications for patient-provider relationship;
    • Emotional implications;
    • Trade-offs.
	Specified through coding the literature

Table 2. Overview of variables

We tackled this challenge by creating metaphors. Lakoff and Johnson's (1980) cognitive theory of metaphors shows how metaphors provide concrete, familiar, and often embodied signifiers that facilitate the understanding of abstract concepts. The creation of cognitive metaphors is thus integral parts of social processes, and scholars have analysed the ethico-political performativity of metaphor creation and, especially, naturalisation (Felt 2014; Puschmann and Burgess 2014; Wyatt 2021). Albeit less conscious and recognizable than poetic metaphors, cognitive metaphors also suggest specific visions and emphasise certain aspects of reality, while hiding others (Lakoff and Johnson 1980). This non-innocence warrants scrutiny of technology metaphors in particular (Wyatt 2004): metaphors hide certain aspects of technologies and naturalise others. Building on this, Wyatt (2021) has recently urged critical scholars to not only deconstruct "metaphors of the powerful," but to engage themselves in the "careful and imaginative" (406) production of new ones.

Our endeavour partially responds to her plea. The metaphors we selected build on often implicit and naturalised metaphors already circulating in each field. So, for instance, mobilising slime moulds enables us to explore the ever-emerging networks central to STS, while theatrical performances articulate the tension between visible and invisible, formal and informal stages that runs across sociological articles. Finally, river engineering allows us to focus on the assumptions and consequences of the concept of workflow — a metaphor as omnipresent in medicine as it is under-problematised. This abductive exercise in metaphor creation thus enables systematising and making explicit metaphors (at least partially) present in each field's predominant conceptualisations. Metaphors let us foreground tensions, assumptions and insights inherent in each field's conceptualisation, thus moving beyond particular cases and bringing the contribution of a specific body of literature to a more abstract yet operationalizable plane. Moreover, the fact that our metaphors build on implicit ones already mobilised, and sometimes naturalised, in each of the disciplines we analysed, ensures that our choice of metaphors is not an arbitrary one. Although we do not aim here to produce metaphors to reimagine technological futures, as Wyatt (2021) calls for, the metaphors we propose, in their describing prevalent conceptualisations in different strands of literature, can hopefully be productive in stimulating scholars' engagement in reflexive metaphor creation.

Results

In this section, we describe the main themes emerging from each body of literature analysed (table 3 in appendix). As summarised in table 2, we focus on the two variables "How is the relationship between the technology and professionals conceptualised?" and "Which implications are described?" (with its sub-variables). Combined, these two variables provide insights into the ways each corpus conceptualises and describes the digitalisation of healthcare work. Even within a single discipline, however, analytical foci and empirical findings often differ greatly. To facilitate our synthesis, we mobilise three metaphors, one for each discipline. We think of these metaphors as focusing tools, pointing our attention to specific contributions of each body of literature and materialising connections amongst the most prevalent themes within it. Thinking through metaphors enables us to articulate assumptions and practical implications emerging from each body of literature.

Thinking like a network: How slime mould helps us bring together the STS literature

Building on the prevalence of network thinking in STS analyses and its importance in tying together the main themes, we propose slime mould as synthetic metaphor. Slime mould (*Physarum Polycephalum*) is a peculiar organism: a single, giant cell comprising many nuclei that share the same cell walls. Thanks to this, information (in the form of a not-yet-quite-specified signalling molecule) can flow across its organism, carried by rhythmic peristaltic movements (Pringle 2019). These movements, and the information exchange they enable, allow *Physarum* to exhibit a behaviour that has been described as "learning," "remembering," "solving problems," "making decisions," despite its lack of a central nervous system (Jabr 2012).

Slime mould is intelligent not because of the presence of a brain, but because of a constant, distributed flow of information across the interconnected parts of its body. It is able to explore its surroundings by sending out its tendrils to explore its habitat, sensing resource-rich patches, or porridge oats positioned by scientists. It "explores territories in multiple directions simultaneously" (Barnett 2014), covering surfaces in complex interconnected patterns. Once it finds food, it rearranges its body into the configuration that allows it to "optimally eat and reproduce" (Pringle 2019) by retracting all but the shortest tendril connecting its body to the food. Slime mould 'works' because it is a network. And slime mould is a network that works. As such, it is uniquely positioned to help us explore the STS literature. In what follows, we explore how this metaphor enables us to synthetise the diverse insights of the STS literature, uncover their assumptions and articulate their implications. We argue that slime mould teaches us to think like a network: to focus our attention on interconnections and their temporariness, to stay open to reconfigurations, to not consider unpredictability a problem in itself.

Conceptualisations: Networks and materiality

The STS literature mostly considers digital healthcare technologies in their process of becoming part of a network of human and nonhuman actors. The networks STS postulates are "a densely interconnected assemblage of actors, actions, and relationships" spanning "users, other technologies, rules and regulations, institutions, and a variety of other heterogeneous elements…" (Nicolini 2006: 2756; cf. also Pols 2011; Winthereik et al. 2007). Like slime mould, which is itself a living network incorporating particles of different origin, STS networks are pulsating, constantly integrating new elements. *Physarum*'s relationship to the surrounding environment is also a good signifier for the exploratory nature of the network-making described in the STS literature. Within the networks that STS postulates, agency is diffused among humans and nonhumans, and materialises in mutual negotiations between technologies and human actors. Similarly, each extremity of a slime mould is endowed with agency, tinkers with the environment (surfaces, food, other parts of the slime mould) and creates new connections.

Not unlike slime moulds behaviour, network-making in STS is also an exploratory endeavour, proceeding in multiple directions: through tinkering, human actors try out different ways of integrating new technologies within pre-existing practices (Danesi 2020: 18). The outcome of this tinkering is hard to predict, and networks are always open to reconfigurations (Greenhalgh and Stone 2010) — just like slime moulds reconfigures its body if a patch is depleted of resources, or if new food is introduced in their environment. Nonetheless, not just any reconfiguration is possible: the semiotic-material aspects of technology are crucial. STS concepts such as technological scripts (Danesi 2020; Galetsi et al. 2019; Greenhalgh and Stones 2010; Nicolini 2006; Oudshoorn 2011; Winthereik et al. 2007) and affordances (Abrishami et al. 2014; Trondsen et al. 2018), emphasise how technologies embody values and visions for practice, and how their material and symbolic properties prescribe specific uses and users (Spatar et al. 2019).

If tinkering happens within pre-designed boundaries (Danesi 2020; Pols 2011; Trondsen et al. 2018), its results are an open-ended, empirical question:

> … the patterns of use inscribed in the artifact by the designers only come to life in the context of the daily activity of the users. When put to work, the concrete anticipations and restrictions of future patterns of use embodied in the technological artifact interact in complex ways with the existing work practices of the users. The result is a process of negotiation between the innovation and the work activity. The outcome of such negotiation determines, on the one hand, how the innovation is used "in practice"; at the same time, it produces some kind of change in the work practice, usually along lines which reflect (to some extent) the desires and intentions of the designers and their sponsor (Nicolini 2006: 2757).

Changing work practices through the introduction of technology is thus anything but a straightforward process with a certain, foreseeable outcome: it depends on a process of negotiation in which multiple actors are involved. However, concepts such as technological scripts and affordances teach us that technology can be a powerful way to steer the growth of a network. Likewise, slime mould does not grow just anywhere: it looks for resource-rich patches in its environment. As bio artists explain, when working with *Physarum*, steering can only be a partial accomplishment. A network's behaviour can be guessed, but not controlled: "the slime mould has the final say in the creative process" (Barnett 2014).

Which implications are described?

Clashing implications for practice

The STS literature does not describe the implications of digital healthcare technologies as coherent and unidirectional. Although it is acknowledged that the introduction of new technologies inevitably entails a change in professionals' (and patients') practices, technologies are mostly observed to open a range of (potentially contradictory) possibilities for change. All sorts of implications are described, to the extent that synthesising them in a coherent narrative is acknowledged as problematic: as Petrakaki and colleagues sum up, "possibilities are endless" (2012: 436).

Thinking through slime mould, however, we can come to embrace unpredictability and situated tinkering. Not unlike STS, slime moulds also point our attention towards distributed agency. In the case of slime mould, agency is about direction. *Physarum* is *Polycephalum*, has many heads and each of them moves simultaneously in a different direction. Each head is responsible for optimally interacting with the environment. In STS terms, they do their own, independent yet interconnected, form of tinkering. Slime mould begs us to shift our attention from the final shape of the network to the ongoing dynamic of network-making. Slime mould can continue its work of exploration as long as information can circulate through its body. And for that flow to happen, interconnections between different parts of the body are crucial.

Though avoiding deterministic stances, STS literature acknowledges that digital healthcare technologies steer healthcare practices in a specific direction. Analyses often point to an increased reliance on quantitative data in the diagnostic and treatment process, and the (potential) loss of qualitative information (Mort et al. 2003; Reich 2012).

Patient-provider relationships: Invisible work and delegation

Invisible work is required for individual professionals to accommodate technologies into their daily practices, especially in patient-provider interactions. This kind of work, necessary but unacknowledged by organisations, is mostly relational, geared towards explaining technologies to patients and reassuring them of their abilities to operate them and making the interaction with patients smoother in spite of interruptions caused by the technology (Mort et al. 2003; Nicolini 2006). Invisible work made necessary by the limitations of a technologically-mediated patient-provider interaction also spans "sensory work" (Maslen 2017) — the exploratory work necessary for professionals to find ways to sense through technological sensors, to integrate their measurements into their diagnostic process, and to establish rapport with patients, technological mediation notwithstanding.

Digital healthcare technologies are often discussed as facilitating patient-provider communication overcoming limitations associated with geographical distances (Nicolini 2006; Pols 2011; Trondsen et al. 2018). However, making things possible is not the same as making things work. STS literature points us towards the fact that not geographical, but also "relational distance" should be considered crucial in patient-provider relationships. Technologies like telemedicine magnify the relational distance already existent between patients and providers: they work in relationships in which trust has been established, but "add to the strangeness or indifference experienced" (Pols 2011: 466) when they connect strangers.

When human interconnections are mediated by technology, they need to be reestablished at both the communication and the sensorial level. Likewise, making stable connections takes work on the part of the slime mould: its body is stretched out, then retracted, severing the interconnections that do not "fit." This is where invisible work emerges in the STS literature: when old relationships are replaced by new, complex and unstable ones, actors in the network (especially professionals) need to go to extra lengths in order to make the new connections "work."

Interprofessional relationships: Boundary renegotiation

When it comes to interprofessional relationships, STS literature generally acknowledges that digital healthcare technologies lead to a renegotiation of professional boundaries. Pre-existing professional roles can be expanded and take up new tasks (Burri 2008; Danesi 2020; Winthereik et al. 2007), new roles can emerge (Galetsi et al. 2019), and tasks and responsibilities can be delegated to other professionals (Burri 2008; Maslen 2017; Mort et al. 2003). These negotiations can lead to tensions between professional groups, that react by mobilising their professional identities (Burri 2008; Greenhalgh and Stone 2010).

Digital healthcare technologies (especially EHRs and telemedicine) are also acknowledged to facilitate interprofessional communication across space and time. They enable more frequent, more structured and faster exchanges of information among professionals, and for the creation of larger, geographically dispersed networks (Lehoux et al. 2008; Nicolini 2006). However, not all information travels easily among professionals. Particularly in the absence of relationships of trust among professionals, whereas quantitative information (measurements, images) is trusted, qualitative information (e.g. opinions, speculations, patient histories) is easily dismissed (Greenhalgh and Stone 2010). Technologies mediating interprofessional communication thus often fail to contextualise quantitative information: they "take the histories out of patients," but that qualitative "knowledge about patients (their psychosocial states, anxieties, worries, and fears; their family circumstances, and so forth) has to be rebuilt later" through "a patchwork of other kinds of activities and materials, such as reassurance, explanation, history taking, intuitive investigation, skin and blood samples" (Mort et al. 2003: 285). This exemplifies the sort of invisible work required to make technologies work in the interactions with other professionals.

Trade-offs of technological innovation

The way slime mould retracts its body when new interconnections have been forged is a good visualisation of the implicit but virtually omnipresent assumption in the STS literature that, when there is a reconfiguration of the network, something is bound to be lost. In the context of the provider-patient relationship, loss is about qualitative information about the patient, the recording of which clashes with the scripts embedded in some digital healthcare technologies. A similar loss of information, this time sensory and non-verbal, stems from the technical limitation of some technologies that mediate the diagnostic process. In the context of technologically-mediated interprofessional communication, qualitative and non-measurable information is also likely to be lost or disregarded, especially in the absence of a pre-existing relationship of trust among professionals. Interprofessional relationships themselves are threatened, in a context in which in-person communication is made unnecessary by technology. In most cases, (possible) loss is seen to directly threaten the diagnostic process and the possibility of providing care — and to require invisible work on the part of professionals.

Main insights

Slime mould's behaviour teaches several things about the way the STS literature looks at the digitalisation of healthcare work. Firstly, we learn how expanding networks is an exploratory endeavour, hard to control or to predict, in which individuals tinker with the environment and its resources, with technologies and their affordances, trying out different possibilities. Far from grand narratives of technologically-afforded innovation, it is through this work of exploration, adaptation, incorporation, and potential establishing of new interconnections that change happens. This is a major insight afforded both by slime mould and STS analyses: changes in healthcare practices cannot be centrally planned, little control can be exerted over the implications that technological innovation has for practices and practitioners. One can work *with* slime mould and its daily, mundane negotiations, not on it.

Secondly, this metaphor encourages us to focus on the interplay between pre-existing relationships and the materiality of newly introduced technologies. Slime moulds live in a state of constant negotiation with the surrounding environment: they explore several, often opposite, possibilities for connections at once. But not just any connection is possible: although specific interconnections cannot be planned, slime mould can be steered to grow in a certain direction. This is acknowledged in the STS literature: the material-semiotic properties of a specific technology can, at least, encourage specific reconfigurations. Some interconnections can be discouraged, or radically changed. Materiality is powerful ways of steering the direction of the network's growth. We thus need to attend to technological scripts and affordances to better understand changes in work practices in healthcare.

Finally, slime mould urges us to reframe the very question guiding our review. The impossibility to predict how the network will ultimately reconfigure sometimes subtracts critical power from the STS analyses we considered. If we trust slime moulds, we can speculate that what matters is not the final shape of the network, but the strengths of the interconnections that allow various forms of information to circulate among its different parts. STS literature is particularly interested in how the network's reconfigurations impact pre-existing relationships between different actors (professionals, patients, technologies), or create new ones. Slime mould helps us foreground STS's call to focus on relationships rather than on outcomes. The strength of interconnections between the nodes of the network is what allows slime mould to behave intelligently even in the absence of centralised decision-making. If strong interconnections are preserved, information can circulate, and the network finds a way to thrive. What emerges from both slime mould and the STS literature is the necessity to preserve and support the interconnections that matter, and to attempt to steer the growth of the network in the desired direction through limited, but existing, technological means.

Understanding the sociological literature through theatrical performances

The synthetic metaphor we propose for sociology centres theatrical performance — a domain central in sociological theory since Goffman's (1959) dramaturgical theory. However, whereas Goffman's discussion centres the dichotomy between frontstage and backstage in order to make a point about self-presentation in different social situation, here we foreground the necessary coexistence of frontstage and backstage as the most striking contribution of the sociological literature to the conceptualisation of the digitalisation of healthcare work. If we usually think of performances as the work of actors on a stage, sociological analyses refocus our attention to include the incessant and unseen work of the backstage crew, and its crucial importance in enabling a smooth performance onstage.

The visible part of a performance takes place on a stage, and consists of the enactment of a script by a cast. Actors need to learn the script by heart, move on the stage and interact with each other and with props. Scripts determine how many actors are necessary, and assign to them different roles, from leads to extras; they can also be more or less strict in the directions they provide, ranging from scenarios, broadly sketching the main lines of the dramatic development, to play texts providing performers with strict directions. Even in the case of strict scripts, however, directors (sometimes in consultation with actors) retain the freedom to deviate from them in their staging of the performance — for instance by reassigning or rewriting lines, suggesting specific movements, and deciding on stage props in consultation with the stage manager.

During the performance, a considerable amount of work is done in the backstage and not seen by the audience. The stage crew, comprising several stagehands, set up the stage and operate lights, sound, different kinds of props and special effects. Their work is coordinated by the stage manager, who oversees the production throughout rehearsals and performances, spreading necessary information throughout artistic and technical departments, bringing the director's vision to life. During the performance, stage managers sit in a booth connected to the backstage and, communicating with the stage crew, they "make sure that the actors turn up whenever there's a scene change, … that everyone discharges their duties correctly, and there's no danger to anyone" (Donaldson 2013). As we elaborate in the following sections, this coordination work is crucial to sociology's conceptualisation of the digitalisation of healthcare work.

Conceptualisations: Technology-in-practice

The "technology-in-practice" approach dominates sociological accounts of the digitalisation of healthcare work (Bailey et al. 2020; Peiris et al. 2011; Reed et al. 2016). This approach is rooted in STS approaches and, as such, presents strong overlaps with the network conceptualisation (cf. above). It encourages researchers to reframe technology as an actor in itself, postulating that "what it does and how it accomplishes something remains an open empirical question" (Timmermans and Berg 2003: 104), thus sidestepping both technological determinism and social essentialism. This translates into a focus on the micropolitics of sociotechnical change, which considers technology's influence on medical practices as "multifaceted and unfolding" (Reed et al. 2016: 738), as well as situated within a network shaping the conditions for action and meaning-making. Technologies are thus investigated as central actors in the "construction and reproduction of novel worlds" (Timmermans and Berg 2003: 108).

Transposed to theatrical performances, the technology-in-practice approach encourages us to extend the concept of acting to a variety of nonhumans. Not only actors enact a performance: this work is distributed across a number of materials, ranging from script, stage and props, as well as audience and backstage crew. Moreover, the same cast is likely to perform the same piece in quite different ways from one day to the next. Similar to technology-in-practice, this metaphor teaches us that performances, albeit scripted and rehearsed, are difficult to control in their unfolding. This warrants open, empirical approaches to the study of the digitalisation of healthcare work.

Conceptualisations: Technology as steering

The technology-in-practice approach coexists with more normative conceptualisations, range from institutional ethnography-inspired views describing technologies as tools to "advance a hospital's strategically designed purposes" (Campbell and Rankin 2017: 366), to an empirically-grounded acknowledgment of the tendency to "deploy technologies to standardise and control work" (Findlay et al. 2017: 118). Despite being analysed as actors, technologies are often described as only taking up specific tasks and being aligned with specific managerial logics. If professionals' interactions with technologies can lead to new sociotechnical configurations, it is also true that this reframing tends to feature specific characteristics: it aims to make care provision more calculable and objective (Campbell and Rankin 2017) and to meet "targets imposed "from above"" (Mueller et al. 2008: 3). Therefore, technologies are conceptualised here as "bring[ing] norms to the clinic … that aim to direct [professionals] and patients to particular ways of caring" (Pols 2010: 377). The digitalisation of healthcare work thus unfolds in a polarised political arena, with professionals caught up "in the midst of a battlefield between their profession and the organisation … set by management" (Fältholm and Jansson 2008: 26).

Thinking through theatrical performances helps us understand how these two dominant conceptualisations need not be mutually exclusive. Like scripts, technologies can be more or less directive in the instructions the provide to performers. This sensitises us to the fact that, in the case of technologies aimed at imposing strong constraints to professional practices, the implications of digitalised healthcare work may be less of an open empirical question than in others. Moreover, even though scripts are embedded in specific dramaturgical traditions and assign roles, lines, and movements to actors, their enactment is an open question, shaped by the distinct but (ideally) aligned choices made by the director, the production, the backstage crew, and the actors themselves. The so-called theatrical hierarchy is thus a good parallel for sociological literature's insight into the managerial agency in reshaping healthcare work by enacting digitalisation through specific technologies.

Which implications are described?

Professional hierarchy and uneven consequences

Sociological analyses thematise the unevenness of technologies' implications for different professionals, materialising in the fact that only some professionals are in a position to be involved in the design and implementations of digital healthcare technologies (Fältholm and Jansson 2008; Petrakaki et al. 2016; Pols 2010). Consequently, some professionals have the opportunity to tailor technologies to their needs and values, while others need to adapt their practices to technologies' requirements (Barrett et al. 2012). Differential professional power thus engenders what Findlay and colleagues (2017) term a "polarisation in [job quality]" (118), with some professionals losing autonomy and control over their tasks while others experience an "enhanced … professionalism" (Barrett et al. 2012: 1463). This polarisation is symptomatic of how innovation, by pitting some groups' occupational professionalism against the norms and values embedded in technologies, ends up restricting some professionals' possibilities for engaging with the technology to resistance, non-use, and workarounds (Dupret 2017; Håland 2012; Mueller et al. 2008). Even when the relationship between professional hierarchy and undesired implications of technologies is not linear, and the latter generate unwanted tasks for professionals high in the hierarchy (Bar-Lev 2015; Petrakaki et al. 2016), a power differential materialises in professional groups' varying abilities to delegate unwanted tasks. Petrakaki and colleagues, for instance, found that doctors using a new EHR system kept delegating to nurses "time-consuming and unchallenging tasks that failed to match their perceptions of their work and role as doctors" (2016: 216), such as data and order entry.

Theatrical performances teach us that not all roles are created equal: some of them have more stage time and more lines than others. This depends greatly on the characteristics of the selected script. Directors, however, have the ability to change the script, and to reassign lines. The possibility, predicated on one's position in the hierarchy, to access those in charge of the performance thus bears significant consequences for how each actor is able to tailor their assigned role to their personal preferences. However, the sociological literature makes clear that, no matter how strongly a director intervenes on a script, the performance itself will always need intense, unseen work in the backstage. Each modification of the script creates new, invisible coordination work for the stage manager and the stagehands, as we unpack in the next section.

Visible and invisible work

The sociological literature centres invisible work, and consistently finds that "[t]echnology does not replace human actors but introduces work for patients and healthcare professionals that is not represented in dominant discourses …" (Oudshoorn 2008: 283; cf. also Fältholm and Jansson 2008). Analyses centre invisible work as emerging in the patient-provider relationship in the form of tasks such as introducing patients to new technologies, reassuring them of their ability to use it, and reminding them to use it (Mossfeldt Nickelsen 2018; Schwennesen 2019). In technologically-mediated patient-provider relationships, emotional work is also needed to (re-)create intimacy (Heath, Luff and Svensson 2003; Lupton and Maslen 2017). In interprofessional relationships, invisible work is about making technologies work in practice, and coordinating them with the work of other professionals (Bailey et al. 2020; Håland 2012). Repair work is also necessitated by professionals' workarounds, for instance when their selective use of technologies invalidates the reliability of alarms and flagging mechanisms (Pols 2010).

The backstage crew supporting and enabling the performance on the stage, and in particular the work of the stage managers, provide the most fitting metaphor for the invisible work described in this literature. The unseen work performed backstage is crucial in bringing together scripts, actors, technical props, and directors' vision in a seamless performance. Without this work of coordination, complex productions could simply not take place, and would be disrupted at the first technical hiccup. The work of the backstage crew is both determined during the rehearsals, thus based on a director's interpretation of the script, and ongoing, constantly adapting to the specific circumstances of a particular performance.

Sociological literature also presents accounts casting technologies geared towards transparency and accountability as redistributing invisible work across the professional hierarchy. In this case, technologies such as the EHR are said to work "as a mechanism that … imposes fairness, ensuring a clear division of labour between professional groups" (Petrakaki and Kornelakis 2016: 234; cf. also Bar-Lev 2015). As mentioned above, the illusion of the seamlessness of theatrical performances is predicated upon the invisibility of the work carried out backstage.

Implications for interprofessional relationships

As we have begun to see in the previous paragraphs, repercussions on professional hierarchies found in this literature are contrasting. Sociological accounts often find that technologies are designed with traditional hierarchies in mind. Once inscribed into technologies, interprofessional relationships ossify, and hierarchies become harder to negotiate. Technological mediation of interprofessional relationships results in a reinforcement of professional hierarchies (Halford 2010). A typical example of this is the fact that EHRs often restrict access to some clinical information to some categories of professionals (Bar-Lev 2015). Formalising interprofessional relationships can thus have contrasting implications, and this has to do with the varying ways professional hierarchies are built into technologies' design. Similarly, scripts need to be interpreted by directors and actors in order to adapt them for a specific performance.

Clashing implications are also described with reference to interprofessional communication. On the one hand, technologies facilitate information transfer, providing "legible clinical notes and requests, fast exchange of information, instant capture and access to data and increased visibility of diagnoses, procedures and test results" (Petrakaki et al. 2016). On the other hand, technologically-mediated interprofessional communication is less tailored to the specific interlocutor and, being accessible by more parties (different professionals, patients, and legal actors in case of lawsuits), tends to be reduced to the transfer of strictly necessary and quantitative information. "Subjective information, uncertain information and additional practical or 'extra' information" (Engesmo and Tjora 2006: 182) is thus lost — an argument against the complete formalisation or technological mediation of interprofessional communication (Bailey et al. 2020; Maiers 2017). Vital information circulates informally, as testified by the work of stage managers.

To communicate with the rest of the crew and with actors, stage managers are connected to the backstage through a complex apparatus of audio technologies, switches and light cues. However, this communication needs to remain unseen in order to be effective and not to disrupt the performance itself.

Trade-offs of technological innovation

Not unlike STS, the sociological literature also reflects on the losses associated with the digitalisation of healthcare work. This pivots on the tension among different modes of knowing embodied by different professionals. Digital technologies are described as implicated in the production of objective knowledge, anchored in quantitative clinical data. This contrasts with the more idiosyncratic, embodied and long-term knowledge of patients and their health that can be developed by healthcare professionals (Halford 2010; Maiers 2017).

Technologies aimed at structuring the knowing, such as strict assessment forms embedded in EHRs (Bar-Lev 2015) or algorithms allocating timeslots for consultations (Campbell and Rankin 2017), constrain the conditions of professional judgment (how and what is known about a patient). For technologies such as clinical decision support systems, that aim at doing (part of) the knowing, the lack of a fit between human and technological modes of knowing results in the selective reliance on the technology itself in the process of diagnosis (Bailey et al. 2020; Maiers 2017). External pressures for legitimation, cost-cutting and efficiency may increase the reliance on these technologies in the process of diagnosis, which would end up "potentially removing idiosyncratic knowledge of particular patients from the constellation of information by which clinicians determine patient conditions" (Maiers 2017: 927). Based on similar reflections, the sociological literature urges to acknowledge both the limitations of technological innovations, and the "abilities, work practices and ... social competencies" of the professionals that some of these technologies aim to replace (David et al. 2009: 935).

Main insights

Looking at sociological analyses through the metaphor of theatrical performances helps bring together some of the important insights they provide. Firstly, sociological conceptualisations of the digitalisation of healthcare work nuance our understanding of the steering potential of technologies. As also articulated in the STS literature, technologies present specific affordances that imprint directionality to the changes in digitalised healthcare work. However, technology, on its own, underdetermines changes in healthcare work. Considering theatrical performances, we have learnt that directors and actors can change how a script is enacted. It is a question of power and time: only the directors and some actors have a say in this process, and after decisions are made in the initial stages of a production, it becomes progressively harder to change the script. In the context of healthcare work, this means that implementation choices (and who is involved in them) matter in how healthcare work changes, and that they progressively ossify, becoming more difficult to unmake.

Theatrical performances and the stage-backstage duality also shows us that invisible work is always needed to coordinate professionals and technologies. This is especially true for the embedding of new technologies into preexisting situated professional practices and organisational structures: this requires extra invisible work, just like the work of stage managers is intensified by modifications to the script or disruptions to the performance. A major lesson from the sociological literature thus concerns the importance of recognising the value of invisible work without trying to make it visible, but rather providing spaces within organisational structures for carrying it out.

Medicine: Understanding workflow digitalisation through river engineering

The medical literature often uses metaphorical language that has consequences for how healthcare work is imagined. In this section, we build on the concept of workflow and its casting of healthcare work as something that needs to “flow.” We thus attempt to probe the assumptions and implications of this conceptualisation by reimagining the digitalisation of healthcare work through the metaphor of river engineering.

In order to make rivers and their cycles more useful and less disruptive to their activities, humans have been engaging in various forms of river engineering for centuries. Engineering interventions have aimed at straightening the course of rivers to improve navigability or speed up their flow. Indeed, although rivers naturally flow in one direction, their course often meanders, which decreases the speed of their flow. Dams and dikes have been built to manage river flow and prevent flooding, which is a part of the natural cycle of most rivers. If the goal of river engineering is to foster the fit between rivers' nature and human needs, river engineering has been often associated with several risks and unintended consequences (EPA 2016). Many of these interventions backfire: straightening rivers makes the water flow more rapidly, thus increasing the risk of floods downstream. Dikes only protect the area around them, but can increase flooding and water pressure both up- and downstream.

In what follows, we show how the idea of workflow bears implications for conceptualisations of healthcare work in the medical literature, and how reading these analyses through the metaphors of river engineering enables us to materialise their insights.

Conceptualisations: Promises of meaningful and efficient work

The medical literature conceptualises digital healthcare technologies as holding two overlapping promises for healthcare professionals: making their work more meaningful and more efficient. Meaningful work generally coincides with patient care (Blease et al. 2019; Grünloh et al. 2016; O'Malley et al. 2015; Westbrook et al. 2013; Zadvinskis et al. 2019). The meaning of patient care emerges negatively, from its opposition to supposedly menial tasks such as documentation and technology-related clerical work (Callen et al. 2006; O'Malley et al. 2015; Sieja et al. 2018; Tai-Seale 2019; Tran et al. 2020; Zadvinskis et al. 2018).

What makes the tension between meaningful and non-meaningful work particularly painful is the fact that time is, for healthcare professionals, a particularly scarce resource. This turns the tension between meaningful and non-meaningful work into a zero-sum game, and sets the terms for identifying tasks which can be delegated to technology (Vogel et al. 2015). Indeed, the automation of these menial tasks is considered desirable and straightforward. The fact that automation is discussed within a context of time scarcity turns the digitalisation of healthcare work into a quest for efficiency, an attempt to get rid of unnecessary tasks, or at least speed them up (Grassl et al. 2018; Hains et al. 2012; O'Malley et al. 2015; Sieja et al. 2018; Vogel et al. 2015).

If we think healthcare work through the metaphor of river engineering, we can think of meaningless work as meandering. River engineering, and the straightening of rivers in particular, point us to the importance of efficiency as discussed in the medical literature. The goal of accomplishing as much as possible as quickly as possible, and with the least possible effort, is never questioned in this literature. However, by nature, rivers meander and slow down. Engineering interventions are needed to speed them up. Minimising the deviations of the river flow allegedly results in more directedness, less waste of energy, a better accomplishment of human goals. Similarly, the quest for efficiency in healthcare is framed as requiring technological innovation: workflows become something that digital technologies can streamline (Hains et al. 2012; Lærum et al. 2003). And streamlining workflows also means reducing deviations, making healthcare work more focused on its meaningful components. Thus, technology is hardly questioned in its role as solution to the problem of time scarcity.

Conceptualisations: Disattended promises and the importance of design

The medical literature acknowledges that digital healthcare technologies often fail to deliver on their promises. Like operations of riverbed straightening, intervening on healthcare work through digital technologies entails risks. In some cases, things do go as planned, and the flow of work is not hindered (Petersson and Erlingsdóttir 2018). In other cases, river engineering can make the river flow more dangerous, increasing the chances of flooding. Technologies are found to cause interruptions to workflow, introduce new errors (Tran et al. 2020) and time-consuming tasks (Strand et al. 2017; Tai-Seale et al. 2019; Westbrook et al. 2013), disrupt interprofessional communication and patient-provider relationships, and sometimes drive the early retirement of physicians. Professionals react to technologies' malfunctioning with resistance, workarounds, or inefficient and selective use

(Fisher Wilson 2009).

The medical literature frames the fact that technologies disattend their promise of more meaningful and efficient work predominantly as a problem of bad design, oblivious to the fact that technology "needs to fit with the workflow of physicians and within the organisational framework of accepted practices, norms and structures" (Callen et al. 2006: 644). Bad technological design is thus anything that contradicts the promise of streamlining, from complicated interfaces (Fisher Wilson 2009) to workflow blocks generating workarounds (Grünloh et al. 2016; Li et al. 2019; Vogelsmeier et al. 2009). Some articles broaden the issue of design by emphasising the need to redesign workflows as part of the implementation process (Vogelsmeier et al. 2009), and to give physicians the possibility to customise the technologies (O'Malley et al. 2015).

Centring technological design as a way to address the disattended promises of digital technologies for healthcare work enables the medical literature to maintain its pro-engineering stance without disavowing the promissory value of technology. Allegedly, the problem does not lie in the river straightening operation in itself, but rather in the bad fit between the specific way the water flows in a certain river and the way the artifacts have been designed and built into the river itself. Bad technology design can be addressed by taking more seriously current workflows and needs of professionals.

Which implications are described?

Work practices: positive implications

Some of the articles we analysed aimed at empirically testing technologies' promises, or the commonly held assumption that these promises are not met in practices. In a few cases, results were positive: patient information was more thorough and easier to access, EHR-initiated reminders were experienced positively, and workflows were simplified (Zadvinskis et al. 2018). Some transcription technologies alleviate burden of documentation (Vogel et al. 2015) and are found to decrease error rates in medication management (Westbrook et al. 2013). Enthusiasm for new technological systems also stemmed from the fact that they did not cause expected disruptions, for instance by not worsening patient-provider relationships nor increasing documentation requirements (Grünloh et al. 2016).

Provider-patient relationships

The medical literature often expresses concerns as to the limitations imposed by technology-mediated communication, and how this affects the relationship to patients (e.g. physicians' ability to pick up on nonverbal cues and emotions, loss of empathy; Schuster et al. 2018; Blease et al. 2019). However, relationships with patients are found not to be negatively affected by the use of digital healthcare technologies, although it is acknowledged that maintaining them entails different degrees of effort from different specialists (Cresswell et al. 2018; Vorderstrasse et al. 2014).

Another major issue in this sense is the alleged erosion of physicians' authority (Cresswell et al. 2018; Grünloh et al. 2016), especially in the context of shared decision-making. This materialises in particular in EHR's potential to give patients access to their own health data. Giving patients access to EHRs often results in changes in physicians' documentation practices. To protect themselves legally and not to cause distress to their patients, physicians tend to limit the amount of information they record, thus "watering down" the EHR (Grünloh et al. 2016; Petersson and Erlingsdóttir 2018). Transposed to the river metaphor, the already very visual concept of watered-down EHRs reminds us of dikes and dams being overwhelmed by an increased waterflow as a result of a new tributary being connected to the river. In this case, new, informal ways of damming the river flow need to be devised.

Desktop medicine, emotional implications and burnout

The medical literature acknowledges that healthcare work is highly stressful in and of itself. However, some articles point to technologies as making the situation worse. One of the main implications of the digitalisation of healthcare work is identified in the increased documentation requirements and prominence of what is often termed "desktop medicine" — "the computer-based clerical work associated with patient care" (Sieja et al. 2019: 793) and that often takes up most of physicians' time (Tran et al. 2020: 809). Desktop medicine, epitomised by the EHR, is the clearest manifestation of technology's disattended promise of increased efficiency: not only does it take time away from patient care, but it creates a 24/7 work environment that, coupled with user interface issues and excessive amounts of notifications, drastically increase work-related stress and frustration:

> A growing research literature suggests that time spent by physicians on the EHR has been linked to their reduced satisfaction with work. … over 50 percent of their time is spent on desktop medicine tasks … it is important to carefully examine the relationship between pivotal aspects of desktop medicine and physicians' well-being (Tai-Seale et al. 2019: 1073-4).

These issues are usually analysed as related to physicians, and are associated with the risk of burnout and early retirement. Emotional consequences for nurses are considered far less frequently, and point to the stress generated by the need to navigate communication and relationships with patients through online portals, as well as frustration when task delegation from doctors is perceived as excessive. Still, delegation of "non-meaningful" tasks to other professionals is frequently proposed as a way of dealing with physicians' risk of burnout (O'Malley et al. 2015; Tai-Seale et al. 2019).

The medical literature acknowledges that both stress and some degree of clerical work are inescapable components of healthcare work: similarly, rivers naturally flood, meander and slow down throughout their course. However, just like with river engineering, technologies can worsen this situation by, for instance, increasing documentation requirements, or introducing the need for remediating technological disruptions to the workflow. Interestingly, however, the solutions proposed for this issue are generally the redistribution of tasks generated by technology, improved technology design, or the automation of time-consuming tasks. Technology can be tweaked and adjusted, but the pro-engineering stance of the medical literature prevents it from substantially questioning the conceptualisation of technology as a solution.

Main insights

River engineering shows us several things about medical literature's conceptualisation of healthcare work. Rivers capture its naturalistic depiction of the essence of this work: like rivers flow in one direction, healthcare work also has a natural course, geared towards care delivery. Like healthcare work, rivers are unpredictable (rainfalls can always cause them to overflow), periodically flooding, naturally meandering, slowing down and speeding up at different points of their course. The engineering metaphor highlights how encouraging the direction of (work)flow is the role that the medical literature assigns to technological interventions. Indeed, the promise of technologically-attained efficiency and meaningfulness is a crucial motor for technological innovation in healthcare. In practice, however, many digital technologies make some water overflow, thus wasting important resources and frustrating professionals.

Rivers and their engineering can help us reflect on some of the tensions inherent to medical literature's conceptualisation of the digitalisation of healthcare work, and give us insight into the reasons of the persistently hard fit between technologies and medical practice. A first tension concerns the pro-engineering stance of the medical literature, and its subsequent singling out of technological design as a crucial dimension for improving the fit between technologies and practices in healthcare work. As pointed out throughout this section, although technologies mostly disattend their promises, their promises of improved care delivery and working conditions are never disavowed, and are instead turned into a quest for more "fitting" design.

Secondly, we have learnt that river engineering does not only intervene on the water flow in the exact point in which the infrastructure is installed. Engineering a specific part of the river, by either straightening its course or tinkering with its banks, despite perhaps decreasing the risk of flooding in that area, can actually increase risks up- or down-stream. This points us to the problem of task delegation as a solution to physicians' technologically-driven increased risk of burnout. Delegating unwanted tasks to other professionals may hide them from sight, but will not get rid of them. The generation of unwanted tasks is not a negligible inconvenience, but a serious consequence of introducing digital healthcare technologies. How we deal with it is a crucial political question.

Synthetising argument

How do metaphors fit together?

Metaphor have taught us several things about the digitalisation of healthcare work. River engineering's lessons centred the fact that digitalisation follows a specific direction, geared towards efficiency, rationalisation, transparency. We have learnt that there is a direction in which the medical field wants healthcare work to develop, made of cost-cutting and time-saving, and that goal-setting is fundamental in fostering innovation. But we have also learnt that forceful streamlining can cause flooding: relying on a single view of healthcare work as efficient and rationalised can engender a host of problems for the healthcare workforce. Slime moulds help us here: as living networks, they taught us that healthcare practices change not through a top-down imposition of a disruptive vision, but through feeding interconnections among humans, and between humans and technologies. From slime moulds we learn that we can steer a network's growth through feeding it in a certain direction — but the shape the network itself will actually assume is a matter of co-creation. Finally, the coexistence of front- and backstage in theatrical performances has taught us about the importance of the unseen work that supports network growth — about what it takes for information to flow within the slime mould's organism. Without stage managers and stage crew, performances on the stage would not be possible. This foregrounds the value of informal, unseen spaces of interaction for the meaningful embedding of technologies.

Following Dixon-Woods et al. (2006), in this section we endeavour to systematise the relationships among the themes surfaced by each discipline into a synthesising argument. Our argument aims at providing a multidisciplinary framework addressing the dynamics of work-related change set in motion by the digitalisation of healthcare work. This will enable us, in turn, to propose conceptual entry points for studying dynamics of digitalisation of healthcare work in a way that tries to reconcile apparent tensions emerging from different bodies of literature.

Directionality and open-endedness

Across the bodies of literature we analysed, an apparent tension emerged between the open-ended negotiations of technologies in practice, and the fact that most of the configurations resulting from these negotiations are aligned in a direction of increased automation, datafication, rationalisation and transparency. Exemplified by the tension between the two conceptual 'souls' of the sociological literature (technology-in-practice and technology as steering), the coexistence of open-endedness and directionality also appears in STS and medical literature. STS acknowledges the steering power of materiality, while medicine laments the disconnect between technological promises and their failure to materialise in practice. Directionality, the property of innovations aimed at enacting a specific type of systemic change (Weber and Rohracher 2012), is present in all three metaphors. Slime mould will grow in the direction of areas that are (made) rich in resources. Theatrical scripts are embedded in specific dramaturgical traditions and steer the performance in a certain direction. River engineering interventions also centre (uni)directionality, especially when straightening the water flow. Nonetheless, the directionality inherent in all three metaphors does not exclude the possibility for situated open-endedness: slime moulds' tendrils explore patches of the surrounding environments in all directions; actors and directors decide how exactly to interpret or deviate from traditions in a specific performance; and, although rivers flow in one direction towards their outlet, they need to be given the space to meander, lest their water overflows.

Based on the three bodies of literature we considered, we propose here that technological scripts, invisible work and informal organisation can be singled out as sensitising concepts enabling scholars in these fields to productively study, in turn, the directionality of the digitalisation of healthcare work and its open-ended negotiation in practices. Technological scripts capture directionality in their materialising values and visions for the future of healthcare. The three literatures agree on this point: STS in its foregrounding how technologies' materiality make only some interconnections possible; sociology in its analysis of the steering power of technologies; and medicine in casting digital technologies as tools to streamline workflows. Scripts thus emerge as an entry point into the directionality of the digitalisation of healthcare work: they are material articulations of (allegedly) desirable future practices, roles, responsibilities (Akrich 1992). As such, they endeavour to give directionality to the digitalisation of healthcare work.

Scripts must, however, be negotiated in practice, and it is in this process of domestication, often accomplished through invisible work, that open-endedness comes in. Scripts interact in unpredictable ways with the practices and relationships that are already there. This suggests that "fitting with" practices is not so much a property of technologies, as the medical literature assumes, but a relational process of entangling materiality and work practices. This resonates with a well-known argument in studies of IT infrastructure, centring

> the ever-present tension between the general (standardised) and the local (situated), which people attempt to bridge through articulation or tinkering (the steps taken to get things done as work unfolds in real-time, despite material limitations, regulatory constraints, imperfect data, conflicting priorities, reluctant colleagues, etc.) (Greenhalgh et al. 2019: 2-3).

We argue that the tension between directionality and open-endedness is what this tension between the general and the situated looks like in the context of digitalised healthcare work. What Greenhalgh and colleagues describe as articulation and tinkering resonates with the different forms of invisible work and informal organisation described by the three bodies of literature we analysed. The focus on the open-ended work of situating directionality emerged from STS's attention to preserving strong interconnections, from sociology's engagement with invisible work, and from medicine's concern with making digital technologies fit with work practices. Informal interactions and invisible ways of coordinating healthcare work are crucial aspects of the open-ended process of situating directionality, and as such, combined with technological scripts, they provide an apt entry point into the tension between directionality and open-endedness.

Discussion and conclusion

In this review, we have analysed articles addressing the digitalisation of healthcare work from three different disciplinary perspectives. If reviews of the impact of digitalisation on work practices are not absent in sociology (e.g. Timmermans and Berg 2003), our CIS harnesses and updates this discussion, shedding light on how classic sociological themes such as invisible work are reconfigured in the context of new digital healthcare technologies. Our CIS also contributes to the STS literature, traditionally shy of middle-range theories (Beaulieu et al. 2007), by trying to move its empirical findings to a more abstract plane. Finally, we contribute to the medical literature by adding an interpretive layer to the systematic reviews usually produced in that field. Synthetising literature from STS, sociology and medicine has shown us how the directionality of innovation trajectories clashes and yet coexists with the open-endedness of situated changes in work-related practices. Based on our synthetising argument, we have proposed that finding dimensions to study both the directionality and open-endedness of the digitalisation of healthcare work is crucial. Technological scripts and different form of invisible work offer us entry points to study interactions between directionality and open-endedness.

Methodologically, we proposed a novel way of performing a CIS of multidisciplinary literature through metaphors. Metaphors enable to systematise the contributions of each body of literature, to articulate underpinning assumptions, and to explore the main insights each discipline provides. We argue that, in the context of CIS, metaphors can be deployed as synthetising constructs, allowing to transform entire disciplinary approaches into new constellations of themes, and to bring them together in a synthetising argument.

Our methodology presents several limitations. Firstly, we have presented bodies of literature as bounded and homogenous entities. However, the disciplinary boundaries we have traced are — as all disciplinary boundaries — porous and somewhat artificial. Several of the journals we looked at are intrinsically interdisciplinary, and several of the authors included habitually cross disciplinary boundaries. As discussed above, we have incorporated these considerations in our methodology. Moreover, since the contribution of our CIS lies, first, in systematising the most recurrent themes in each discipline and, second, in synthetising insights across fields, interdisciplinarity and fuzzy boundaries do not end up substantially impacting our findings.

Secondly, in our search strategy we relied on impact factors to identify the top journals for each discipline. This inclusion criteria builds on the assumption that top-tier journals are more likely to capture major debates in a field. Nonetheless, this search strategy might have led us to exclude relevant contributions, not lastly because it skewed our results towards English-language publications. Moreover, the digitalisation of healthcare work spans many more disciplines than the three we have analysed (e.g. HTA, implementation studies, management science, innovation studies, design studies). Although STS, sociology and medicine provided relevant insights, we might disregard some important dimensions of this dynamic. We hope to have inspired scholars to explore whether the tension between directionality and open-endedness also resonates in other disciplines.

Finally, as discussed above, metaphors are not transparent devices (Puschmann and Borgess 2014). Although our metaphors build upon and deconstruct concepts already present in each literature, none of them was able to completely cover the often quite disparate foci emerging from each corpus. A biological (though not evolutionary, cf. Wyatt 2004) metaphor convey STS's focus on ever-emerging networks, a social one the sociological attention for social relations, and a technical one the engineering stance inherent to medicine, giving substance to abstract conceptualisations (Lakoff and Johnson 1980; Wyatt 2014). However, they are to be read as syntheses, not exact reflections of each body of literature: each of them leaves something out. Our analytical strategy ensured that the selected metaphors align with the most prevalent conceptualisations in each discipline and are recognizable to scholars in the field.

To conclude, we formulate several recommendations for practitioners, designers and policymakers. We build on the general observation that even seemingly non-disruptive technologies (Hwang and Christensen 2008) call for fundamental readjustments in work practices. Even when technologies apparently "fit" with and streamline preexisting routines, they end up causing disruptions in the practices of healthcare staff (and of patients) — and disruptions require repair work. Our first recommendation thus aligns with Elish and Watkins's (2020) emphasis on the political and practical value of acknowledging, valuing and supporting the invisible work necessary for digital technologies to become embedded in practices (without striving to make such forms of work fully transparent).

Relatedly, our reading of the literature invites to reflect on the epistemic trade-offs associated with the digitalisation of healthcare work. Crucial knowledge exceeding quantitative data is likely to be lost through technological mediation. In implementing digital technologies, it is thus necessary to preserve some spaces of informal interprofessional and patient-provider interaction, where information can be exchanged without leaving (legally binding) traces.

Our final recommendation entails a rethinking of the idea of “fitting” innovation within preexisting work practices. Any new digital technology is likely to reconfigure the network in which it is embedded. Because of this fluidity, the concept of “fitting” appears misplaced. We urge developers, practitioners and managers alike to think of innovation in relational terms, as an ongoing process of experimentation and network-making, guided by a specific directionality but nonetheless intrinsically open-ended.

Chapter two
Uncertain visions[8]

[8] An earlier version of this chapter was published as: Carboni C, Wehrens R, van der Veen R and de Bont A (2022). Eye for an AI: More-than-seeing, fauxtomation, and the enactment of uncertain data in digital pathology. Social Studies of Science 53(5): 712-737.

Introduction

AI and data-driven technologies are increasingly being developed in the medical domain to assist professionals with complex diagnostic tasks (Keyes 2021; Shastry and Sanjay 2022). Though many of these technologies are still not in use in daily clinical practice, scholars within and beyond STS are investigating how they reconstitute organisations, professional relationships, knowledge, and responsibilities in the context of healthcare work (Bailey et al. 2020; Elish and Watkins 2020; Maiers 2017; cf. also chapter one).

Questions around knowledge and responsibility have emerged as particularly productive in a variety of academic debates concerned with decision-support systems. At the epistemic level, scholars have pushed against an exclusive reliance on “big,” quantitative data, casting qualitative, in-depth data as potentially better suited for some questions (e.g. boyd and Crawford 2012). Others have challenged AI and (big) data”s claims to objectivity by pointing to the omnipresence of interpretation, as well as to the fundamental issue of potentially “bad” or biased data (e.g. Mittelstadt and Floridi 2016). Not unrelatedly, questions of responsibility are often foregrounded in discussions of how professionals should incorporate algorithmic analyses in their decision-making, and who or what is to be held accountable when these decisions result in poor health outcomes (Gaube et al. 2018; Grote and Berens 2019). However, this focus on bias, outcomes, and legal frameworks risks reducing ethical conversations to technical and legalistic ones (Scheuerman et al. 2021). In this article, we attempt to speak to these questions of responsibility and knowing through a different approach. Building on an ethnographic study of the digitisation process undertaken by a pathology department in the Netherlands, we endeavour to unpack knowing and responsibility in digital and datafied healthcare by zooming in on the epistemic and ethical disruptions that AI (even in its absence) sets in motion in clinical practice.

Given the visual nature of its diagnostic processes, and recent developments in machine vision, pathology has become an obvious target for the application of AI-assisted diagnostics (Parwani 2019). Currently, however, digitisation is still only on the agenda of many clinical pathology departments. Pathology departments, anticipating a future in which AI-assisted diagnostics become the norm, embark on efforts to digitise the slides they produce through whole-slide imaging (WSI). We discuss the relationship between such digitisation (i.e. the transformation of analogue objects into digital ones) and other aspects of digitalisation, (i.e. broader dynamics of datafication and introduction of digital analytic; cf. Trittin-Ulbrich et al. 2020) in our overview of pathological methods below. Pathologists are thus increasingly requested to perform their diagnostic activity on digital, rather than glass, slides. Although one step removed from the actual application of AI software and thus considered fairly undisruptive (Ghaznavi et al. 2013), WSI nonetheless bears significant consequences for how pathologists make a diagnosis and for the (re)organisation of the department itself.

To shed lights on epistemic disruptions produced by WSI, as well as on their material-semiotic foundations and on their ethical ramifications, we turn to Barad's (2007) agential realist framework. Agential realism articulates how epistemic acts are generative of the reality they describe and, as such, inherently moral. By articulating the material-semiotic forces constituting the enactment of specific realities, agential realism, and specifically the concept of apparatus, lets us pin down how changes in organisational structures and epistemic practices are enabled by specific narratives and ambitions around digitisation—which simultaneously disqualify the enactment of other types of change (cf. chapter one; Giraud 2019). In this frame, knowing emerges as intrinsically moral—an act that requires taking responsibility for the realities one co-enacts.

Attending to the different constellations of humans, technologies, materialities, and narratives constituting digital and non-digital apparatuses, in what follows we begin to catch a glimpse of the roots of the ontological otherness of the objects they produce. We then turn to our analysis of pathologists' (digital) diagnostic practices. Building on pathologists' expressed concerns around working on digital epistemic objects, we identify three issues: sharpness, which we read as *sensorial uncertainty*, ontological depth, which we read as *intra-active uncertainty*, and fidelity, which we read as *fauxtomated uncertainty*. In our analysis, sensorial and intra-active uncertainty both stem from the ontological otherness of digital objects, materialised in their different affordances, and result in pathologists' inability to meaningfully engage with slides. Though not unrelated, fauxtomated uncertainty emerges from an agential cut enacting a nonhuman subject position, which complicates the question of taking responsibility for epistemic objects. Because of the moral nature of knowing, we argue, experiencing uncertainty hinders pathologists' use of digital slides in their diagnostic practices. To conclude, we elaborate on the relevance of agential realism for rethinking digitisation and "good" data in organisational settings.

Theoretical underpinnings

Particles of agential realism

Agential realism is a new-materialist, "posthumanist performative" framework, committed to "a genealogical investigation into the practices through which "humans" and "nonhumans" are delineated and differentially constituted" (Barad 2007: 32). Through performativity, Barad offers an alternative to representationalism, the epistemological position that casts epistemic subjects and objects as pre-existing, separate entities, the relationship between which is mediated by representations. Conversely, performative frameworks insist that practices of representation (i.e. knowledge) are ontologically performative: they participate in the creation of reality—epistemic objects and subjects, their properties and boundaries are enacted in the act of knowing. Building on Bohr's complementarity principle, Barad's main contention is that knowing and being are intertwined. Performativity thus spans, simultaneously, the realm of matter and meaning. In Barad's framework, the world is not made up of individual human or nonhuman entities, but of entangled agencies with multiple potentialities ("phenomena") that engage in and become determinate through specific intra-actions (a term Barard uses instead of "interactions", to avoid presuming the existence of separable entities). Intra-actions crystallise human-nonhuman distinctions, as well as the boundaries of phenomena's components, their properties, and the meanings associated with them. Intra-actions are respectively enabled or constrained by an apparatus—another Bohr-derived concept, here stretched to span boundary-making material-discursive practices. Apparatuses "produce, rather than merely describe, the subjects and objects in knowledge practices" (147).

In intra-actions, the boundaries between objects and subjects, humans and nonhumans, are defined through agential cuts. Just as human and nonhuman are not pre-determined substances, subject and object are not pre-determined positions that apply to individual entities with pre-defined boundaries. Barad, like Bohr, explains this through a thought experiment asking readers to picture a person holding a stick in a dark room. If the stick is held tight, it becomes an aid in navigating the room, and thus comes to occupy the subject position in the practice of knowing the dark room. If the person holds the stick loosely to sense its features, a different agential cut is enacted that relegates the stick to the object position. In agential realism, subject and object are mutually exclusive positions, enacted differentially in different practices (e.g. sensing the room vs sensing the stick). The boundaries of epistemic subjects do not coincide with the boundaries of what we assume to be "human": "human" itself is a category emerging within intra-actions. Objects and subjects do not pre-exist their interrelating, and different intra-actions, through different apparatuses, will enact ontologically different subjects and objects.

Objects and subjects thus cannot be divorced from the intra-actions that enact both. Ontology and epistemology thus implicate each other: there are no entities and no knowable characteristics outside of practices of knowing them. Epistemic practices, with the apparatuses sustaining them, are implicated into crystallising specific configurations of the world—at the expense of other possible ones. Agential cuts are for Barad a matter of ethics: Actors involved need take responsibility for the realities they do and do not enact.

Apparatus and epistemic cultures

The study of epistemic cultures (Knorr Cetina 1999) makes central epistemic practices, the "logics and arrangements through which knowledge comes into being and is circulated, approached and collectively recognised" (Nerland and Jensen 2014: 104). Knorr Cetina postulates such practices as pivoting on a specific "epistemic machinery" that constitutes, in a certain area, "how we know what we know" (1999: 1). This concept resonates with Barad's notion of apparatus. Like an apparatus, epistemic machinery shapes, enables, and constrains epistemic acts and the knowledge produced, creating the conditions under which "practitioners distinguish signal from noise, … or … decide which figure to trust when experimental outcomes are uncertain" (Knorr Cetina and Reichmann 2015: 18). The introduction of new technologies in an epistemic culture produces the rearrangement of epistemic machineries and their related cultures (Knorr Cetina and Amann 1990; Stevens et al. 2022). The concept of apparatus, however, emphasises how "machineries" are simultaneously material and semiotic, enabling us to probe the material-discursive practices shaping both the knowledge produced in a specific culture, and the objects and subjects within the culture itself. Moreover, Baradian agential cuts do not always result in epistemic subjects that neatly line up with the boundaries of the human—and sometimes are not occupied by "the human" at all (2007: 379).

As for epistemic objects, Knorr Cetina (2001) acknowledges they do not fully pre-exist the epistemic practices that investigate them: To some extent, their very ontology emerges within epistemic practices. Centring intra-actions, agential realism enables us to specify how epistemic objects unfold. Intra-actions are moments of embodied engagement, through which subjects, objects and their representation are enacted. We usually think of these moments as instances of sensory knowing. Recent scholarship on the more-than-human sensorium has argued that knowing agencies are not confined within the boundaries of what we traditionally think of as human. Especially in digitised settings, healthcare professionals engage in a sense-making that is at once embodied and technological (Harris 2021; Maslen 2017; Maslen and Harris 2021). However, increasing reliance on digital technologies in professionals' sense-making[9] is often found to lead to a progressive devaluing of sensory knowing (Beaulieu 2002; van Dijck 2011), especially when under the pressure of organisational policies (Campbell and Rankin 2017) or of personnel shortages (Maiers 2017).

In this article, we simultaneously embrace and problematise the crucial role that senses play in medical diagnostics. We postulate that professionals' expertise transcends the usually accepted boundaries of the human, by almost

[9] Many studies have also investigated laypeople's incorporation of data and digital technologies in their sense-making around their own health, often finding that increasing reliance on these technologies comes at the cost of sensorial and intuitive knowledge (e.g. Smith and Vonthethoff 2017).

seamlessly incorporating devices in their sensing and sense-making. By centring the specifics of *how* sensing with digital technologies is enacted in clinical pathology, we broaden the idea of professional sensing beyond the boundaries of perceiving, to span the enactment of diagnostic realities. By teasing out how pathologists and their microscopes, rather than just perceiving, actively enact the properties of the object they analyse, we argue that more-than-human sensing practices should not only be investigated in terms of their hybrid nature, but also of their enactment of specific realities—and thus of their intrinsically ethical nature.

Agential cuts, affordances, and automation

We propose thinking about epistemic objects' enacted properties in terms of affordances emerging from an agential cut (Hollin et al. 2017). Affordances are usually mobilised to describe what an artifact enables or encourages users to do (Bucher and Helmond 2018). However, since agential realism casts knowing and intervening as fundamentally entangled, here we expand the concept of affordances to include the possibilities for intra-acting with, and thus knowing, a specific (epistemic) object. In agential realist terms, objects' affordances are consistent with the material-semiotic practices we are referring to as apparatus; they support the epistemic practices expected by the users enacted as epistemic subjects within that very apparatus.

For pathology, a crucial difference in the apparatuses enacting digital versus glass slides concerns a degree of automation that often comes with digitisation. We argue that automation of this work qualifies as fauxtomation, that is, the practices around the "myth of human obsolescence" mobilised around technologies that (promise to) automate work (Taylor 2018).[10] Fauxtomation allows us to see that mainstream narratives around automation obliterate the amount of human labour that automation technologies still require. This "reinforces the perception that work has no value if it is unpaid and acclimates us to the idea that one day we won't be needed", leading to a

[10] The phenomenological ascendence of this point (e.g. Ihde 2016) is beautifully spoken to by Harris (2011) in her autoethnographic analysis of medical professionals' work of bodily attunement to the materialities of medical practice. In a similar vein, Schubert (2011) offers a compelling analysis of how anaesthesiologists embed monitoring systems in their practices of attending to patients' bodies in ORs.

devaluation of human, and especially menial, labour (Taylor 2018).

Since apparatuses are performative of epistemic subjects, fauxtomation is also involved in agential cuts that have concrete consequences for who (or what) is enacted as a subject. The idea that human labour can be made obsolete and automated can result in an agential cut that enacts nonhuman epistemic subjects. If the fauxtomation of diagnosis is, in clinical practice, still confined to the realm of future possibilities, it often underpins the enactment of digitised epistemic objects—which thus are often produced by technologies in the first place. Although humans are not absent from the organisational configurations stemming from these agential cuts (we are talking of a fake automation, after all), they occupy a position ancillary to the technology and its needs. This has significant ethical consequences since, as we have seen, epistemic subjects need to be accountable for the realities they create, but machines are famously unable to do that (Floridi and Sanders 2004). A representationalist critique of fauxtomation narratives might reject the agential cut those narratives draw between humans and supposedly automated work, but in this work we maintain an agential realist perspective by using these categories descriptively, rather than analytically. We trace how a self-contained machinic epistemic subject emerges through specific intra-actions that, being enabled by the fauxtomation narrative, are predicated on and enact a strong boundary between machines and the humans operating them. Thus, we call fauxtomation faux not because of the cuts it makes but because it imagines a particular cut to separate definite, a priori categories.

Fauxtomation provides us with an entry point into the dynamics of organisational change that underpin processes of digitisation, and into their onto-epistemic performativity. It reminds us that myths and narratives have tangible consequences for how we design and implement the relationship between human and machinic labour. We thus mobilise fauxtomation to examine the intertwinement of organisational and epistemic dynamics, and to problematise the boundary between automation of menial and of knowledge work. We argue that ways in which some tasks are considered more automatable than others are predicated on a devaluing of the sensory and contextual knowledge inherent to those tasks. This bears serious consequences for the reliability of the epistemic objects produced in the department, as well as for professionals' ability to take responsibility for the knowledge they produce.

The theoretical framework that we have brought together in this section suggests that the root of the ontological otherness of digital objects is to be found in changes in the workflow, materialities, rules, and people brought together to produce them, and simultaneously in the narratives and expectations steering these changes. Through this semiotic-material apparatus, specific agential cuts are enacted, resulting in specific distributions of affordances and responsibilities. This entails that, 1) ontological differences materialise in objects' affordances, which are to be considered not only possibilities for acting on and with an object, but also possibilities for knowing the object; 2) if epistemic questions are never separated from ontological ones, we need to bring into focus the organisational conditions (both material and semiotic) of production of the epistemic objects at hand; and 3) because of the ethical dimension inherent to epistemic acts, responsibility for the realities that newly-configured epistemic subjects enact can be foregrounded in the analysis of digitisation efforts and organisational change. In what follows, we attempt to apply these insights to gain a deeper understanding of the dynamics of change that digitisation initiates in different locales of a pathology department.

Data and methods

This article builds on an ethnographic study of a hospital pathology department in the Netherlands. One of us (Carboni) engaged with the department in various ways over the course of six months in 2021, as the department purchased new scanners and endeavoured to tighten its policies around digital diagnostics. Because of the relatively long duration of the data collection, we were able to analyse our data abductively (Tavory and Timmermans 2014), moving back and forth between the "field" and our analysis, and progressively refining our research design and strategy. The data collection started with 15 exploratory interviews with various members of the department, selected and recruited with the help of the department's head. The participants included in this first phase were thus judged by the department's head to be meaningfully involved in the digitisation process. They included pathologists who were both users and non-users of digital slides, as well as the manager overseeing the process, the lab managers, and lab technicians involved in the purchasing, set-up, and validation of the scanners. These interviews had a broad focus on participants' experience and expectations around digital pathology from different angles and were conducted via Microsoft Teams. Based on them, we were able to trace the current state of the digitisation process within the department, as well as to get some insights into issues currently experienced by pathologists.

After this exploratory stage, Carboni started attending the weekly meetings of the working group responsible for the digitisation process (consisting of a few clinical pathologists, two managers, a few lab technicians, the head secretary, the head of the tissue bank, and an image analysis specialist). During these meetings, relevant issues ranged from purchasing, to testing and validation, to pathologists' needs and workflow reorganisation, to machine learning in image analysis. Thanks to these observations, Carboni was introduced to the expectations and the most current issues experienced by the department in terms of digitalisation, and could observe the strategies discussed and implemented to deal with them. At the end of each meeting, Carboni had an informal hour-long conversation with the senior pathologist leading the working group, during which she deepened her understanding of the technical issues that had been treated and discussed more general points about pathology. Scientific literature, both suggested by participants and retrieved by the authors, was also consulted to learn more about the history of pathology and the developments around WSI and AI.

Carboni also conducted ethnographic observations in the department, with the aim of mapping the various stages of the workflow and to gain an in-depth understanding of the practices and tasks involved in each step. To do this, she adopted a "follow the tissue" methodology, which entailed tracing patients' samples (and their metadata) as they got (re)processed and progressively turned into slides. She observed front-desk secretaries as they registered tissues coming into the department, assistant pathologists, technicians and trainees as they prepared samples, lab technicians mounting samples on slides and checking their quality, and secretaries scanning the slides. She also joined the morning handover meetings during which interesting cases were presented by either pathologists or trainees, and attended one training session for ophthalmic pathology.

Ethnographic observations enabled us to trace the distributed epistemic and material practices at the heart of the department. During the working group's meetings and exploratory interviews, however, participants expressed concerns about whether pathologists could see "well" enough to make a diagnosis when analysing digital slides. To better understand what "seeing" entails in clinical pathology, Carboni conducted five object-elicitation interviews, using digital and analogue slides as concrete props to scaffold the conversation. These conversations circumvented a shortcoming of observation: Pathologists "see" quickly, mostly while hunched over a microscope—an unfavourable situation for an external observer. As often found in qualitative research, graphic or object elicitation in interviews support participants in articulating their lived experiences (Woodward 2016). In our case, discussing concrete cases with pathologists, probing them on which cues enabled them to make a diagnosis, and how they went about getting these cues, enabled us to reach a deeper understanding of the embodied practices underpinning slide analysis. Our interpretation was enthusiastically confirmed by pathologists themselves during the two presentations Carboni held at department-wide meetings. Drafts of this article were shared with the research participants for a member check. Participants provided a few technical corrections but did not request the removal of any quotes or fieldnotes from the analysis.

"Turning meat into information:" Pathology work as intra-actions

The profession of pathology, rooted in anatomical studies of disease processes through autopsies, has grown and differentiated into a complex and varied medical specialty, whose practitioners contribute to individual patient care, the detailed description of diseases, and even the ontology of disease itself (Van der Tweel and Taylor 2010). In this article, we focus on histopathology (henceforth, for simplicity's sake, "pathology"), the examination of cellular structures under a microscope, since it is the subsection of pathology on which digitisation efforts currently focus. Since the 1990s, pathology has entered an era of intense technological innovation through the gradual move to WSI (referred to here as digitisation), automated image analysis, and development of AI for assisting diagnostics (Ghaznavi et al. 2013; Pantanowitz et al. 2020). These forms of digitalisation promise increased quality of diagnosis, reduced work burden for pathologists, and a move towards personalised medicine.

Nonetheless, in practice, even digitisation (WSI) proves difficult. The department where we conducted our fieldwork had been engaging in this shift for more than ten years. Throughout our fieldwork, we witnessed the discomfort of pathologists and other actors, such as secretaries and lab technicians, with the changes digitisation forces into their practices. This section explores the department's workflow, with an eye on the intra-actions through which digital and non-digital slides are enacted. Crucially, albeit not being part of the diagnostic process in the strict sense, these intra-actions are both a testament to the distributed nature of diagnosis, and to the inextricability of matter and meaning. Knowledge is being applied and produced at every step of slide-making.

Much of a pathology department's work revolves around creating (good) slides. This is a sort of datafication of human bodies, as explained by a senior pathologist who described his work as "turning meat into information". In the hospital where we conducted our fieldwork, tissue samples of disparate sizes, from biopsies to entire organs, are collected from various wards and delivered to the department's front desk several times a day. Front desk secretaries register the case in the department's lab and information management system (LIMS). They connect the case file in the LIMS, which details the extent and nature of the sample, with the patient's information in the hospital's electronic health record, and assign to it a specific 2D barcode that is then printed out and stuck on each jar making up a case. Scanning this barcode automatically opens the case file in the LIMS.

Next, secretaries transfer tissue samples to the grossing room, where assistant pathologists, residents and lab technicians perform an examination of the samples and fix samples in formalin, making them hard and resistant to proteolysis. While small samples such as biopsies are directly embedded in blocks, larger samples must be dissected, put into a cassette, and embedded in paraffin, creating multiple blocks (Figure 1). This is a first kind of intra-action, a moment in which enacting an object and knowing it are inextricable. Pathologists and assistant pathologists describe how touching bones, cartilages and tumours in the specimen makes grossing an act of bodily engagement, steered but not determined by protocols, scalpels, pathology saws, and numbering practices. These forces also steer the selection of which parts of the specimens to turn into blocks, and which to reject. The process of inclusion and exclusion, of enacting only some properties as belonging to the epistemic object in question, is here extremely visible. Crucially, it is also painstakingly documented through pictures, reports in the LIMS and, in case of doubt, by involving more senior pathologists. Although no definitive cut is enacted (after sampling, large specimens are temporarily preserved in the grossing room in case pathologists request additional slides), the thorough documentation of this step enables accountability for the epistemic objects produced.

Figure 1: Small samples of tissues are embedded in paraffin and turned into a block. The different colors of the blocks flag different types of tissue. On trays, blocks are transferred from the grossing room to the lab, where technicians turn them into slides.

Blocks are taken to the lab, where technicians cut their content into very thin slices using rotary microtomes. They then mount these thin tissue slices onto small glass rectangles, to which they also apply labels with the case's barcode. These half-slides are fed to an automated staining machine. Once processed by the staining machine, tissues are visible and coverslipped: A glass slide has been created. The agential cut in this second intra-action enacts a subject position that is less clearly "human", since the process of staining is automated. Colour, which as we shall see plays a fundamental role in epistemic practices, is thus enacted in an automated intra-action. When sliced, tissue samples appear transparent: Staining is thus an intra-action through which some of their parts emerge as coloured, and thus legible. Through staining, tissue samples are enacted as containing specific cells, and structures with specific characteristics.

The automated staining machine does not document this intra-action. There are however systems in place to retrospectively check that the apparatus enacts consistent epistemic objects: In case of "special" stains, glasses are equipped with a control tissue, positioned next to the specimen. And once the staining machine has finished its operations, a lab technician collects the slides and checks them under the microscope, to examine the epistemic objects and certify the legibility of their enacted properties.

The digital workflow entails an additional step—digitising the slides. The person assigned to this task varies from place to place: if elsewhere specifically-trained biotechnicians are in charge of this process (Kusta, personal communication), in this department a secretary was reassigned from the secretariat to a dedicated "digital pathology" room. This "scanning" secretary waited approximately 10 minutes during our first observations before stating how much she disliked her job. Her main task entails feeding the glass slides to the four big pathology scanners crowding the room. Of these, two are new-generation, highly automated scanners, one is a specialty scanner, used exclusively to scan dermatopathology slides, and the last one is an older, drastically less automated one. Operating the latter requires a great amount of work from the secretary (Figure 2), as emerges from the fieldnotes below (all names below are pseudonyms):

> Jane fills up the scanner's rack with the slides, paying attention to make them match the order of the trays, so it's easier to put them back in the right trays when the scanning is completed. This older scanner is cumbersome, she explains, because it does not automatically focus on the tissue on a specific slide. So she needs to do this by hand: She enters case number associated with each slide, waits until the scanner takes a "picture" of each slide, then opens each of them and starts clicking on the tissues, tracing both its contours and the body itself. With each click, a small asterisk appears on the picture of the slide. On this one, a pathologist has already drawn circles around areas of tissue of particular

> interest. Jane makes sure to put a bit more asterisks in those areas. On the side, there is some additional tissue, but she tells me they're controls, and don't need to be scanned, so she does not put any asterisks on them. It is a painstaking task, it takes a lot of time, and Jane jokes that sometimes it makes her fall asleep. Once she is done, the scanning starts. On the screen, the scanned asterisks turn green. (fieldnotes)

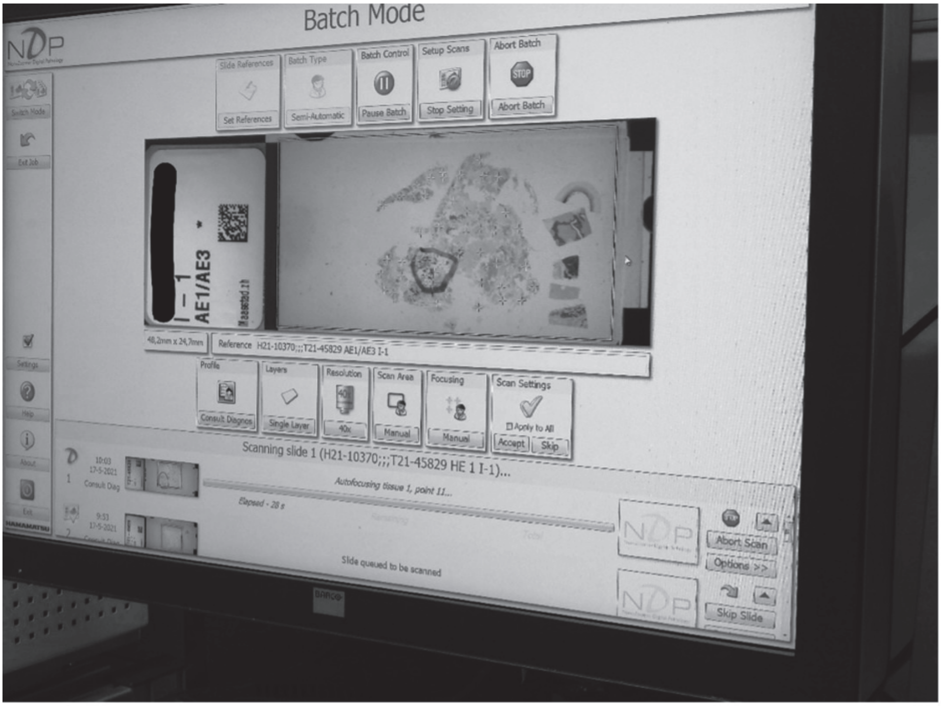


Figure 2: In the older scanner interface, scanner and secretary produce a digital slide: While the scanner implements an "autofocusing" function, the secretary further directs it selecting areas of particular interest; in this case, part of the tissue has been circled by a pathologist: the secretary places more focal points around that area.

Despite reportedly a supremely unexciting task, it is clear how much work the scanning secretary needs to perform in order for the scanner to produce a good digital image. The agential cut, in the case of non-automated scanners, enacts a hybrid subject position, in which the scanning secretary and the machine are not easily distinguishable in the digital slide-making. Jane can lend her eyes and her contextualised seeing to the scanner: She is familiar with pathologists' habit of circling parts of the specimen that require particular attention and knows that a circle drawn on a slide warrants more focal points, so that the image produced will be sharper in that area. She is also aware that control tissue is placed on the slide when it undergoes a specific staining process, is able to distinguish the control tissue from the specimen, and knows that there is no need to include it in the digital image. Her contextualised knowing and seeing enable the older scanner to produce good digital images.

Newer scanners, however, require less work on her part. They are part of an apparatus crucially informed by fauxtomation. New generation scanners (Figure 3) can identify the tissue's position on the slide (and thus which areas do not need to be scanned). They also automatically select focal points on the slide, based on an automated scanning protocol, so Jane only needs to load them with glass slides and start the process. After the scanning is completed, the scanner also rates the quality (i.e. sharpness and contrast) of the image, assigning to it a number up to 100. In this intra-action, a largely nonhuman subject position is enacted: the fauxtomated apparatus makes Jane's perceptions and tacit knowledge superfluous. This process of automation is *faux*-, however, because it still substantially relies on human labour. We observed that these scanners in practice run a lot less smoothly than expected: they reject some slides, and sometimes enact digital slides of insufficient quality. All of these cases call for Jane's intervention.

The digitising intra-action is thus the only one in the department in which humans are only marginally enacted as epistemic subjects and of which there is no documentation. Crucially, digitisation also entails a different apparatus, this time made up of scanners, their protocols, image recognition software and related algorithms, software for image quality control, rules for priority, inscription devices, secretaries supporting these operations, the need for sharpness, the fauxtomation myth, and the expectation of AI-driven diagnostics. As we show in the next section, this enacts ontologically different epistemic objects, with properties and affordances different from their analogue counterparts, which in turn enable (and disable) different epistemic practices.

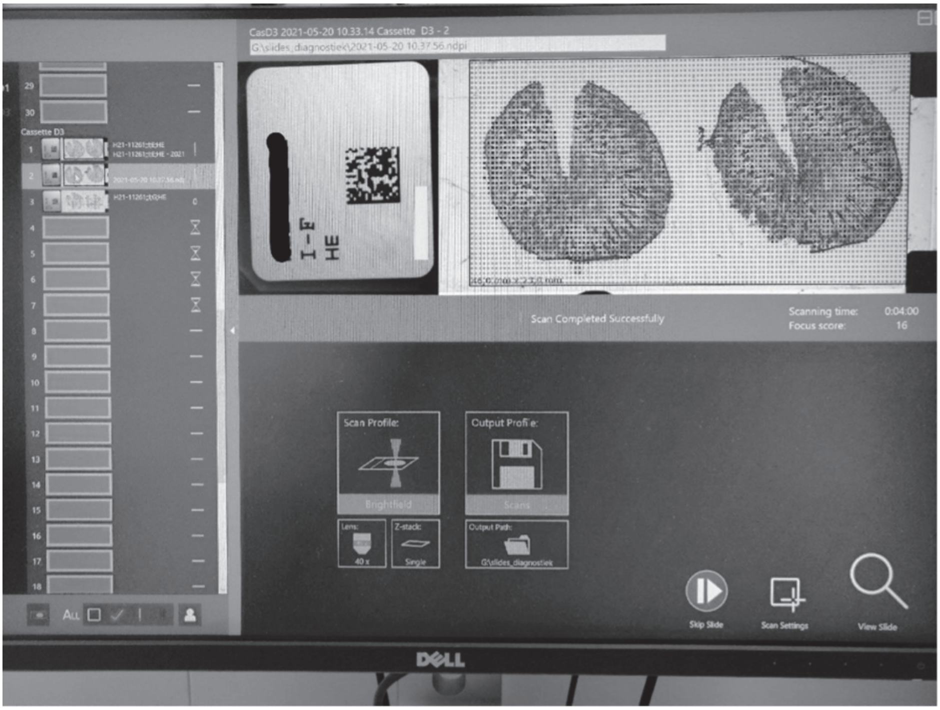


Figure 3: New-generation scanners are enacted, in faxutomated intra-actions, as firmly occupying the subject position. They autonomously select the boundaries of the tissue on the slide, and the areas on which more focal points (the blue and yellow crosses) are necessary. They also attribute a focus score to the resulting image, based on the contrast achieved in the scanning process. The secretary is only tasked with loading them, and restarting the scanning if the score is too low.

Diagnostic intra-actions

In this section we provide an agential-realist reading of pathologists' diagnostic practices. We analyse diagnosis as an instance of intra-acting. If conceptualising the lab's workflow as a collection of intra-actions helped us understand how epistemic objects end up having specific properties, here we focus on how epistemic objects are enacted as legible for a specific epistemic culture, and on how they are involved in diagnostic processes. Specifically, we attend to how digital and non-digital objects afford different types of intra-active engagement from different epistemic subjects.

This operation entails unearthing diagnostic practices in pathology that, at a first glance, mainly concern seeing: Pathology labs work to make tissues and their microscopic structures visible, and pathologists themselves appear to spend their days carefully looking at slides. In this section, however, we attend to what it takes to see in (digital) pathology. We juxtapose digital and non-digital epistemic objects and the intra-actions they afford, endeavouring to tease out the epistemic implications of the ontological difference between the two. The ensuing epistemic dynamics produce, respectively, sensorial, intra-active, and *fauxtomated* uncertainty.

Situated seeing: Sensorial uncertainty

Haraway (1989) famously reminded us that a view is always from somewhere. This very concretely applies to our case, insofar as specific epistemic cultures develop particular ways of seeing (Friedrich 2010), as well as epistemic objects that enable those ways of seeing. Apparatuses co-produce not only objects and subjects, but also objects that are legible within an epistemic culture and culturally specific practices of reading them. In pathology, traditional epistemic objects (i.e. glass slides) present affordances that can be meaningfully harnessed within pathologists' intra-actions. They lend themselves to be known in a way that makes sense within pathology's epistemic culture.

In their intra-actions, pathologists focus on histological structures (different types of cells, their subcomponents, and the architectures they make up), especially their qualitative aspects (Apfeldorfer et al. 2008). Color plays an important role, as Jody, a young pathologist specialising in pulmonary and dermatopathology, explained as we were looking at a glass slide through her microscope:

> [Malignant cells] look different. I'm sorry, I don't have metastases [on this slide]. But they are really pink. So, here everything is a little bit more purple, and they would be really pink. But I will show you the tumour. … they are enlarged, and they are quite pink.

Being the result of the staining process, colour makes for an interesting visual cue: It has been enacted as the result of a previous intra-action. However, it is not the only relevant visual cue. Altered shapes and cell architectures, histological architectures, and the relationships among different cells are all important signals of pathogenic changes as Harold, a senior pathologist specialising in gynaecological pathology, articulates:

> [Y]ou can see that [cervical cancer cells] have less cytoplasm, so they are less pinkish. They do have dark, sometimes irregular nuclei, here the nuclei are more open. It is darker and … you have a change in nuclear size and shape.

The department's non-digital apparatus thus enacts glass slides in which colour, shapes and nature of lines are meaningful characteristics; pathologists and microscopes are enacted as prospective users of these epistemic objects, and qualitative detail as relevant in the department's epistemic practices.[11]

Since epistemic objects and epistemic practices are co-enacted through an apparatus, it may come as no surprise that the digital apparatus constitutes a disruptive element in the department's diagnostic processes. Although they resemble each other, digital and glass slides are enacted through different apparatuses, and thus present different affordances. As a result, pathologists rarely perform their diagnostics on digital slides only. When in doubt, even pathologists who consider themselves enthusiastic adopters move back to the microscope to examine glass slides, as Laura, a young pathologist specialising in gastrointestinal pathology, did during our interview:

> This is suspicious. Look, this is a lined lumen and there is a tumour cell. I have to be sure. And then … I have to find this in this slide again. I still prefer to see the slide as well, in this case. … Just to check, actually. To be certain. But it's not really necessary, I mean, I can also—this would be sufficient to do it here.

Pathologists often brush off this widespread uncertainty as a matter of technical limitations, of insufficient image quality, bound to be overcome with technological developments. Here we attempt to take this uncertainty seriously, as rooted in the ontological difference between digital and non-digital objects, and in the lack of fit between digital objects' affordances and pathologists' intra-actions. Digital slides are enacted by an apparatus that also enacts AI as

[11] Quantitative factors are not absent from pathologists' analysis: Cells' size, mitotic figures, and the distance of tumour cells from resection margins are all relevant quantifiable. However, these quantifications are imprecise, performed by putting a ruler under the lens, or using an eyepiece reticle, or eyeballing (i.e. estimating the number of) mitotic figures. Arguably, these are cases of qualitative quantifications, pivoting on distinctions such as large vs. small, many vs. few, far vs. near.

a potential epistemic subject. Subsequently, their affordances encourage epistemic practices less tied to the largely qualitative ways of articulation intrinsic to pathologists' expert seeing. Pathologists articulate this shift when reflecting on what digital slides do not enable them to see, or what they discourage them from looking at. Consider, for instance, this quote by Hanna, a long-time user of digital sides, specialising in dermatopathology:

> [Looking at digital slides] is something that you have to get used to. Because I mean, if you look at a granulocyte [under the microscope], for example, the nucleus looks like a pair of glasses, and it has a bit of red around it; normally, under the microscope, you can really see the granules—and [on digital slides] you just see it as … red dots, but not as granules. And you just have to get used to that and ... adjust this in your head. … you still can recognise them, but it's not fully in the same manner. For mitosis, I don't have a solution. I think that I can now recognise granulocytes as easily on digital [slide] as under the microscope, with a good scanning. But the mitosis—you really miss it … And sometimes I still have to take the glass slide, because I'd have it anyway—and put it under the microscope.

Digital slides do not allow pathologists to examine the same amount of qualitative detail as glass slides would, especially when looking a cells' nuclei. As Hanna explains, pathologists can learn to recognise specific features, even though they look different on a screen—but other features are still difficult to spot, even after years of practice.

At the same time, digital images and associated viewers afford increased precision in the quantitative mode of analysing slides. The epistemic object comprising digital slides and imaging software lends itself to quantification much more than glass slides and microscopes, thanks to built-in quantification instruments, such as precise measuring devices. Besides "secondary advantages" (Harold), like increased traceability of slides, shareability with colleagues and students, and remote accessibility, digital pathology also enables pathologists to perform more quantitatively precise measurements, which can significantly influence the recommended choice of treatment. We can thus think of digital pathology as fostering epistemic practices that prioritise quantitative precision (how many things are there, what they measure) over qualitative precision (the examination of how things look).

Though they do not fit with pathologists' current epistemic practices, digital slides' affordances are central for AI, a prospective user envisioned central in the department, as John explains:

> We put a lot of money in machines and actually … your diagnosis doesn't improve at this point. And it's all to generate these slides images for future analysis, because you have to make the step to artificial intelligence. And you can only do it when you have a fully digitalised pathology department.

Automated image analysis, which at least in this department is still in the realm of future possibilities, would, in its infancy, mainly be concerned with quantification tasks, such as counting specific cells or mitotic figures (Ibrahim et al. 2021). As we learned during a demonstration given during one of the meetings of the department's digitalisation team, the level of qualitative detail machines and deep learning need in order recognise specific types of cells is limited to the presence of specific colours on the slide (e.g. the red of granulocytes Hanna mentions above). Further qualitative detail is, at least at this point, superfluous. By enacting AI as a potential prospective epistemic subject, the digital apparatus thus also enacts epistemic objects with affordances that do not always fit easily with pathologists' intra-actions, giving rise to sensorial uncertainty.

More than looking: Intra-active uncertainty

While the section above testifies to the relevance of slides' qualitative properties for pathologists' epistemic practices, casting these properties as something inherently present or absent in slides would lead us back to the representationalist fallacy that agential realism seeks to overcome. In this section, we harness intra-action to conceptualise glass slides and the visual information on them as something that is enacted through the epistemic practice of light microscopy. Crucially, we argue that qualitative detail is not inherent to glass slides, waiting to be surfaced: an untrained observer cannot just look at a slide and see meaningful detail, no one can examine a slide without a microscope, and it is hard to engage meaningfully with a glass slide when someone else is using the microscope. Rather, qualitative detail is enacted as a property of glass slides during intra-actions.

A close observation of pathologists' practices of engagement with microscopes and glass slides shows how this entails more than just zooming in and out. Seeing with a microscope is not just a matter of looking: It is about receiving "good" specimens, about selecting the right lenses, about using the microscope's light in the correct way, about inspecting the specimen without neglecting relevant areas. "Seeing", for pathologists, is therefore not just about seeing. One becomes a pathologist with and through a microscope.

These intra-actions are such a fundamental part of the practice of pathology that it can be hard for both pathologist and ethnographer to pin them down. There are, however, specific operations that can provide us with a glimpse of the intra-active nature of the practice of light microscopy. As Jody explained, pathologists know how to "play a little bit with the microscope", and this enables them to make specific characteristics of the slide visible, while backgrounding others:

> So, for example, if I'm looking at the mitotic figures … it's like this, [you get] this picture. But … if I turn this knob here, I can actually scroll through the nuclei. So now I'm getting this a little bit sharper, and if I [turn again], now, this gets a little [in] background and I'm getting this a little bit sharper. And if I'm not sure about the mitotic figures, I can use this technique to scroll through the nuclei to see how things are changing. And that's something I can't do with the computer.

We can think of the visibility of nuclei as a property of the slide that is not just surfaced but is enacted in this intra-action. In light microscopy, pathologists and microscopes are enacted as a joint subject, responsible for (and capable of) crystallising as relevant specific properties of the glass slides. In the quote above, the glass slide emerges as an epistemic object with visible nuclei. Similarly, we can start to understand how qualitative detail, rather than being inherent to glass slides, is enacted within intra-actions in which pathologists and their microscopes share a hybrid subject position. In the following quote, Jody explains how light also participates in intra-action, not just in terms of making colours visible, but in a less deterministic and more tentative way:

> And another thing I can't do with the computer is—if we have foreign material in the lung, then we use … polarised light … This is probably some dirt, but you see … it shining up. So, if I have crystal structures, for example—some pleural diseases go paired with crystal structures—I can make them visible only with this technique. And I can't … see them in the computer, because I can't use this technique breaking the light in a special manner. … if I have an infection with a granulomatosis aspect, I always want to look in polarised light, because I want to know if there's foreign material inside the granuloma … So if I only had the chance to look on the computer, I'd still have to ask the secretary to give me the

> [glass] slide, so that I can look at it. … I want to be sure, I don't want to miss anything.

The possibility of engaging in intra-actions in such an embodied way is not shared by digital slides. Digital slides' fixity and flatness bears crucial consequences for pathologists' epistemic practices, as John, another dermatopathologist and long-term user, describes:

> The biggest disadvantage is that sometimes the scanner cannot get the section into focus. … Because if …, for example, the tissue is a bit fragmented, then you have different layers of tissue. ... One piece is little bit thicker than the other piece, and then the scanner really has trouble getting all the pieces into focus. It focuses on the higher piece and not on the other pieces, for example. And under the microscope, you can migrate through the tissue. … some scanners can also scan multiple planes … you just make [an image with] different layers. And then you can also do this migration through the sections in a scan. But the biggest disadvantage is that you need … a lot of storage.

Digital slides' affordances are fixed upstream: The scanner, as an enacted epistemic subject, is responsible for making some properties visible and other invisible—once and for all. If pathologists' seeing rests on the possibility of participating in an intra-action that enacts qualitative detail, digital slides clash with their practices not so much because qualitative detail is "absent", but rather because they are ontologically "fixed" epistemic objects. That is, pathologists can *interact* with them, zooming in and zooming out, annotating and measuring them with specific software. But they cannot *intra*-act with them: they cannot harness their affordances to enact properties that were not yet defined.

This points us towards a fundamental characteristic of digital epistemic objects: their ontological closure and their drastically limited possibility for intra-action. The impossibility of intra-acting with digital slides, and of enacting relevant qualitative detail, makes it harder for pathologists to take responsibility for their digital diagnoses—especially in complex cases. As we shall see in the next section, this has to do with the fact that digitisation is supported by an apparatus that enacts a nonhuman subject position.

Digital artifacts: Fauxtomated uncertainty

As mentioned above, diagnostic intra-actions rest not only on pathologists' practices, but also on the correct mounting of glass slides. In imaging-based medical specialties (Cartwright 1995; van Dijck 2011), slides are a proxied way of peering into human bodies and the tissue on the slide is manipulated throughout several iterations. Pathologists' diagnostic practices rely on the (correct) performance of these manipulations: tissues need to be cut in the right direction, applied on the glass with as few wrinkles as possible, and stained in the correct way. In agential realist terms, this means that the apparatuses underpinning the different intra-actions in the lab's workflow need to perform consistently.

Deviations from apparatuses' consistent performativity results in artifacts, "artificial structure[s] or tissue alteration[s] on a prepared microscopic slide as a result of an extraneous factor" (Seoane et al. 2004). Artifacts stem from errors in the way the tissue sample is handled since even before the time of excision, all the way through staining and mounting on the slide. Agential realism allows us to consider artifacts as a way of flagging deviations from the way different apparatuses are expected to perform, and the objects they enact are expected to be legible. Epistemic practices also entail identifying and rejecting epistemic objects with artifacts. Doubting whether something is an artifact can make a slide unknowable, as emerged during our interview with Laura:

> Laura: Here, I have the impression that this is the vessel … and that there is a tumour in this vessel. But it's not easy, because there is something called retraction artifacts and then it also looks like a clear space or lumen, so I'm not really sure if this is really a vessel wall with a tumour in it or that it's just … a fixation retraction artifact, and that the shrinkage of the tissue makes that this looks like an empty space.
>
> CC: What is a retraction artifact?
>
> Laura: [It's] because of the fixation of the tissue—you need to fix the tissue to be able to make a slide ... And when [you fix it], the tissue shrinks, and sometimes this makes it tear a bit. And because this is dense stroma, then sometimes it can shrink or tear and then this looks like a space — but maybe it's not a real space. But then I can do some immunohistochemical stains to prove if this is a luminal vessel, so lined by endothelial cells, or if this is just stroma. So then if I am in

> doubt — is it really a vessel or is it just a shrinkage artifact? — then I can do an additional stain to help me with that.

Laura knows that tears in the tissue are a possibility, and what tears would look like on a slide (they could be confused with a luminal vessel). In addition, although it is not always obvious whether a property of a slide is an artifact or not, she is aware of the conditions under which slides can be enacted with a similar artifact (during the intra-action in the grossing room). Crucially, solving this uncertainty means going back to the block, turning it into new slides, and staining them differently. It requires a new, additional intra-action that enables crystallising different properties.

Glass slides are thus not trustworthy per se — but the intra-actions they undergo are well-known to pathologists, and thoroughly documented in the LIMS. Moreover, these intra-actions involve humans in different roles, both participating in them and checking their results. The quality control step, in particular, constitutes a formal evaluation of the correct processing of the slide. Although it does not always guarantee that the slides, even when enacted "correctly", afford transparent images of diseases, nor that slides get to the pathologists completely free of artifacts, it features actors that can be held accountable for the artifacts that pathologists spot. Pathologists thus not only know the workflow and the semiotic-material apparatus scaffolding slide-making, but they can access the LIMS and identify the assistant pathologists who have performed the quality control and approved that glass slide as suitable for diagnostics. Most of these conditions do not apply where digital slides are concerned.

The affordances of fauxtomatically produced digital epistemic objects thus affect the epistemic practices of professionals and their ability to take responsibility for the knowledge they produce in the diagnostic process. As per Barad, onto-epistemology is also a matter of ethics: It is necessary, as epistemic subjects, to take responsibility for the realities that are enacted and for the ones that are not. However, this also means that one can only fully take responsibility for knowledge when one is meaningfully involved in the intra-action enacting epistemic objects—or, at the very least, when one has sufficient insight into the intra-action and its onto-epistemic consequences. Since fauxtomated intra-actions enact a nonhuman subject, neither of these conditions applies to the creation of digital slides.

This can make digital artifacts particularly tricky business. Some of them, such as the absence of the specimen from the digital slide, are easy to identify for pathologists because they make slides unequivocally illegible, thus clashing enough with the department's epistemic culture and not generating any uncertainties or questions around epistemic responsibility. Other scanning artifacts, however, are not clear-cut and easy to spot. During our ethnographic observations, we had a chance to witness pathologists' unease with some of them.

> John takes the floor to warn the rest of the team about the risks that switching to digital diagnostics could entail. He explains that, if a biopsy is really small, sometimes the scanner will ignore it. Thus, if they stop distributing glass slides, and pathologists can't compare digital slides to them, they won't be able to tell if something is missing from a digital one. "Maybe there is a tumour, and you just don't see it", he explains. As Harry opens a random digital slide on the IMS, … Jesse exclaims: "This is such a good example!" We look more closely: the tissue sample on the slide is made up by two round shapes next to each other. However, one of these pieces of tissue has very sharp and straight edges—straight in a way in which human tissues don't grow. Harry zooms in and confirms that, apparently, a spot in the middle of the slide wasn't scanned. Jesse comments that here you can easily tell that some tissue is missing, but that's not always the case: "You don't always see that something's missing. That's scary." No one seems to be able to explain what happened, not even Harry or Robert, the IT experts. It's even more disconcerting for them because it turns out this slide was scanned in the 250 scanner, the special one for dermatopathology—which is supposed to produce very high-quality digital slides. (fieldnotes, eTeam meeting)

The idea of working fully digitally, and the risk of receiving incomplete digital slides without realising, is obviously troubling for pathologists. In cases in which the scanners fail to enact parts of the tissue as belonging to the digital epistemic object, pathologists simply cannot know whether epistemically relevant characteristics of the tissue have not been enacted. As we have seen in the previous section, digital epistemic objects radically restrict the possibilities for intra-action: the properties that have not been enacted by the scanner are simply lost.

Fauxtomated intra-actions thus turn out to be problematic on two levels. First, because of the fauxtomation narrative that constitutes the apparatus underpinning them, they enact a nonhuman epistemic subject, preventing the "human" from meaningfully intervening in the enactment and quality control of digital epistemic objects. Neither pathologists nor secretaries have a say in which and how properties of these objects are to be enacted, nor are they tasked to check whether the results of this performativity meet the epistemic requirements of the department. Fauxtomation thus obliterates the possibility of dealing with the uncertainty stemming from automation and its inherent decision-making.

Second, digitisation, and its attendant fauxtomated intra-action, embraces a representationalist logic: It assumes relevant information to be intrinsic to epistemic objects, and casts the intra-actions that make up the diagnostic process as a simple matter of zooming in and out. As such, it disregards the performativity of both the apparatus crystallising specific properties (and not others) in digital epistemic objects, and the performative, embodied engagement with epistemic objects that is the foundation of intra-actions. We can thus think of pathologists' discomfort with performing diagnoses on digital slides as stemming from the fact that, while excluded from the intra-actions enacting slides' properties, they are still required to take responsibility for the knowledge they produce based on them. The concluding section elaborate on the consequences of this insight for broader debates on knowing and automation.

Conclusion: Towards a new materialist agenda for digitisation

In this article, we have examined digital and non-digital epistemic practices in pathology through an agential realist lens. We considered both digital and glass slides and diagnosis as the result of intra-actions in which knowledge, materialities and narratives are enacted. We thus shed light on ontological, epistemic, and ethical issues that complicate digitisation in pathology. Digital and non-digital slides resemble each other but they afford different epistemic practices, since their respective apparatuses incorporate different narratives and make space for differently assembled epistemic subjects. In this section, we elaborate on how agential realism helps us to articulate a post-representationalist, new-materialist agenda for digitisation. In concluding, we endeavour to tease out possible ways to materialise the political potential of the Baradian framework in this context.

Agential realism has enabled us to trace the epistemic disruptions produced by digitisation back to digitisation's own representationalist assumptions. As we have argued, representationalism plagues digitisation efforts whenever epistemic objects are considered self-contained entities, that is, when their properties are presumed to pre-exist their analysis. In our case, the attempt to replace glass slides with digital slides builds on a representationalist logic. It assumes that the relevant data is *already* present in epistemic objects, ready to be perceived, and that it can be captured satisfactorily without further intervening. This representationalist fallacy has implications for the work of both pathologists and secretaries: for pathologists, it affects the quality of care they can provide; for secretaries, it opens up the possibility of de-professionalisation through partial automation. We will articulate these points in turn.

Representationalism casts glass and digital slides as interchangeable—as long as the representation is of sufficient quality. The whole idea of digitisation rests fundamentally on this assumption of ontological sameness—or, at least, comparability. At the heart of this is the assumption that both analysing and reproducing epistemic objects is something that can be done leaving their (presumed) original essence intact. As we have seen, the idea that scanning enables capturing without intervening fails to consider how the capture itself entails the intervention (and latent decision-making) of both scanners and (to varying degree) their secretary. Conversely, here we have teased out the more-than-human decision-making inherent to the scanning process, for instance in the selection of focal points.

Representationalist assumptions ignore not only the epistemic import of (digital) object making, but the material performativity of knowledge making (i.e. diagnosing). We have learned that not all relevant properties have necessarily been enacted in glass slides once they reach pathologists' desks. Whereas tissues have been enacted as being of a specific colour, additional properties, such as the presence or absence of crystal structures, have not been enacted. This taught us how diagnosing entails more than looking, and relevant information crystallises in the slide through the embodied intra-action involving pathologists and their microscopes.

If the task of pathology is to "turn meat into information," this process of turning is performed in more-than-human, embodied, and organisationally situated epistemic practices. Not considering the epistemic import of these practices increases the degree of uncertainty pathologists experience in the diagnostic process (e.g. has all the relevant tissue been scanned?), and introduces a potential threat to the quality of care provided. The potential false negatives rooted in intra-actions that are either hidden (e.g. the scanner deciding what the specimen is) or missed (e.g. the one involving microscope, pathologist, slide, and potential crystal structures) represent possible sources of error in the diagnostic process. As such, pathologists' uncertainty and distrust of digital slides proves a crucial, yet by itself insufficient, protective mechanism. Requesting glass slides or additional stains enables them to multiply the number of intra-actions a tissue undergoes. Were this distrust to dwindle, as managers hope, once pathologists become more used to working digitally, the quality of care itself could come under threat—especially when complex cases are concerned.

At the organisational level, representationalism enables specific implementation choices. Casting (high quality) representations as innocent stand-ins for analogue epistemic objects leads to two related erasures. First, it obliterates the degree of domain knowledge necessary for enacting good-quality digital objects. Second, it disentangles the knowing and doing that enable digital slides to come into existence. This disentangling serves the myth of fauxtomation, by enabling the obliteration of situated forms of knowing crucial even for allegedly menial tasks, as shown in our analysis of secretaries' changing work. If the boundaries between humans and nonhumans are not predefined, but emerging within specific practices, the concept of fauxtomation enables us to probe the political potential of agential realism and of its move away from representationalism. It is to this point we now turn.

Social science literature on digitisation and automation has long engaged with the question of whether the introduction of increasingly "smart" technologies results in the de-skilling and de-professionalisation of the workers operating or engaging with them (Findlay et al. 2017; Petrakaki et al. 2012). Recently, Delfanti and Frey (2021) have produced a compelling argument about the continued cultural acceptability of de-skilling in the context of automation. Our analysis resonates with much of this literature in finding that the work of digitising is, for a great part, experienced as repetitive and isolating by the ones doing it (cf. Barrett et al. 2012). However, as we have argued, there is no intrinsic reason for fauxtomation to lead to the de-skilling of these workers. Although new-generation scanners' functionalities and interfaces are conducive to a diminished involvement on the part of the secretaries, excluding them from the enactment of digital slides is a choice. Not making them responsible for the quality control of the digital slides, relinquishing that role to the scanner itself, is, as we argue, a move enabled by the fauxtomation myth. The directionality of the digitalisation of (healthcare) work (Campbell and Rankin 2017; cf. also chapter one), enabled by semiotic forces such as fauxtomation, is not always resisted at the level of local implementation policies. As we have shown, not only does compliance with the myth of fauxtomation make the secretaries' work less meaningful, but it also produces frustration and uncertainty for pathologists. In times of soaring burn-out cases amongst healthcare workers, this might be something to consider more carefully.

Agential realism enables us not only to see the representationalism at the heart of digitisation efforts, but also to envision alternatives to it. Starting from intra-action allows us to ask different questions of and for digitisation efforts, and to reframe issues of data quality. Embracing the performative nature of diagnosis, and the enacted nature of data, we can see how technical questions (such as the ones around sharpness) fall dramatically short of guaranteeing "good" data. In the case we examined here, the exclusive focus on sharpness fails to give due consideration to the embodied diagnostic practices underpinning pathologists' certainty—practices connected to and enabled by the situated knowing of other human and nonhuman actors earlier in the workflow. What constitutes "good" data, thus, becomes a question to be considered in the context of specific, interactional epistemic practices and organisational structures.

Conversely, uncritically foregrounding the promises of AI-assisted diagnostics and automation to boost workers' enthusiasm around innovation (a strategy often suggested in the change management literature, e.g. Chiu 2018) risks resulting in apparatuses that devalue pre-existing epistemic practices, generating either frustration or downright worse-quality jobs. Based on our analysis, taking seriously the different knowledges that make up epistemic cultures, and investing in supporting them with data (rather than trying to replace them; see Pedersen and Bossen 2021), would arguably be conducive to improved digitisation efforts. The challenge organisations face is thus both how to fruitfully mobilise undervalued domain knowledge, and how to ensure humans agencies are meaningfully included in the enactment of knowledge and materialities.

Similar points have been made in STS and Critical Data Studies (CDS), which have questioned forms of representationalism underpinning digitised knowing (boyd and Crawford 2012; Stevens et al. 2018). As we have shown, a new-materialist approach contributes to these post-representationalist sensitivities. First, agential realism, and the ethnographic approach it suggests, enables us to examine data enactment at the organisational level with great granularity, tracing less visible moments of knowledge production beyond the boundaries of the diagnostic process narrowly conceived. This includes re-examining instances of more-than-human intra-action as moments of knowledge production—though they might fall outside of the scope of what is usually considered epistemic subjects and practices. Second, and related, agential realism sheds light on how data are enacted in more than-human, organisationally situated practices. Opening up intra-actions as moments of knowledge (and data) production, and as crucial for the diagnostic process, allows us to see the synergies between seemingly disconnected issues such as the automation of menial work, epistemic uncertainty and resistance, and quality of care. Finally, agential realism encourages us to centre responsibility as an empirical question in thinking about digitisation efforts' epistemic disruptions. As we have seen, responsibility for data enactment is a relevant focus not only during the diagnostic process, but also at the point of digitisation itself. What sort of agential cuts specific implementation policies enable, what sort of subjects they enact, and how those subjects can be held accountable for the realities they co-enact emerge as fundamental questions for researchers and managers alike.

Chapter three
(Dis-)attuning the workforce[12]

[12] Under review as: Carboni C, Wehrens R, van der Veen R, and de Bont A. From attention to attunement: Unpacking data-driven efficiency and more-than-human care provision in the work of ICU nurses. *Science, Technology & Human Values.*

Introduction

Workforce shortages are one of the most prominent aspects of the state of crisis affecting healthcare systems today (WHO 2022). During the Covid-19 pandemic, intensive care units (ICUs) in particular were shown to be dramatically affected by shortages, especially in nursing staff. Whereas protocols mandate a 1:1 or 1:2 nurse to patient ratio in ICUs, at the height of Covid-19 nurses were often asked to attend to up to six patients per shift, to "create more beds" (Ford 2021). The experience of providing care in such conditions resulted in skyrocketing rates of PTSD and burnout amongst ICU nurses (Bae 2021).

Since intervening at the level of staffing is not considered feasible in current policy discourse (Dowling 2019), the scarcity of healthcare personnel has become a trope in framing the necessity of investing in artificial intelligence (henceforth, AI) tools (cf. Topol Review 2019; Deloitte 2023). As the argument goes, human resources are insufficient for the volume and complexity of current and future care demand, thus necessitating the introduction of technologies. AI is expected to automate some clinical tasks, freeing up clinicians' time and enabling them to provide efficient and high-quality care in suboptimal conditions (Lorkowski and Jugowicz 2020; Nagele and Thamm 2022).

STS and medical sociological literature suggest caution in equating digital and automation technologies with reduced need for labour. Although the implications of technologies for labour should always be an open empirical question (Timmermans and Berg 2003), empirical studies of digitalisation have often pointed out how the introduction of new technologies often results, for instance, in the proliferation of extra tasks that are often concealed in mainstream discourse (Oudshoorn 2008; Suchman 2007) and delegated to workers at the bottom of the professional hierarchy (Burri 2008; Maslen 2017; Mort et al. 2003; Petrakaki, Klecun and Cornford 2016). Moreover, new technologies are known to negatively affect the quality of work for lower-ranking professionals (Barrett et al. 2012; cf. also chapter two).

This paper endeavours to bring earlier STS reflections on the digitalisation of healthcare work to bear specifically on the algorithmisation of the (clinical) workplace (Jarrahi et al. 2021). Inspired by the work of Louise Amoore (e.g.

2023), we foreground the implications of attempts at automating away issues that are inherently political, such as workforce shortages. Algorithms restricts viable approaches to such issues: ”[w]hen algorithmic systems reduce the intractability and pluridimensionality of politics to a machine learning model, they foreclose the potentiality of other claims and alternative projects that could be built” (Amoore 2023 22). In our context, this means asking how the lived reality of care provision on ICUs is framed and, progressively, remade in the attempt to intervene on personnel shortages through data and machine learning.

We empirically explore these dynamics as they materialise in an innovation project that we followed ethnographically. The project aimed at intervening on ICU nurses’ shortages by making their work more efficient through the mobilisation of data analytics aimed at enabling individual nurses to attend to more patients. Interrogating this aim through our analytical sensitivity towards earlier STS work on disattended promises in combination with Amoore’s emphasis on recognising pluridimensionality and alternatives in algorithmic systems, this paper seeks to interrogate this “attending.” What kind of objects of professional and organisational intervention are surfaced through data-intensive forms of care provision? What do they leave in the background, or attempt to obliterate altogether? We keep these questions centrals as we trace the divergences between ICU nurses’ daily care practices and the way algorithmic management tools (Jarrahi et al. 2021) proposed in the project reimagined these very practices.

We find nurses’ attention to emerge as the object that algorithmic management intervenes upon to attempt to enable nurses to care for more patients. Attention is a concept with a long history in Western modernity, which we sketch in the first part of the paper. In analysing the AI technology proposed in the project we examine here, we argue that the project’s framing imports a notion of attention as a passive filtering capacity in information- and stimuli-rich environments. Simultaneously, it turns attention from a scarce commodity into a scarce resource (Liboiron 2021). In the second part of the paper, we propose an alternative, empirically-grounded conceptualisation of data and care practices in information-rich and technology-dense environments. Weaving together literature from STS, cultural studies and phenomenological anthropology, we assemble a concept of attunement as an embodied, affectively-laden disposition towards a situated environment. Attunement allows for staying open to a multitude of rhythms and emerging patterns, rather than closing down on an exclusive focus. Attunement enables us to show how ICU nurses' care practices overflow the narrow conceptualisations of attention that have been dominant in the last century (Crary 2001; Halpern 2014). In closing, we reflect on how moving from attention to attunement might not only do justice to the complexity of nurses' care and data practices, but also provide more empirically solid foundations for policy-making on healthcare workforce shortages. Since our discussion of both attention and attunement are situated at the intersection of extant literature and our own empirical material, we split our theoretical section thematically. We begin with an exposition of our case and methods.

Methods

This paper builds on four months of intensive ethnographic fieldwork conducted by Carboni in the adult ICU department of a large Dutch academic hospital. During these four months, she was exposed, simultaneously, to the largest amount of continuously produced data and sophisticated technologies, and to the largest number of alarms, unsettling bodily fluids and smells, and both temporarily and permanently unconscious bodies that she had ever witnessed. Donning a white nurse scrub and shadowing nurses throughout their eight-hour shift, she quickly realised how physically and emotionally taxing (field)work in an ICU is. With time, however, she progressively got used to being cold and tired in the ICU corridor, to the strange feeling in her stomach when patients died, and to trying to mask how dizzy stepping inside the room of a neurotrauma patient made her feel. Even her extremely watered-down, responsibility-free experience in the ICU thus proved analytically generative, sensitising her to the experience of care provision in an acute care setting, and to the experience of learning an embodied disposition to such a complex environment.

During her 150 hours of observation, she took extensive fieldnotes, and had long informal conversations with nurses (around 10 in each shift) about their experiences and their strategies for caring for and interpreting both patients' bodies and technologies. She observed nurses' routine checks on patients (described at length below), consultations with intensivists, patient transfers to other departments (discharges to general wards, physiotherapy, or radiology for CT scans). She also joined nurses during breaks and was able to observe their informal discussions.

Carboni had been allowed in the ICU as part of a project aimed at bringing about a “future-proof” ICU. As she soon found out, this implied chiefly making nurses’ work more “data-driven.” Since the Covid-19 pandemic, local nurses complained about feeling exploited by the hospital management. The project thus had to navigate a complex political environment. On the one hand, according to intensivists, the pandemic had created momentum for adopting AI technologies. On the other hand, nurses’ malcontent (not to mention soaring rates of burnout and long Covid amongst them), meant that they had to be “included” in the project. To do this, a few “inspiration sessions” were organised, during which nurses were informed about AI technologies, and asked to write down “points for improvement” for a future-proof ICU. Carboni observed six out of the seven sessions organised, and helped collecting nurses’ input by transcribing the post-its they filled in. These post-its were not discussed during these sessions. Carboni took extensive fieldnotes of the nurses’ reaction to the information presented to them, and of the issues emerging in these brief sessions.

Carboni was asked to contribute to the project through a visualisation of the nurses’ workflow based on her fieldwork. The aim of the visualisation, as emerged during talks with nurse coordinators and managers, was to identify parts of the workflow that could be made “more efficient” through an increased (re)use of data. To gather material for this visualisation, Carboni thus focused, in her observations, on how nurses produced, used, and ensured the reliability of the data mobilised in ICU patients’ care. One of the intensivists involved in the innovation project at one point approached Carboni asking whether her visualisation could be used to incorporate more “learning moments” in the workflow, so that nurses would not be seen sitting around and “drinking coffee” as much. The ongoing demand for efficiency and prominent disapproval of downtime that populated the discourse around the “future-proof ICU” clashed with the exhaustion she observed daily in the nurses she shadowed. This clash also emerged in her fieldnotes, coded inductively in Atlas.ti. Resulting codes centred on issues such as data production, data error, sensing, tinkering, time, efficiency, information flow, and decision-making.

In her visualisation, Carboni attempted to show the complexity and value of nurses' practices, rather than pinpointing moments to be optimised. She presented and discussed this visualisation in two private meetings held, respectively, with the head and the director of the ICU, and with two intensivists, one manager, and one nurse. During these discussions, the local personnel responded positively to her analysis of nurses' data and care practices as in tension with data-driven efficiency. However, as far as we are able to tell, these discussions and the visualisation did not steer the direction of the project's development.

Attention and AI

A brief genealogy of attention

Historical analyses of perception and emerging technologies have identified the historically- and culturally-contingent character of ways we "intently listen to, look at, or concentrate on anything" (Crary 2001, 1; cf. also Halpern 2014). Jonathan Crary has identified attention, in its current, culturally-dominant meaning of "a process of selection" (24), as a formation contextual to Western modernity. More specifically, "new technological forms of spectacle, display, projection, attraction, and recording" (2) emerging in the late nineteenth century have contributed to crystallising techniques and imperatives of "paying attention" as a "disengagement from a broader field of attraction, whether visual or auditory, for the sake of isolating or focusing on a reduced number of stimuli" (Crary 2001, 1). Currently taken for granted notions of attention as the capacity to focus on a single task excluding the noise of other stimuli need thus to be considered as historically-and culturally-situated, and in conjunction with technologies that have influenced ideas and techniques around perception.

Albeit emerging in the nineteenth century as part of discourses of modernity that emphasised the overwhelming number of stimuli people were exposed to in modern metropolises, this conceptualisation of attention becomes particularly prominent in the 1970s, in discussions around the "attention economy." Economist and psychologist Herbert Simon (1971) first described the attention economy as inverting the terms of the information economy (Davenport and Beck 2001; Goldhaber 1997): with new media, information ceases to be a scarce commodity, and is replaced by attention, defined as "whatever it is that information consumes" (Simon 1971). Discussions of the attention economy have since then seeped into the public discourse, usually and centring on the relationship between attention, advertisement, and consumption in media ecosystems (Crogan and Kinsley 2012).

These theories operate three baseline conceptual moves. First, they turn attention into a scarce, and thus quantifiable, commodity (Zulli 2018). Second, they conceive of attention as something that is eminently cognitive, can be either on or off, without any degree to it, and operates as a passive mechanism of selection of one focus over other stimuli (Bucher 2012; Crogan and Kinsley 2012). Finally, they connect attention and technologies: digital technologies are identified as both responsible for bringing about information overload, and as potential solutions that would "only provide users with the information that they need to know" (Bucher 2012).

This connection with technologies, which Crary (2001) already identifies, becomes particularly interesting in the context of AI. As Orit Halpern's work shows, AI has been modelled on models casting attention as filter (2014). Since early discussions in cybernetics, the reflection on AI has been heavily concerned with the issue of perception and information storage in contexts characterised by a multitude of stimuli. AI's algorithms function by filtering out what they consider "of interest" in an environment through "feature extraction, reduction, and condensation" (Amoore 2020, 16), and discarding "much of the material to which [they have] been exposed" (17). Clearly, AI's working bear surprising similarities with descriptions of attention as a filtering mechanism, disengaging from unnecessary noise. Given the historical and conceptual affinities between notions of attention and the workings of many AI technologies suggest investigating how attention is being remade with the ingression of these technologies in contemporary clinical practice. It is to this question that we turn below, by analysing plans and technologies aimed at making ICUs "future-proof" through AI.

Attention as a resource: On the algorithmic management of scarcity

As detailed above, one leg of our fieldwork entailed attending six "inspiration sessions" that were meant to kickstart the innovation project we were studying. Organised by Jim, the intensivist in charge of the project, with the aim of "involving nurses from the start of the innovation trajectory," these 30-minutes sessions were held weekly for almost two months. At the beginning of each session, nurses were invited to think about aspects of their work that would be different in an ideal "future-proof ICU." This was followed by a 20-minute presentation by Jim on AI applications in ICUs. Unsurprisingly, especially after Covid-19 had thrown into sharp relief the issue of workforce shortage, when asked what they would be different in the ICU of the future, nurses tended to point out the necessity to hire more people:

> Jim asks the nurses to write down three areas they would like to see improved in a "future-proof ICU." … Someone asks if they can also say that they just need more personnel. Jim doesn't seem enthusiastic, but says that "that's also a possibility." … [In a later session] Jim mentions laughing that nurses in the previous sessions have come up with points about parking spots and more personnel. Regardless, the point about more personnel also comes back in the post-its nurses hand over to me at the end of the session. (fieldnotes, inspiration sessions)

Although the concerns raised by nurses were not explored in depth in the duration of these sessions, similar exchanges speak to how personnel scarcity is central to nurses' current work experiences and to the way they relate to their future (figure 1). Jim's dismissive response is a prime example of current tendencies towards the side-lining of personnel scarcity as an unsolvable issue. What is more, this very side-lining, for Jim, opens up a possibility to intervene through technological means, thus justifying the project itself — as well as the investments it entails.

While this was not explicitly communicated at the start of the project, it soon emerged that the steering team (constituted of intensivists and managers, as well as one nurse) had already settled on a technology to be introduced: a dashboard that had already been piloted in another Dutch ICU. Developed with the support of a big tech corporation and a large data analytics company, this dashboard uses data analytics to visualise each patient's "care needs" at a given moment. As one of the department's intensivists explained in an interview to a Dutch magazine (not referenced here for purposes of anonymity), the dashboard would mobilise patients' health data in real time, producing an analysis of each patient's current state. More stable patients would be displayed as green, the ones "in need of more attention as orange," and the ones in a critical situation as red. Thus, the dashboard, visible in each ward's corridor, would allow nurses to allocate their time to the tasks, and the patients, that needed it the most, even in other units. In his interview, the intensivist explained that being "green" does not entail that a patient needs "no attention," but rather that nurses can "focus on the human dimension" (own translation). Although our ethnography had to stop before the dashboard's implementation, this very plan begs questions around the norms, assumptions and expectations scripted into its very (imagined) workings.

Figure 1: A magnet on the wall in the general ICU corridor, reading "We need a good labour agreement for university hospitals." Similar calls, sometimes sponsored by Dutch unions, abounded in the ICUs, and in the hospital in general.

Focusing on the assumptions about users' practices and responsibilities inscribed the dashboard, we can begin to unpack the mechanics through which the dashboard promises to intervene in the ICUs' care practices. The dashboard promises an ongoing analysis of real-time data. On the basis of this analysis, it assigns patients to categories of (in)stability. In other words, the dashboard sifts through the myriad of health data points continuously produced for each patient, combines them, and filters out the patients who require nurses' immediate attention. This means, first, that the analytics built into the dashboard function, in themselves, as a filter of sorts, reminiscent of modern notions of attention (Halpern 2014). Second, the filtering that the dashboard operates is mobilised for algorithmic management purposes: based on these analytics, nurses are allegedly made able to optimise their task- and time allocation. Thus, it is not simply the case that the dashboard functions following a model of attention as a filter; it also intervenes on nurses' attention, making it an object of algorithmic management.

Working with a modernist notion of attention allows the dashboard to promise making nurses' attention more efficient, thus qualifying it as a potential solution to workforce shortages. As we have seen above, attention is described in this discourse as a passive capacity, something eminently cognitive, and that functions by singling out one focus from a multitude of stimuli. Based on this, the dashboard can assume that attention can be easily refocused from one task (and from one patient) to another. Provided a worthy focus is identified, nurses should be able to disengage from other stimuli, and from previous task, to attend to the ones data analytics flag as urgent. This assumes that nurses' care practices can be neatly divided into tasks that can be easily exchanged among different nurses. It also seems to propose that nurses have no affective investment in the care of patients assigned to them in a shift, and that they can easily disengage from their care to turn to more pressing matters. Finally, it also appears to reduce nurses' work to attending to (unstable) patients: if they are to utilise their attention efficiently, nurses are to move quickly from one urgent care task to the next. Indeed, chopping up ICU patients, as it were, into urgent tasks, appears as the only way in which nurses could be able to attend to more patients than allowed by current care safety protocols.

This cursory analysis bears implications for thinking through the continuities and discontinuities of notions of attention as mobilised in the dashboard's design. Although attention's mechanics as a filtering device, as it is theorised in the dominant discourse that underpins the attention economy, stay relatively unchanged, there is a subtle shift in the ontology and the affordances of attention — that is, what attention is thought to be and to be able to achieve. In the attention economy discourse attention figures as a commodity, something to be attracted and possessed. Conversely, in the wake of algorithmic management, and especially in organisational contexts characterised by scarcity, attention becomes a resource, something that can be mobilised to achieve something else (in this case, efficient patient care), and that must be used to the maximum extent possible, lest it is wasted (Liboiron 2021).

A crucial corollary of this point is that, to achieve the maximum use of nurses' attention, the dashboard's script operates a redistribution of responsibility, particularly as it relates to patient safety. In fact, a considerable part of the responsibility for deciding what counts as "urgent" and what deserves nurses' attention is diverted from the nurses themselves to the dashboard's data analytics. Although nurses' attention has been described as characterised by micropolitics (Felder et al. 2023) that can be biased and unjust, it is worth noticing here how this dashboard attempts to turn what is inherently matter of professional ethics (i.e. who or what needs attention more urgently in a context of scarcity) into a technical question that is solely predicated on data. Interestingly, during sessions, nurses put forth the request to be themselves the ones to decide which patients qualified as green, orange, or red — something that would defeat the purpose of mobilising data analytics. Data analytics, in the dashboard, take over part of nurses' decision making, turning them into passive recipients of an analysis. Working with a passive notion of attention, the dashboard also fails to take into account how data itself is imbricated in and sustained by nurses' care practices. Below, we theorise attunement as a notion that sensitises us to how these practices, in overflowing mainstream notions of attention, prove crucial for both patient safety and data production.

Attunement, care, and data

Assembling attunement

Rather than conceiving attention as a passive capacity or as a filtering device, scholarship in anthropology brought forward an alternative conceptualization of attention as an active doing (Pedersen, Albris and Seaver 2021). In this section, inspired by the experience of conducting ethnography in an ICU, we assemble a concept of attunement that helps us rethink what attention might be in datafied clinical settings.

A concept often mobilised in affect theory, attunement is etymologically cognate to sound and sonic experiences (at-tune), having to do with instances of shared affect amongst bodies "fleshy or otherwise," in which sounds plays some role (Garcia 2020; Gibbs 2010). Sarah Ahmed describes attunement as a body's orientation towards something that causes it to inhabit a space in a particular way and, subsequently, to pick up on some affects rather than others (2014). Different bodies can be differentially attuned, even in a shared environment, because of their diverging histories (2010). Attunement thus has to do with how specific experience, habitus, and work on the self have shaped a body. Attunement is also a matter of proximity, of how this being affected reshapes a body and what comes into contact with it, making the two somehow fit together. Ahmed's work helps us specify the bodily dimension of attunement: attunement emerges as an embodied disposition developed within and towards a specific environment, which entails a learning process as much as a reshaping of one's body. Attunement is an ongoing accomplishment, and lies in the progressive, reciprocal adaptation of a body and the objects it interacts with. Attunement, crucially, entails energy expenditure, and points to a non-innocent process of wear and progressive exhaustion.

The notion of attunement resonates with scholarship on sensory anthropology in clinical settings. For instance, scholars like Harris (2021) and Maslen (2017) show how professionalisation also means learning to incorporate various kinds of instruments in one's practice. Sensory anthropology reminds us of the ongoing agency professionals display in choosing to be affected by something. Attunement broadens this sensory perspective, emphasising the emotional and energetic toll that being affected entails. Especially in acute care settings, where professionals attune to vulnerable bodies and vulnerable technologies, attunement needs to counter constant (and life-threatening) risk of breakdown (Mesman 2014; Wiedemann 2021).

STS literature can assist us in thinking through the epistemic dimensions of attunement. Anna Tsing's work on the "arts of noticing" help us thinking through what might mean in information-rich and technology-dense environments, such as an ICU. Arts of noticing, for Tsing, are an essentially multispecies accomplishment that entails "watching the interplay of temporal rhythms and scales in the divergent lifeways that gather" (23). Arts of noticing point to a way of knowing that, unlike attention, is not only about isolating a specific object, but about considering it in its temporal, spatial, and ontological relations with other objects. Tsing's arts of noticing centre on indeterminacy, a "particular kind of attention to the here and now of encounter, in all its contingencies and surprises" (46). This is an open mode of perceiving that enables picking up and even centring on the unexpected, rather than filtering it out as noise. Tsing assists us in conceiving of the epistemic dimension of attunement in a way that exceeds dominant notions of attention. First, attunement is an open mode of perceiving that does not need to preselect a single focus, but that rather takes in a complexity of interconnections. Second, and related, attunement is not an exclusively human doing, but is a multidirectional achievement performed jointly with other humans and nonhumans.

The type of perceiving Tsing describes resonates with Karin Bijsterveld's (2018) taxonomy of listening practices. Bijsterveld problematises traditional dichotomies pitting active (or attentive) listening against passive (or distracted) listening. Modes of listening vary according to their purpose (from indeterminate to more analytical listening), and according to the way the listening is done (zooming out on all the complex of all sound in an ensemble, or isolating a particular sound within an orchestration). Moreover, Bijsterveld shows that different modes of listening can co-exist, and that listeners can develop virtuoso-like ways of switching between them. Given the etymological and conceptual affinity between listening and attunement, Bijsterveld analysis helps us think through the co-existence of different modes of attunement — that is, attunement might look different for objects at different scales, and it could shift between from analytically-focused to more indeterminate modes, without necessarily give way to "distraction." In what follows, we give substance to the notion of attunement by tracing it in ICU nurses' care practices.

Tracing attunement in nurses' practices

ICU patients are connected to a myriad of technologies, some of which take over or support their vital functions (e.g. the extracorporeal membrane oxygenation machine, ECMO; and mechanical ventilators); others measuring and displaying their vital signs (e.g. heart rate, blood pressure, oxygen saturation) in real time (figure 2). Finally, patients are connected to intravenous (IV) pumps administering medications. Nurses' object of care in the ICU is composite, made up of bodies, technologies and data. The very technologies that keep patients alive also produce the real-time data that provides a proxy for monitoring patients' health — and, incidentally, also makes the ICU an attractive environment for the application of AI technologies.

In this section, we show how nurses become attuned to their individual patients and to their attending technologies and data by fostering the mutual attunement amongst components of their composite care object, and intervening to fix moments of dis-attunement amongst them.

(Dis-)attuned machines

Although they prepare for it during their one-on-one handover with their colleague from the previous shift, and by scanning previous reports in the EHR, nurses' attunement to their care object begins when they step inside the patient's room. Nurses' workflow is structured around these moments of attunement: at the start of the shift, and then every two hours, they conduct a "patient round," during which they check the settings of IV pumps, ventilators, ECMO, and set alarm thresholds. The first patient round is time-consuming: nurses first attune to each nonhuman component of their care object, checking that their settings correspond to what is reported in the EHR, and that they are well-positioned and working properly. The latter aim, especially, requires shifting between attunement to different components (including the patient's body), to notice possible dissonances between them:

> Tom moves to the ventilator, examining not only the tubes and the values displayed on its screen, but also the pressure of the cuff — an inflatable balloon

> that's inside the patient's throat, and that, Tom explains, makes sure the tubes don't move. … He measures the cuff's pressure with a barometer, explaining that usually, if the pressure is not sufficient, "you can just hear it." A while later, when we come back in to wash the patient, Tom invites me to get closer to the patient's face, and I hear a gurgling sound. "This means there's not enough pressure," Tom explains, and proceeds to pump air into the cuff. (fieldnotes)

Figure 2: Some of the technologies inside an ICU patient's room: a vital sign monitor, a ventilator, and a computer displaying the patient's EHR.

Although the pressure of the cuff inside the patient's throat returns a "good" number, Tom remains attuned to it. Whereas the protocol guiding the patient round would require nurses to tick examined items off their list (both mentally and in an EHR form), Tom's attunement to his care object means persists even when a task is completed. Attunement is not selective in way attention is, and it does not subdivide care provision into separate tasks: Tom keeps listening for dissonant sounds that might not align with recently obtained information. Interestingly, information sensed from the patient's body takes precedence over the quantitative data returned by a device. The dissonance between sensed data and other information returned by the care object is interpreted by Tom as a sign of dis-attunement, requiring intervention. When the dissonant gurgling stops, the care object is considered re-attuned.

(Dis-)attuned patients' bodies

In other cases, it is the patient's body that might need to be re-attuned. If, as we saw above, checking life-supporting machinery might entail getting close to and sensing the patient's body, the second part of the patient round, which aims at directly examining the patient's bodily functions, combines moments of sensing the patient's body (e.g. looking at pupils, feeling limbs' temperature) with an emerging attunement to the data that act as a proxy for its physiological functions: vital signs. Nurses become attuned to technologies by listening to patients' bodies and, as we see below, become attuned to patients' bodies by looking at numbers on a monitor:

> Having checked all the machines' settings, Tom moves to the patient's bed. … Since the patient's oxygen saturation levels are low, he informs me he will check for mucus in his ventilator tube. With a little suction tube, he removes some mucus from the tube. The saturation data on the vital signs monitor go up quickly. (fieldnotes)

If data allow Tom to attune to this patient's body, he still does not take this information at face value. Low blood oxygen does not necessarily have to do with the ventilator's settings: patients' bodily processes can interfere with measurements. Because of the dissonance between vital signs (saturation is low, but blood pressure and heart rate are not concerning), Tom suspects an emerging dis-attunement in his composite care object, and re-attunes it by intervening on the patient's body. This, in turn, improves the quality of care, enabling, simultaneously, the patient to breathe better, and the produced real-time data to fall into a more acceptable range.

During patients' rounds, thus, nurses attune themselves to their care object, but, while doing so, also attune different components of it, ensuring they are working well together. Attunement also entails attuning the care object to oneself, making it work for oneself by tapping into its automated agencies — for instance by replacing medication pumps close to running out at a convenient time, or setting alarm thresholds. Thresholds vary also because different ranges of "normal" apply to different patients: as a nurse explained, "the normal saturation level for a patient with COPD could be cause for concern in someone who is healthier." Setting thresholds shows how the very possibility of alarms meaningfully participating in the care object is predicated on nurses' being attuned to a patient's body, understanding what can be expected from it.

(Dis-)attuned data

Instances of dis-attunement are routinely spotted and intervened upon during patient rounds, while nurses are themselves attuning to the care object. But some signs of dis-attunement, which are subtle and do not trigger an alarm, are most easily spotted by nurses in between patient rounds:

> Looking at her patient's vital signs monitor, Nina notices something off in his arterial line tracing. Since the peaks of the tracing are overshooting, she suspects there might be an artifact. We go inside the room, where she tries flushing the patient's arterial line. The tracing on the screen becomes a little flatter, but apparently not enough. She explains that these artifacts sometimes have to do with the way patients hold their hands: the cannula is inserted in their wrist so, if they bend it, sometimes the measurements are altered. She tries to put a cushion underneath the patient's arm — but also this doesn't seem to change things much. She figures that the cannula must have moved slightly as the patient moved his hand, "and he's had it for a while already." (fieldnotes)

Nurses validate data every two hours, during their patient rounds. Beyond the protocolised task, however, the attunement that nurses maintain during their shift appears to be what enables them to spot artifacts, that is, unreliable, dis-attuned data. Artifacts can be indicative of vulnerabilities in the system: the strange arterial line tracing noticed by Nina, called "underdamping," for instance, can lead to wrong estimations of blood pressure (Saugel et al. 2020). Obviously, unreliable blood pressure measurements are extremely risky for ICU patients: a slight change in blood pressure can be an early sign of a rapidly deteriorating situation (Mesman 2014). Getting rid of artifacts, and producing reliable data, is thus a crucial matter of safety.

Identifying dis-attuned data, however, is not sufficient: re-attuning the care object is an embodied activity that needs to account for the porous boundaries between patient's body and technology — thus probing various potential origins of the dis-attunement. Trying to deal with artifacts entails opening up the enacted nature of data, produced at the interface of potentially faulty sensors, potentially misplaced cannulas, and inconveniently bent joints. Nina, above, knows that for data to provide a reliable insight into the patient's body a balance must be struck amongst different components of a care object. Investigating the causes of artifacts, and trying to eliminate them, entails a time-consuming, open-ended process of tinkering with both technologies and the patient's body.

(Dis-)attuned nurses

Managing the attunement of the care object is a complex task and one that must be learned, difficult to master for less experienced nurses:

> The blood pressure of the patient starts crashing, announced by a myriad of alarms. Ralph, a nurse in training, tries to increase the dosage of noradrenaline, but this doesn't seem to do any good. His supervisor, Will, tells him that sometimes, if you increase the dose of noradrenaline too quickly, you can paradoxically end up with a drop in blood pressure. Ralph stops fiddling with the IV pump's settings, while Will leaves the room. … Soon, a yellow alarm informs us that the patient's blood pressure is crashing again. Ralph increases the noradrenaline, and turns off the alarm, waiting for the medication to have its effect. I can see he's very nervous … He tries to adjust the patient's arm and the height the line's valves, but the blood pressure does not improve. He pages Will, who comes in immediately. They try changing the position of the bed so that the patient's legs point upwards. They wait a bit, and then increase the noradrenaline again — to no avail. Will notices in the arterial line tracing a pattern that might indicate a heart tamponade: some of its peaks are much lower than others. They decide to call the intensivists. … The intensivists don't come in until much later and suggest lowering the ventilator's settings. This makes the patient's blood pressure stable again. The alarm stops. (fieldnotes)

The same awareness of data's enacted nature, which made Nina suspicious of her patient's blood pressure data, makes diagnosing the situation a complex and sometimes confusing task. When is an abnormal blood pressure value a sign of a deterioration in the patient's state, to be intervened upon by administering noradrenaline? Not unlike Nina, Ralph and Will do not take real-time data for granted: they doubt the reliability of the measured blood pressure and try to rule out measurement-related artifacts by readjusting the patient's body in various configurations.

Only after this probing of the care object — adjusting both medications' settings and the patient's body — does data become data. Only at this point does Will turn his gaze to the arterial line tracing. Only when traced back to a possible sign of cardiac tamponade (a possibly fatal build-up of fluid in the pericardium) does data become a cause for concern. In this case, taking the arterial line tracing at face value, and the concern that this specific tracing causes in the nurses, results in a moment of dis-attunement: Will and Ralph do not check the ventilator's settings, which later turn out to be the culprit.

Despite focusing on the dis-attunement of different components of the care object (machines, patients' bodies, data and nurses), all the vignettes in this section point to how an attuned care object relies on the coordination of its different components, producing both a stable patient body and reliable data about it. All vignettes show the affectively-charged and time-consuming work that nurses perform to attune and become attuned to their care objects. Crucially, part of this attunement lies in questioning the reliability of the signs displayed by different components. This is particularly true for the real-time data providing insight into patients' bodies: these data are appraised as enacted at the intersection of always possibly dis-attuned bodies and technologies.

Shifting modes of attunement

So far, we have teased out different aspects of nurses' (dis-)attunement to the care object inside ICU rooms. In this last section, we follow this care object as it extends outside of patients' rooms in the form of pagers, monitors, and their alarms. When in the corridor, nurses are supposedly not attending to their patients — that is, not providing direct patient care. Sitting at their desks, they perform administrative tasks (filling in the EHR, ordering tests, and checking lab results). In the corridor, handovers take place, and nurses talk to each other and take breaks. However, even in the corridor, nurses maintain a level of attunement to their care objects — a certain degree of orientation towards technologies and data and, through them, towards the bodies inside the rooms. Crucially, through monitors and alarms, the object of nurses' attunement is also expanded to the unit's aggregated patient population (figure 3).

Figure 3: View of ICU nurses' desk. Above the computers through which EHRs can be accessed, monitors display the vital signs of all the unit's patients in real time.

We can think of alarms as technologies of attention, rather than attunement. Yet, by being on the ICU, nurses develop a certain kind of attunement even to alarms. Indeed, since they are virtually constantly going off, becoming attuned to the ICU environment also means learning to distinguish different types of alarms, and becoming able to ignore some of them. Different alarms point to different things: a blue alarm signals a disconnected sensor, a yellow one a deviation of a value from its set thresholds; only red ones indicate emergencies. While blue alarms only make the vital signs monitor beep once, yellow and red alarms set off the monitor's beeping (though at different speeds) and the nurses' pagers. Learning about alarms is a matter of becoming attuned: gradually, one starts ignoring the blue ones and stops being startled by yellow alarms, unless they keep recurring. Being what Ahmed (2010) would term an "affect alien" (30), a body whose responses are unaligned with the other bodies in a specific environment, brings into relief nurses' attunement to data and alarms:

> Sitting at the desk, Velma and Victor are sharing grievances about the pharmacy. I keep being distracted by a red alarm going off in room 10. Above our heads, the part of the monitor displaying the electrocardiogram keeps turning red and displaying "ASYSTOLE," and Velma and Victor's pagers keep beeping, though they both keep silencing them. … At last, I bring up how "things seem not going well in room 10." Velma giggles and tells me not to worry: "It's just the sensor malfunctioning. This patient has blood pressure; had it been an asystole, everything would be flatlining." She turns off the monitor's alarm and resumes her conversation. A bit later, the alarm goes off again while Velma and Victor and in their respective patients' rooms. This time, they both jump out quickly to check what's going on and if their help is needed. (fieldnotes)

Attunement to the extended care object, and to the dissonances among different components, is precisely what enables nurses to know what not to pick up on. The vignette above also shows how different locations differentially enable attunement at the level of the unit. When sitting in the corridor, nurse have access to monitors, and are thus able to notice the dissonance between different data. Once inside patients' rooms, where only one patients' vital signs are displayed, the very same auditory cue cannot be contextualised as easily, and causes a very different reaction. Re-attuning from patient to unit level requires nurses to interrupt patient care to be able to dismiss a false alarm.

During downtime, when nurses are in in the corridor, alarms do manage to spur their re-attunement, surfacing a specific object of care:

> Deborah and I, together with most of the other nurses, are sitting in the middle of the corridor drinking our coffees and chatting, when an alarm goes off. The monitor above the desk in front of which we're sitting starts blinking, the blood pressure value turning red. All the nurses all of a sudden go silent, turn their heads towards the monitor, and immediately jump out of their chairs. A second later, Jamie emerges from the room, and yells "cart!" Deborah runs to get the crash cart and pushes it in front of the door, while another nurse gets the defibrillator (which is still plugged into the wall, so another nurse needs to run after her and unplug it). As I follow them, I see a doctor performing a heart massage amongst a cacophony of alarms. (fieldnotes)

Above, alarms work as imagined, apparently redirecting distracted nurses' attention successfully. When dealing with vulnerable bodies and vulnerable technologies, the constant threats of breakdown make it impossible to dis-attune oneself from the unit-level care object. The red alarm suddenly changes an atmosphere that had been until then convivial and relaxed — it turns a shift that Deborah had described as "probably a bit boring" into one that requires calling in a heart surgeons to perform emergency surgery, and that left nurses and doctors alike utterly exhausted. A red alarm re-orients bodies, it affects them, pulls them towards the monitor and out of their chairs. A red alarm initiates but does not, alone, accomplish re-attunement: not only are alarms, like data, rarely taken at face value, but, even in acute care settings, atmospheres are sticky (Ahmed 2014), and changing them is anything but a frictionless process. It remains hard to believe that a reanimation is happening until Jamie asks for a crash cart.

If we were to think with dominant notions of attention embedded in the dashboard, it would be hard to account for the inertia of re-attunement. We would leave out the incredulity that the unexpected produces, and neglect how emotionally taxing re-attunement can be:

> After dinner, the whole team, including the doctors, sit down in a circle in the doctors' room. … Amy, one of the nurses, shares how stressful and unexpected this reanimation setting has been for them: "All of a sudden we saw the pressure drop on the monitor and had a moment of hesitation: is this really happening?" The others agree that they went into panic mode. (fieldnotes)

Having to re-orient one's body to pick up what it did not expect to pick up generates panic. When alarms do their job, re-orienting a body is not a simple shift from "distracted" to "attentive." Red alarms and the panic they generate teach us that care objects themselves are affectively charged, and that switching from downtime to emergency takes a toll on nurses. Moments like this teach us about how task-switching is not just about turning attention off and on, or selecting a different object: it is a matter of re-attuning, and thus cognitively- and affectively-charged.

Discussion and conclusion

Building on an ethnographic case study, this paper has discussed current tendencies towards mobilising data analytics for intervening on healthcare workforce shortages. We have thematised attention as an emerging object surfaced and managed through algorithms in contexts of human resource scarcity. The data analytics in the dashboard we analysed enabled sustaining a notion of attention, long-lived in Western modernity, as a rational cognitive state easily redirectable through data. Compared to earlier notions of attention as a filter, algorithmic technologies appear to take over the filtering of stimuli that attention was supposed to perform, effectively attempting to turn attention into a resource to be extracted to the maximum.

Since the dashboard was not in use during our fieldwork, we could only offer a glimpse of its potential implications through an ethnographically-grounded analysis of nurses' current care and data practices — the very practices that the dashboard attempts to make efficient. To make more tangible the chasm between these practices and the notion of attention underpinning the dashboard, we have offered a notion of attunement assembled at the intersection of STS, cultural studies, and our own empirical material. Attunement represents an alternative and more sound heuristic to think about work in data-rich and technology-dense care settings. Unlike attention, attunement is an embodied, more-than-human accomplishment, in which different actors in a situated environment adjust themselves to each other, becoming accustomed and alert to their respective rhythms, expressions and needs. Attunement is less exclusionary than attention, but has a dark side in the affective and physical toll it can take on who performs it. Whereas frameworks centred on attention postulate nurses' externality to the technologies and the data deployed in the ICU, our analysis has shown how attunement is rooted in the experience of affecting and being affected by the ICU care objects — objects that compound and confound technologies, data, and bodies. Table 1 offers a summary of the main differences between attention and attunement.

Our analysis contributes to STS and sociological discussions around the digitalisation of (healthcare) work by providing an empirically-grounded reflection on the organisational and professional ramifications of emerging machine learning technologies (see chapter one). If there is still a dearth of analyses of machine learning in practice (Jaton and Sormani 2023), we have proposed here that a generative angle for examining such technologies might lie at the intersection of the interest in attention emerging within the STS literature (Jablonski, Karppi and Seaver 2022) and sociological discussions around the algorithmisation of the workplace (Jarrahi et al. 2021). Turning algorithmically-directed attention into an object of inquiry enables us to analyse how work and workers are being reimagined in the wake of machine learning, whose work these technologies try to make more efficient, and what work falls outside of efficiency's scope.

Attention	Attunement
Acting as a filter; process of selection of focus from multiplicity of stimuli	Indeterminate in focus
Passive capacity, consuming information	Different modes, including active doing, producing and probing information
Commodity made scarce by digital media and information proliferation	Engages with polyphony (multiplicity of rhythms and scales)
Cognitive on/off state, with no degree to it	Embodied, affectively-charged
Must be used to the maximum and made efficient (resource)	Flourishes in moments of downtime, incompatible with efficiency
Deals with parts (taskification)	Shifts between wholes and parts, focuses on relations
Optimisable: shifting foci is a seamless process	Sticky: re-attuning is cognitively and affectively taxing
Human capacity (though AI's feature extraction models it)	Situtated in multidirectional more-than-human relations

Table 1: Comparing attention and attunement

Perhaps unsurprisingly, in our case, nurses were singled out as targets of increased efficiency (Maslen 2017; Mort et al. 2003). Efficiency rested on a disregard both for the invisible work (Star and Strauss 1991) of attunement, and for the visible, but undervalued, work of data production (Bossen et al.

2019). Charting users' reactions to attempts at algorithmically managing their attention, as well as the different materialisations of attunement in different contexts, which both falls outside the scope of the present paper, strikes us as a worthy focus for future research.

Our analysis also bears important implications for practice. Investing in technologies, and specifically machine learning, is increasingly cast as the last resort for ensuring the future sustainability of healthcare systems through increased efficiency (Agyeman-Manu et al. 2023; Deloitte 2023; Gupta Strategists and Edwards Lifesciences 2021). Our analysis of attunement suggests caution around narratives of technologically-achieved efficiency. Our empirical material shows that efficiency is often thought of as an elimination of downtime in an attempt to keep professionals in a constant state of productivity. Conversely, our analysis of ICU nurses' care and data practices shows that tasks that are not urgent, such as dwelling on dissonances and investigating their possible origin, are crucial for both patient safety and for the production of reliable data. The importance of this care-and-data work should be inscribed within any technology that attempts to intervene on attention in acute care settings. Although proposing alternatives for technology design falls beyond both the scope of this paper and our expertise, making the concept of attunement central to technologies in clinical settings might provide avenues for technology design to be more in synch with the needs and rhythms of care provision. We invite clinicians and technology developers to think, jointly, about what might happen if we started thinking about care work as driven not solely by data, but by attunement to bodies, technologies, and data themselves.

Chapter four
Doubt or punish[13]

[13] An earlier version of this chapter was published as: Carboni C, Wehrens R, van der Veen R, and de Bont A (2024). Doubt or punish: On algorithmic pre-emption in acute psychiatry. *AI & Society*.

Introduction

The relationship between (healthcare) professionals and emerging artificial intelligence (AI) tools has become central in contemporary academic and public debates. Questions address the ways in which these technologies are likely to replace professionals (Chockley and Emanuel 2016), take over their tasks (Wong et al. 2019), alter the scope of professional decision-making and the nature of their work (Bullock 2019; Chan and Siegel 2019) or professionals' ability to domesticate these tools in practice (Avnoon and Oliver 2023; Topol Review 2019). These questions are usually discussed speculatively, and appear to be anchored in a future in which the use of novel technologies would have normalised. However, regardless of the slow pace of their implementation (Koutsouleris et al. 2022), AI tools are entering clinical practice along less formal routes, such as pilots or living labs (Archibald et al. 2021). This phenomenon also spans fields traditionally resistant to digitalisation, such as psychiatry (Bourla et al. 2018; May et al. 2001; Pickersgill 2018), where AI tools are being mobilised, for instance, to predict mood shifts in people with bipolar disorder (Semel 2021), classify and predict behaviour in patients (Fernandes et al. 2017; Mulinari 2023), predict suicide attempts (D'Hotman and Loh 2020) or violent incidents (Borger et al. 2022).

If analyses of AI tools in practice are still few and far between (Jaton and Sormani 2023), empirical studies have begun to show the complexity inherent to processes of embedding AI systems in professional decision-making. Studies have shown, for instance, how AI can introduce new ambiguities in routine decision-making (Lebovitz 2019) and to their selective reliance on algorithmic outputs (Maiers 2017). This scholarship crucially nuances our understanding of the ways in which algorithmic outputs are reshaping clinical practice in ways that exceed dominant narratives. Simultaneously, it alerts us to the potential marginalisation of other forms of knowing currently present in clinical practice, which, in times of crisis and workforce shortages, might end up sacrificed at the altar of algorithmically-achieved efficiency (Henriksen and Bechmann 2020; Maiers 2017; Russell 2012; Schwennesen 2019)

Contributing to this emergent strand of scholarship, in this paper we analyse the pilot of an algorithm for the prediction of inpatient violence[14] in acute psychiatry clinics. By analysing nurses' reports in the electronic health record (EHR) and identifying "predictive terms," the algorithm was meant to provide a risk score for individual patients, thus flagging which patients warranted professionals' attention on a specific day. In acute psychiatry, professionals' decisions are not only clinically relevant, but sanction sometimes violent disciplinary measures, and have direct consequences for patients' freedom. Zooming in on a pilot enables us to tease out the tensions between pre-existing practices around violence and emerging algorithmic logics. Attending to the different ways in which violence is constituted as an object for intervention, respectively, by the algorithmic output signal and by the local psychiatric nurses, this paper thus centres on the question of the implications of algorithmic outputs for professional decision-making. Specifically, we focus on how, and through which kinds of information, different agents (nurses, quantification instruments, algorithmic risk scores) construct inpatient violence, and on the implication of these different constructions in suggesting specific kinds of interventions.

In what follows, we first detail our theoretical footholds. To illuminate organisational and professional techniques for dealing with future risk, we combine work on future-oriented knowing and acting (Clark 2016; Star 1991) with examinations of quantifications' and algorithms' ethics and performativity. By dwelling on reflections on the madness on decisions and its obfuscation in algorithmic logics (Amoore 2020; Derrida 2001), we come to characterise decision-making, particularly in acute clinical settings, as a complex moment of ethical deliberation, which algorithmic predictions attempt to obliterate. After laying out our methodology and describing our case more in depth, we turn to our empirical material, analysing, respectively, psychiatric nurses' practices around recognising, understanding, and dealing with inpatient violence, their use of quantification instruments, and the way the piloted algorithm attempts to intervene in nurses' practices.

[14] In this paper, we use the term "violence" to indicate a behaviour resulting in an incident (i.e. physical attacks to objects or people, or clear verbal attacks). With the term "aggressivity," we refer to behaviour that can be identified as warning signs for violence (e.g. restlessness, tone of voice, inappropriate language).

Theoretical footholds

Future-orientedness

Future orientation has long been identified as a core aspect of many areas of social life. Forms of anticipatory governance (Flyverbom and Garsten 2021) are fundamental in assembling organisational knowledge and shaping organisational futures. In the context of computer-supported cooperative work, Adele Clarke (2016) has offered the concept of anticipation work, a type of invisible work engaging with the future as a space for (professional) action, and needed "to optimise and live in preparation" (90; cf. also Star 1991; Strauss 1988). Clarke describes anticipation work as made up of three components: abduction, simplification, and hope. Abduction, a notion she borrows from Peirce, happens especially in conditions of "genuine doubt or uncertainty or fear or great pressure to act" (quoted in Clarke 2016: 92; emphasis in original). It entails collecting empirical information and producing theories about it in an ongoing, yet tentative, way. Because of the inherent uncertainty of the conditions under which it is performed, all hypotheses generated abductively are adopted "on probation" (91) and are thus always open for reconsideration. Simplification works by setting the boundaries and managing the complexity of the situation engaged. Finally, for Clarke, "anticipation comes pre-wrapped in affect — hopefully inflected" (97) — that is, anticipation work is affectively laden and fuelled by some optimistic belief that the future might (be made to) be better than the present — at least for some. In our reading, anticipation work is also ethically laden, and ethics manifest in the probatory nature of hypotheses and theorisations anticipation work produces, which acknowledge doubts and the partiality of any knowledge.

If anticipation work helps us think through future-oriented epistemic practices, this type of abductive knowing concretises in action-oriented decisions. In this sense, Anderson (2010) proposes a typology of anticipatory action to flesh out the practices in the here and now "paradoxically" justified by a (possible) future (778). Like anticipation work, anticipatory action emerges in conditions of uncertainty. Clarke's notion of hope becomes here a matter of staving off undesired futures, making sure "that no bad surprises happen" (782). Anderson identifies three forms of anticipatory action: pre-emption (acting to neutralise threats that are yet to emerge), precaution (intervening before a perceived threat reaches a point of irreversibility), and preparedness (preparing for the aftermath of an event). In the context of acute psychiatry, precaution might mean increasing the sedation of a patient who has been raising their voice; pre-emptive action would entail sedating a patient that an algorithm flagged as at risk of violence; preparedness entails making sure that sedatives are available on the ward.

Anticipation work is ongoing in acute psychiatric wards. Anticipation work, however, needs to consolidate into a decision — a moment that has been described as having a uniquely ethical character. Louise Amoore (2020) has recently revisited Derrida's (2001) reflection on the madness inherent to all forms of deciding in the wake of algorithmic decision-making. Since both the future and the implications of any decision are unknowable, Amoore argues that any form of anticipatory action is by nature a moment of ethical deliberation that exceeds purely epistemic consideration. As she states, "to decide is to confront the impossibility of the resolution of difficulty; it is madness in the specific sense that it has no unified grounds" (Amoore 2020: 112). Though we return to Amoore's work later in this section, her discussion of the madness of decisions enables us to point out that anticipatory action entails ethical deliberation and requires ways to be response-able for the ramifications of one's decisions (Suchman 2023). Following Amoore, foregrounding the doubt that spurs attempts to trace and open up alternative accounts, as well as and being cognizant of futures foreclosed by human and algorithmic any decision are crucial aspects of such response-ability.

In what follows, we work towards a framework aimed at embedding different risk prediction instruments currently used in psychiatric practice within a theory of anticipatory action. Guided by our empirical material, we focus mainly on two forms of risk prediction, namely, quantification-based risk prediction, and algorithmically-enabled pre-emption.

Quantification and capture

Risk assessment instruments in psychiatry quantify aggressivity, providing nurses with checklists to score behavioural expressions presumed to be predictive of violence (see also section 4.1). Quantification, that is, the "production and communication of numbers" (Espeland and Stevens 2008:

401), has generated far-ranging sociological interest (Popp Berman and Hirschman 2018). Gazi Islam (2022) breaks down the phenomenon of quantification into processes of capture, specification and appropriation. Processes of capture are central to how aggressivity is made into a graspable object in psychiatric clinics. Defined as the "process of objectifying [a] social phenomenon so as to express it as a numerical quantity," capture often entails "high levels of processing, manipulation, or abstraction" of aspects of social life (199). Moreover, quantitative capture decontextualises one aspect of a lived experience in flux, and simplifies by decontextualising it. By attributing value to the numerical expression of knowledge and experience, capture thus risks dismissing more complex and relational accounts.

Forms of quantification can be seen as integral to the simplification component of anticipation work, in which it participates by performing the "empirical," by simplifying and producing what reality "is." Thus, albeit not future-oriented per se, quantification techniques can be integrated into forms of anticipation work. Because of this lack of inherent orientation to the future, however, we can speculate that quantification, in and of itself, does not suggest relating to the future through either precaution, pre-emption, or preparedness, but can be mobilised by actors as part of any type of anticipatory action. Moreover, the numbers that quantification instruments produce have, in themselves, no direct claim to the future. These numbers enact objects (e.g. aggressive patients) that are supposed to have existed in the past, and perhaps in the present. As we detail in the next section, this lack of claims to the future is a major way in which quantification instruments diverge from algorithms.

Algorithmic pre-emption

Unlike quantification instruments, machine learning algorithms can be seen as engaging in anticipation work. Analyses of computational technologies has emphasised the experimental way in which algorithms engage with reality. For Amoore (2020), the experimental nature of algorithms has to do with their ability to fine-tune and adjust parameters in an ongoing manner, thus subscribing to an eternally-shifting version of the truth. Similarly, for Luciana Parisi (2019), in algorithmic engagements with the world, contingency and fallibility become productive forces in machine learning: algorithms learn through trial and error, generatively incorporating their own failures and thus engaging in an ongoing mode of optimisation that can be deferred *ad infinituum* (Halpern and Mitchell 2022).

Though both Amoore and Parisi chiefly discuss deep learning, the points they raise around the nature of learning apply to a broader class of machine learning algorithms. For our purposes, we can distil two ideas from these discussions. First, as Parisi argues, these algorithms perform their own particular abduction, adopting hypotheses on probation and remaining open to learning from them when they prove wrong. Second, as Amoore points out, the truth against which these hypotheses are tested is removed from the lived world and anchored in a ground truth dataset. This entails that algorithmic outputs cannot be judged as "true" or "false," but must be seen as a (tuneable) function of their "probabilistic proximity to, or distance from, a ground truth" (2020: 136).

If algorithmic abduction resembles other forms of anticipation work, algorithmic output represents a unique form of anticipatory action. As Amoore (2013) argues, algorithms' relating to the future, especially when mobilise to intervene on some form of risk, is fundamentally pre-emptive. That is, by generating novel targets for action, algorithms tend to be geared towards creating the possibilities for action on yet unknown (or even still non-existing) threats. In Amoore's view, creating (and justifying) the possibility to act on a threat matters more, in an algorithmic logic, than the actual materialising of the threat itself. Rather than merely offering a representation of reality, thus, algorithms open up a space of action — and they do so through the generation of a univocal output signal. As Amoore puts it,

> among the most significant harms of contemporary decision-making algorithms is that they deny and disavow the madness that haunts all decisions. To be responsible, a decision must be made in recognition that its full effects and consequences cannot be known in advance. (2020: 120)

It is precisely in this effacing of the myriad of doubts and alternative explanations they engaged in during processes of learning that algorithms fail us at the ethical level. Presenting a technically-optimised output obfuscates the ethical nature of anticipatory action. Even more, algorithms propose pre-emption as the optimal future orientation, assuming the possibility of a future in which risk and ethical difficulties can be computationally. What is problematic here is not only the replacing of ethical doubts with the stochastic management of uncertainty, but also the lack of accountability for alternative futures that the univocality of algorithmic output (and of fantasies of optimisation) forecloses. The doubt that we teased out, above, as constitutive of the ethics of decision, is here effaced. In discussing possibilities for ethics in the wake of algorithmic decision making, Amoore (2020) suggests reinstating doubt at every point of machine learning's engagement with the world: from data to ground truths, to datasets, to outputs. In our analysis, we follow this doubt as it differentially manifests, and sometimes is dealt with, in different practices around aggressivity and violence on acute psychiatric wards.

Settings and data

This paper builds on the ethnographic study of a three-month long pilot that took place in two acute care clinics in a Dutch general psychiatric hospital. Despite both comprising closed wards and being very connected in practice (they quite regularly exchange patients depending on whether they stabilise or destabilise), these two clinics are different in size and patient population. Clinic 1 is a high-intensity care (HIC) clinic made up of 3 wards sharing 4 isolation cells (figure 1). Each ward has between 8 and 10 rooms with en-suite bathrooms, as well as a communal living and dining room, and indoors smoking rooms. As a HIC clinic, clinic 1 receives all the involuntary admissions from the area — people brought in by ambulances or police after some public incident. More rarely, patients admit themselves. Although "voluntaries" are, in principle, able to leave whenever they want, the clinic's personnel are always liable for discharging them. This means that, when staff suspects patients could be at risk of suicide or violence, they can apply for a "care authorisation" (*zorgmachtiging*), which revokes the patient's right to discharge themselves, forcing them to receive care. This is a long legal process in which lawyers and external psychiatrists are involved. In case of emergency situations, for instance in the case of acute admissions, "crisis measures" (*crisismaatregelen*) are usually granted, which however only last three working days (VWS n.d.). In clinic 2, the algorithm was piloted in a medium- to high-intensity ward, meant for patients who display less aggressivity or resistance to medications, but who are considered not stable enough to be discharged yet. This ward houses more patients, and has smaller rooms and shared bathrooms. Legal provisions in force in clinic 1 also apply to patients here. All the wards tend to operate at capacity.

Figure 1: A view of the back of clinic 1. Although all doors lock automatically in acute clinics, high-intensity care units are surrounded by fences.

Because of the constant threat of new admissions, particularly on weekends, when people tend to use substances, clinical staff is always under pressure to discharge patients and thus free beds. Both clinics work with a system of "freedoms:" as they stabilise, patients are allowed to do progressively more things: take walks on the hospital's terrain with a supervisor, go get their own groceries, go to the hairdresser, or go home for the weekend. Freedoms can always be revoked — for instance, when patients come back from a weekend away with a positive drug test. Clinics are usually staffed by one psychiatrist, one resident doctor, and a number of nurses ranging from one to three per ward. Because of personnel shortages, flex workers are also called in each day to assist nurses. These people are often younger and have generic backgrounds as social workers, so they do not know the patients' group and have little to no experience in dealing with aggressivity.

The pilot itself pivoted the introduction of an algorithm, developed within the same organisation and geared towards the prediction and pre-emption of inpatient violence (see section 4.2 for more details). Carboni had the opportunity to follow the pilot since its first presentations in both clinics. She conducted 50 hours of observations in both clinics, with the aim of observing how the risk scores produced by the algorithm would be discussed and used in clinical practice. She thus attended daily nurse-doctor handovers, as well as meetings in which wards' capacity and logistics were discussed. Nurses and psychiatrists also invited her to join their daily morning rounds to the isolation cells and, in the first days of fieldwork, some patient consultations. However, she stopped attending the latter after realising that, although medical students would also be in the room, her presence as a silent observer tended to distract patients, and could have potentially impacted the consultation.

Instead, she started spending the shift in the nurses' station (figure 2). Unlike psychiatrists' offices, which are removed from the ward, nurses' stations are either next to or inside the wards. Moreover, they have monitors showing isolated patients in real time. Conducting observations in these settings allowed her to be static in clinics in which most doors need to be locked, thus disrupting care practices as little as possible. Nurses' stations are also the place where the most nurse-patient interactions take place, as well as data registration. Being situated there, she was able to informally discuss their care and data practices, as well as their thoughts about difficult patients, organisational dynamics, aggressivity and algorithms.

Carboni also kept in close contact with the data scientist leading the pilot. She had regular conversations with her about what she was observing in the clinics, and about the rationale for various design and implementation choices. These conversations sparked the curiosity of the data scientist, who, upon authorisation of the head of the clinics, joined Carboni on two days of observation. Carboni also conducted one two-hour interview with her, discussing more in depth the development and architecture of the model, as well as the lessons she had learned from the pilot.

Figure 2: A view of the nurses' office in clinic 1. One of the monitors shows images from the cameras in the isolation cells. Nurses sitting in the office control the doors to the wards to the right and to the left of the office.

The data scientist organised two mid-term evaluations (one for each clinic) and three final evaluations of the pilot (one for each clinic, plus one with the patient council of the conglomerate the hospital belongs to). During these evaluations, local professionals discussed their thoughts about algorithmic prediction and inpatient violence. Carboni attended all evaluations, and presented some insights from her fieldwork during the latter. Bringing up issues of non-use during these meetings — as she had previously done in informal conversations — was met with somewhat defensive attitudes, particularly by one senior psychiatrist. Nurses did not object to her observations. After the end of the pilot, Wehrens], together with the data scientist and another doctor involved in the algorithm's development, organised three one-hour focus groups (with professionals, managers, and patients, respectively) to share further results from Carboni's analysis and discuss their views around the BVC and its automation more in depth. Although they do not constitute the core of this analysis, the results of these focus groups have informed our thinking in developing this paper. The rest of the material (fieldnotes, observations from the evaluation meetings, and interview transcript) were coded abductively with the software Atlas.ti. Emerging themes, such as nurses' affective and embodied practices around aggressivity, dislike of the BVC, and playful attitudes towards algorithmic outputs are at the core of our analysis.

Violence: Understanding and doubting

Psychiatric nurses on acute wards do find themselves having to anticipate or deal with aggressivity many times a day. Since their stations are located in the wards, nurses are the professionals who interact most extensively with the clinics' patients. They are responsible for administering medications and executing routine drug tests; they also make sure all patients get enough food, organise discharges and respond to the many requests coming from patients. Although only psychiatrists have formal consultations with patients, nurses informally talk to them throughout the shift, and have a "5-minute contact" round every morning, to get a sense of patients' conditions.

Partially due to the topic of the pilot, aggressivity and violence on the ward was a common theme in our informal conversations with nurses. Being able to understand and deal with signs of impending violence constitutes a central component of nurses' work. Especially in a clinical context in which many patients are psychotic, sings of aggressivity cannot be assessed in universal or absolute terms. Instead, assessing aggressivity entails picking up on subtle embodied signs that signal deviations from an individually-defined norm: "Does the patient breathe more quickly, or more deeply than usual? Do they look at you askance? Do their muscles appear tense?" As nurses often emphasised in our conversations, "what isn't threatening in one patient could very well be a threat in another one." For instance, raising one's voice could qualify as aggressive behaviour in one patient, but only indicate hypoglycaemia in a diabetic one, and thus amount to an isolated incident that does not warrant disciplinary intervention — perhaps only a snack.

The example of the diabetic patient highlights how nurses' approach to the risk of violence entails attempting to understand the underlying causes of behavioural expressions that might be perceived as aggressive. For instance, nurses recount how there might be deeper emotional causes (e.g. being scared or sad) behind aggressive behaviour: "if you address that, also the aggressivity goes away." Taking signs of aggressivity at face value, rather than trying to understand their cause, thus emerges, first, as epistemically inadequate, because it effaces a multitude of alternative explanations by assuming a univocal link between an internal (emotional, physiological) state and its external behavioural manifestation. Moreover, and crucially, it is also ethically problematic, in that it would suggest violent interventions (sedation, isolation) that would be likely to escalate the situation.

The doubt mushrooming across nurses' accounts of how they assess the link between observed aggressivity and its cause is ethically charged and rooted in the experienced fallibility of attempts to understanding and relating to an object volatile yet threatening, such as inpatient violence. In their disciplinary power, nurses have to assess under which circumstances aggressive behaviour might be justified, and when it should be reined in. Distinguishing aggressivity as a risk factor for inpatient violence from aggressivity stemming from emotions that patients have a right to experience is as a tricky tightrope professionals have to walk in their decision-making. During handovers, for instance, professionals usually ignore the aggressivity displayed by patients who appeared "quick to anger" as long as they are able to quickly apologise for their behaviour. Even when expressed in confrontational ways, anger does not necessarily warrant escalation on the part of professionals:

> In the second isolation cell there's a woman who has been moved here because of some violent behaviour. … [when the nurse opens the door], the patient appears pretty calm. … At the end of the conversation, when the psychiatrist offers her medications, the woman gets upset. She starts crying and tells them that she doesn't want them and doesn't need them. Nonetheless, she opens her hand and receives the two white pills a nurse hands her. Kneeling down, the nurse also passes her a paper cup with some water. The woman takes a small sip and then flings the cup to the ground floor, spilling all the water on the nurse's foot. This brings the conversation to a halt, and everyone leaves the cell, locking the woman inside. But even after this incident …, the psychiatrist states that the patient appeared much calmer. The staff agrees that they are going to see how things evolve in the evening, and possibly move her back to the ward the following day. (fieldnotes)

Although throwing a cup is punished by interrupting the conversation and leaving the patient in the cell, the small incident does not affect the overall positive impression the psychiatrist had of the woman — to the extent that she suggests trying to move her back to the ward. A rebellion does not qualify as aggressivity worth intervening on. In a similar way, signs of aggressive behaviour are also considered not worthy of intervention if another relational justification for them could be identified: sometimes, patients are upset by relatives' visits, or by tensions within the patient group. Appraising aggressivity in its relational dimension thus enables nurses to understand its causes through contextualisation efforts.

At the start of the pilot, nurses had agreed to consider (and thus start reporting) as violent episodes verbal threats and physical attacks to people and objects. However, it was not uncommon for nurses to ignore acts that might reasonably fall into one of these categories. For instance, one day Ralph, one of the nurses, came back to the nurses' post saying that a newly admitted patient had just spat in his face. "So gross!" he commented, after sitting down at his desk and continuing his administrative work. The point here is not just that nurses working in acute settings might become desensitised to instances of violence, particularly when they do not put them at risk of physical harm. Our point is also not simply that they have an interest to work with as narrow a definition of violent incident as possible in order to save themselves some administrative work. Rather, we argue that working with such a narrow definition, and thus de facto ignoring episodes that could be read as violent incidents, has to do with how ethically-charged professionals perceived their violence-related interventions to be. This emerged during one of the first handovers we observed:

> When, at the end of the handover, I ask the nurse and the doctor about how they deal with aggression, they explain that what counts as a "significant episode" is very complicated: "If someone hits the window, like today, is it aggression to objects, or is it just that they got a bit angry?" Josh says. Karin, the doctor, agrees: "Do you come in and pump them full of lorazepam?" "Exactly, or put them in a straitjacket?" (fieldnotes)

What emerges here is the complexity of the negotiations nurses and physicians engage in when dealing with aggressivity. These negotiations are made complex by, on the one hand, the epistemic doubt surrounding attempts at classifying what they are seeing and, on the other, the ethical doubt that is about deciding whether interventions that are in themselves violent in their infringing on patients' autonomy, either at the physical or the neurochemical level, are warranted. Indeed, when nurses flag a patient as at risk of violent behaviour to a psychiatrist, this can set off a chain of measures that can be considered as violent in themselves: for instance, injecting medications in the bodies of patients who refuse to take them orally (an operation that requires several nurses to immobilise the patient in question), or dissolving them in their food; increasing their sedation levels; or, in the more extreme cases, locking them in an isolation cell.

Professionals' discussions of violence and aggressivity are guided by an awareness of their own fallibility, by the doubt inherent to any attempt to understand a complex reality — doubt stemming from the potential multiplicity of explanations underpinning any behavioural expression they observe. This doubt materialises in the rejection of universal warning signs (replaced by knowledge of individuals' personal and clinical histories); in their probing of counterfactual explanations for aggressive behaviour (e.g. other feelings); in their attempts at finding contextualising explanations beyond the individual that might justify aggressivity. We suggest that this is warranted in light of these decisions being perceived as an ethical moment by nurses themselves.

Quantifying aggressivity, pre-empting violence

Because of their close contact with patients, nurses are responsible for the bulk of data registration in the clinic. Indeed, since psychiatrists only write notes on the few patients they see for consultations, nurses are the only ones consistently assessing all patients and registering these assessments. Throughout the shift, one of the ward's nurses writes up their observations for each patient in a Word document, which, before the end of the shift, is both copy-pasted in the EHR and printed out to be brought to the handover. These reports are relatively standardised. First, they describe the 5-minute contact round in reasonably standardised language (e.g. "confused," "friendly," "psychotic," "verbally hyper but can be reasonably corrected"). Following this, they succinctly write up any episode worth of attention (*bijzonderheden*), which is usually articulated more in depth orally during handovers. This more narrative part of the report is followed by two highly structured sections: “risks” (*gevaar*), detailing the reason for admission or other points for attention (e.g. psychosis or suicidality); and, finally, the result of the currently deployed violence risk assessment instrument: the Brøset violence checklist (BVC) score.

In this section, we briefly describe the functioning of the BVC, its enactment of aggression as a quantifiable risk object, and the hesitations the use of this instrument brings up in nurses. We then move to the algorithm at the centre of the pilot we followed, and try to understand in what ways the risk object it creates, and the use scripted in its organisational embedding, might deviate from previous instruments of quantification.

The BVC: Quantification and its discontents

Developed in the late 1990s, the BVC is a violence risk assessment instrument aiming to offer a rough estimate of the upcoming 12-24h. It requires ward nurses to assess patients based on the presence or absence of six items: confusion, irritability, boisterousness, physical threats, verbal threats and attacks to objects. Items are to be factored in only if they represent a deviation from what is to be considered normal for a patient, e.g. boisterousness for a patient known to be quiet. Each observed item adds one point to the final score. The final score (ranging from 0 to 6) should assist further decision-making. Almvik and colleagues (2000) suggest taking “preventive measures” for scores 1-2, and activating “plans for handling an attack” for scores 3 and higher, for instance increasing sedation or isolating patients.

Similar to nurses’ practices, the BVC turns violence into a risk object that can be intervened upon before its escalation through appropriate anticipatory action (Anderson 2010). Unlike nurses, however, it does so by establishing a linear relation between factors (behavioural expressions) that it identifies as predictive, and episodes of violence themselves. Observing one or more of these expressions directly translates into an increased risk. Arguably, the direct connection the BVC operates between these behaviours and violence bears ontological implication for violence. First, identifying individual behaviour as a predictor of violence, the BVC roots its causes firmly in individuals. Second, operating within a precautionary logic and overtly disregarding “the motivations behind violent incidents” (Linaker and Busch-Iversen 1995), the

BVC casts observation of individually-defined risk factors as a sufficient condition for anticipatory action (De La Fabián et al. 2023). Finally, since the risk resides in the individual, the possible interventions that scores 3 and higher call for will also target the individual patient who expressed the problematic behaviour. On a more general level, moreover, the BVC also assumes both the risk factors it identifies and violence itself to be clear-cut, commonsensical categories that can be uncontroversially applied to categorise patient behaviour.

Perhaps unsurprisingly, given the clash between how nurses think about violence and how the BVC operates, nurses had a complicated relationship with the instruments. Some nurses refused to fill it in, while others described it in informal conversations as "a box-ticking exercise … with a lot of copy-paste going on, especially when no significant episodes happen during a shift." Both during informal conversations and more formal evaluations of the pilot, nurses objected to the awkward temporality of the BVC, which in their view necessarily only captures past events, with no predictive value:

> Lola claims that the BVC doesn't help nurses, because it always refers to a situation they have already dealt with. Once we sit down for lunch, Ralph also explains that nurses score it at the end of their shift. However, if a patient shows any behaviour that could possibly escalate, they act on it immediately. If a patient knocks some objects off a table, or threatens other people, nurses will probably give them some time out in their room, or in some cases even in the isolation cell. By the time they score the BVC, the situation has often already resolved. (fieldnotes)

As we learn from studies of quantification (Islam 2022), the BVC simplifies the complex phenomenon of patient behaviour, and de-contextualises it, extracting it from the flux of life on the ward. The BVC requires nurses to score aggressivity as if it was an external entity that they themselves were not experiencing and interacting with. As we have learned, on the ward, observations of aggressivity are rarely divorced by interventions aimed at countering or deescalating it. As a result, a high BVC ends up describing, ex-post, episodes that have already triggered cautionary measures. As we often observed during handovers, patients that were observed to be too restless were immediately given timeouts, and isolation cells were prepared as soon as patients started "thinking and acting chaotically" and becoming unintelligible for nurses.

The somewhat artificial separation between life on the ward and its punctual, quantified representation in the BVC is, to a certain extent, bridged by nurses. Crucially, nurses are both involved in interaction with patients, in the embodied, affective experience of aggressivity and violence, and are supposed to be the motor of its datafication. If quantifying something as complex and intangible as aggressivity often proves a challenging enterprise, it gives them nonetheless an opportunity to instil the doubt that characterises their approach to aggressivity in the BVC data they produce:

> [During the mid-term evaluation] Tim, a nurse, recounts how they had a patient who was proving really difficult to handle; it was hard for the nurses to pin down exactly why — and the patient would score 0 every day, though he gave them a lot to do. He goes on to talk about another patient, who sometimes would kick a door, or throw an ashtray on the floor. "That was more of a rebellion than aggressivity," according to Tim. Regardless, this other patient would score a 6 — a score that, according to protocol, mandates isolating patients. "In this case we did give him medications, but we decided to go back and change the score. They were all separate incidents, after all." (fieldnotes)

If Tim points out how separate acts of rebellion over the course of a shift do not necessarily add up to an increased risk of violence, something more radical emerges from the fieldnotes above about how nurses produce and adjust the numbers in the BVC. Tweaking scores appears justified in cases in which observed aggressivity is either hard to articulate through pre-established categories, or when it is doubted as a predictor of violence. These are moments, as we saw above, in which epistemic and ethical doubt, respectively, appear to guide nurses' anticipation work. What we want to suggest here is that these moments of tweaking amount to attempts to make BVC scores not so much a more accurate representation of reality, but a more accurate representation of nurses' doubt. By instilling doubt in the numbers they produce, nurses try to make BVC scores more internal to the life on the ward and to aggression as a relational, experienced phenomenon.

This matters because, although barely brought up in handovers, BVC scores can be used by nurses when "it's busy and [they] have so little time and personnel," and thus "need to understand quickly what's going on." Due to personnel shortages and the very nature of work on acute psychiatric wards, nurses sometimes benefit from oversimplified but quick ways to assess the situation. In what follows, we analyse the algorithm that was aimed at automating the BVC scoring, and trace how its differential relation to doubt and externality influenced its reception by nurses.

"Just a stupid outcome indicator:" Training a violence prediction algorithm

Inpatient violence emerged as a target for the algorithm partially because the organisation's data science department had access to a similar model (paper not referenced here for purposes of anonymity). The original model was intended to predict violence episodes in patients over a period of two weeks, and had previously been tested on data from the psychiatric hospital in which the pilot was run. The organisation's data science department thus decided to retrain it to predict the following 24 hours — which was considered a more actionable timeframe. The algorithm was thus developed with the aim of replacing the BVC with an automated risk assessment, and of assisting professionals in their decision-making during handovers.

When we asked Emilia, the data scientist behind the repurposed algorithm, about the early days of the project, she explained that the topic of violence had been selected by the organisation's management mostly because of the model's availability, and secondarily due to the clinical prevalence of violent episodes. From her perspective, however, violence was not an experienced threat, but rather an abstract entity — as she described, it was "a bad outcome indicator." To train the algorithm, Emilia had to mobilise machine learning techniques such as random forest and bag of words, which teach the algorithm predictive terms based on their correlation with an indicator — that is, something that is chosen as a datafied proxy for violent episodes. In the clinics we studied, nurses, as the professionals closest to patients, were responsible for reporting violent incidents through an online portal. However, as mentioned, they rarely did. This "bad outcome indicator" (i.e. a painfully incomplete dataset) caused issues in the training process and, down the line, in the performance of the algorithm:

> [The algorithm] is not doing what it should do [because] we don't have the incidents and the almost-incidents. [So] the model says, okay, I see words like "aggressive" and … "throwing" or "hit" — and there is no incident. Huh? What

> do I do? It's weird. And it *is* weird, because there was an incident, but it was not reported. (Emilia)

Concretely, the algorithm was trained on a dataset correlating nurses' notes for each patient with presence or absence of a VIM report for that patient during that shift. Based on this, it identified a series of words that, supposedly, either increase or decrease the likelihood that that patient would initiate an incident. Predictive terms included terms such as "aggressive," "reacts," "offered" (likely referring to medications), "angry," "emergency medications," and "colleague" (likely referring to the fact that nurses always approach aggressive patients accompanied by at least a colleague). Being linked to behaviour and interactions, these terms are supposed to capture aspects of aggressivity that would be overlooked in traditional quantified risk assessments.

Here, we encounter a first aspect in which the algorithm diverges from the BVC. The focus on words used in nurses' reports seems to attempt to push the boundaries of quantification by spanning more subtle signs of aggressivity and relational causes of violence. However, the way it appraises nurses' reports is crucial. First, the algorithm assumes a linear relationship between the presence of predictive terms (and the behaviour and interactions they refer to) and the increase of violence risk. In other words, whenever the word "colleague" is mentioned in a report, the risk score for that patient automatically goes up. This predictive logic, which it shares with the BVC, causes performance issues for the algorithm. In Emilia's view, based on her perusal of the dataset, "there are people that … quite frankly just … have some kind of one-off explosion of violence." Indeed, this resonates with nurses' experiences of some situations of tensions resolving themselves and others being impossible to predict. Interestingly, in Emilia's view, only the latter ("false negatives") are an issue: what worries her is not being able to identify patients that do become violent, rather than the chance of flagging, and potentially intervening upon, patients that would not have engaged in violence after all.

A second way in which the algorithm's appraisal of risk crucially differs from practices on the ward has to do with its very identification of predictive terms. These terms emerged as "predictive" based on techniques of dataset creation, model training and tuning mobilised by Emilia. If nurses can tweak the BVC to provide a better representation of their doubts, tweaking, (or tuning) happens here in the relation between data scientists and algorithm. Tuning is chiefly a matter of statistical accuracy, referred to the (shaky) ground truth of violence reports, rather than to the situated and embodied relations nurses have with a specific patient. This, of course, matters, in its externality to life on the ward, in that it ends up constituting what counts as risk in a way that is separated from the lived experience of that risk.

It is, of course, impossible to detail what tuning actually does when working with random forest algorithms. What we can, however, look at, are the risk scores the algorithm ended up suggesting throughout the pilot. As Emilia disclosed during the final evaluation meeting, only during those three months, the algorithm had flagged as at risk of violence (≥0.5) more than 500 cases (i.e. patients/day) that had been given a BVC of less than 3. In other words, the algorithm had suggested adopting violence-pre-empting measures on 500 occasions in which nurses would have not initiated any intervention. Risk scores are thus a major aspect in which the BVC diverges from the algorithm. Clearly, this has to do with the way the algorithm is tuned: we have seen how Emilia is worried about false negatives, about the algorithm possibly not picking up on all risk behaviour. The risk of false negatives, thus, drastically outweighs the reality of false positives.

Reinstating doubt in algorithmic predictions

Risk scores were shared with each clinic every morning, before the general handover, in the form of an Excel sheet listing all the clinic's patients' names in one column, ranked from highest to lowest risk, and with their scores (a number from 0 to 1) in the other. Scores higher than 0,5 were considered representative of risk. In Emilia's and the managers' plans, professionals in the two clinics would discuss the aggression scores during the morning handovers, during which nurses share with psychiatrists their reports from the previous day, and discuss which patients should have a psychiatric consult, and whose treatment plans should be changed.

If patients' aggressivity is a major topic of discussion during handovers, it is never a topic tackled by referring to BVC scores. Perhaps unsurprisingly, thus, the algorithmically-produced risks scores were not part of this decision-making throughout the whole pilot. Even in the first days, when the algorithm was still new and exciting for professionals, the knowledge it produced was, apparently, never quite taken seriously by professionals in their deliberations:

> The algorithm only came up towards the end of the meeting. Josh, the nurse, noticed that it was weird that the patient with the 2nd and 4th highest scores were scoring so high. Especially for the 4th one, he commented that "it's just a matter of personality, he spends a lot of time in his room." (fieldnotes)

Although the algorithm was intended to support decision-making, it is interesting to notice how scores actually generate decisions by coming up with targets for intervention (cf. Amoore 2013; Ratner and Elmholdt 2023). If professionals are the ones to finally implement any decision, it is also true that a major way in which algorithmic scores are supposed to contribute to decision-making is by deviating from professional judgements — that is, by flagging target that are not on their radar. However, as emerges from the quote above whenever the algorithm diverged from professional's assessment of a patient's situation, professionals simply assumed it was wrong. Some of them even attempted to theorise *why* the algorithm would assess some patients wrongly:

> Josh explained how nurses are often surprised by some of the patients who had been scored as high-risk. He shared his theory about why: "Sometimes [the algorithm] will score as high-risk someone who really isn't, because it picks up on the "aggressive" in a sentence that actually says "not aggressive."" He showed me the file with the risk scores for that day, comparing it with the printout from the handover. He saw the name of the patient scored as the highest risk: "This one, for example, doesn't make any sense. This is a patient who is asking to be isolated from the group. That has nothing to do with violence!" (fieldnotes)

Investigating whether the algorithm was getting things wrong is beyond our analytical scope. These attempts at reverse-engineering algorithmic reasoning show that professionals consider risk scores on a lower epistemic level than their assessment. Such attitude, which was the only way nurses and psychiatrists engaged with risk scores, starkly rejects algorithmic pre-emption and the assumption that machine learning might pick up on aggressivity *before* nurses.

During the pilot's final evaluation, Carl, the head of one of the clinics reflected that this mode of engagement likely stemmed from the gap between algorithmic and professional ways of assessing aggressivity: "[the algorithm] sees something in the [EHR] notes and [the score] goes up. And when we see it, we think — yeah, but of course, that's not aggressivity." Himself a nurse, Carl questions here the linear connection between words used to describe patients' behaviour and violence risk. First, the words the algorithm considers "predictive" are often generic (e.g. "colleague," or "offered"). Second, what is registered in the EHR is a simplified, de-contextualised version of life on the wards: nurses routinely add a wealth of contextualising details to their account orally during handovers. Since understanding the possible motivations behind a patient's behaviour is central to nurses' techniques around aggressivity, they consider the algorithm's sole reliance on EHR data insufficient to assess aggressivity.

Inspired by Amoore's (2020) articulation of ethics in the wake of machine learning as founded on a reinstating of doubts and alternatives against the certainty of output signals, we read nurses' rejection of the linear logics of prediction and insistence on contextualisation as an ethically-laden attempt at opening up the multiplicities that algorithmic scores reduce to one. The only meaningful way to engage with a score that reduces to a number an issue nurses know to be so complex is then, as we saw Josh do above, trying to understand which relational dynamics the algorithm might be objectivising, which stories it might be ignoring, which alternative futures it might be pre-empting. Nurses are, as it were, reinstating doubt in algorithmic outputs. We elaborate on the theoretical and practical implications of this in our discussion.

Discussion: On doubt, futures, and data

In this paper, we have examined three different ways of relating to and dealing with aggressivity and violence risk: nurses care practices, the BVC (and nurses' interference with it), and the algorithmic risk scores. We have argued that nurses mobilise a kind of doubt that is simultaneously epistemic and ethical. Doubt troubles both the linearity between expressed sign of aggressivity and underlying states experienced by patients, and the opportunity of intervening on these signs (thus potentially escalating situations that might have resolved themselves). Nurses' practices illuminate how deciding entails dealing with a multiplicity of potential alternative accounts, and is, as such, an ethical moment. The probing of alternative explanations is central to professionals', and specifically nurses', practices around aggressivity and violence on acute psychiatric wards. In their interactions with patients, in their use of quantification instruments, and even in their relating to algorithmic risk scores, nurses embody an attitude of opening up possibilities, resisting to linear logics and insisting on a proliferation of possible alternative accounts. Indeed, it is this very opening up of alternatives that makes it possible for them to morally navigate heavily ethically-charged decision-making moments. Because of this doubtful attitude, nurses' anticipatory action tends to be informed by *precaution*: they monitor the situation by attempting to notice warning signs, yet regard intervention as warranted only when it might prevent a *perceived* threat from escalating.

Due to their constant presence on the ward, nurses understand themselves as internal to the world they are appraising and attempting to intervene on. As such, they are aware of their need to account for the ramifications of their actions — or lack thereof. This internality and accountability are progressively displaced with instruments for quantification and risk pre-emption. The logic of precaution also applies to the BVC: although it crystallises life on the ward in strange temporalities, doubt can be effectively incorporated in the final BVC score. Nurses thus meaningfully shape quantitatively-mediated anticipatory action. Conversely, the algorithm imports a pre-emptive orientation that is unrecognisable to nurses. Pre-emption means that the space for doubt is limited to the training process, and contained in the complex algorithmic architecture and its finetuning. Its univocal output, by generating decontextualised risk scores that aim to efface doubt and alternative accounts, intervenes at the level of anticipatory action itself. Although professionals try to reinstate doubt in the score, the univocality of the output leaves them with a dichotomous option to either subscribe to or altogether reject the score. Moreover, the algorithm's *pre-emption* attempts to identify patterns that might be beyond professionals' awareness, thus generating targets for intervention that are not necessarily known (or, as it were, not yet existing).

If other applications of machine learning in acute psychiatry are possible, it is worth here taking seriously the implications of introducing pre-emptive orientations in clinical settings, like this particular algorithm (and others, e.g. Borger et al. 2022) attempts to do. This appears particularly urgent in acute clinical settings in which shortages of trained personnel translate into emphasised need for securitisation. First, algorithms in psychiatry have to face the problem of the limits of calculability (Amoore 2020): modelling objects that, like violence, are known to be volatile and hard to pin down. Even with machine learning, these objects continue to elude predictions (the reader may remember Emilia's comment on some patients showing "explosions" of violence). When the objects of algorithmic outputs are ethically laden, like in the case of violence, we could ask whether it is worth it to try eliminating their inherent difficulty through algorithmisation. In the case examined in this paper, we might argue that the difficulties emerging from this algorithm are not so much technical ones. Even if the algorithm had access to more and more diverse datapoints (e.g. physiological data related to aggressivity), it is its very attempt at addressing, and somehow simplifying, such an ethically-charged decision-making moment that deserves critical scrutiny.

Moreover, noticeably, the pre-emptive orientation introduced by the algorithm examined here reshapes violence as an object of intervention in ways that warrant reflection. On the one hand, the algorithm's pre-emptive orientation enables professionals to imagine a less violent future for acute psychiatry. Nurses and psychiatrists often shared their unease with practices of patient isolation, which they recognised as problematically violent. In this sense, catching threats before they emerge would allow them to phase out this controversial practice. Although pre-emption would likely entail more sedation, we can see that the introduction of risk scores is, at least partially, justified by a will to move acute psychiatry towards more ethical ways of dealing with unstable patients.[15]

On the other hand, algorithms' pre-emptive future orientation is a cause for concern, especially when they are meant to support decision-making in acute psychiatric settings. If the trade-off between sensitivity and specificity (i.e. between false positives and false negatives) is a well-known issue in statistics, here it emerges starkly in its ethical import. Algorithmic risk scores presuppose linear relationships between signs of aggressivity and futures of violence; they recommend action by foreclosing the possibility for alternative interpretations of patients' behaviour, and for differential, doubt-ful ways of weighing particular episodes. Untouched by nurses epistemic and ethical doubts, the algorithm ends up displaying a drastically more punitive logic than nurses

[15] Whether sedation amounts to a less violent measure is a point worth discussing, but beyond the scope of this paper (cf. Dumit 2012).

subscribe to. We can speculate that this might have to do with the model being tuned without nurses' involvement, and that a differential tuning of such a simple algorithm might have produced less strikingly off outputs. But the point stands that, albeit attempting to recast aggressivity as relational and subtle through a focus on words, the algorithm falls dramatically short of considering the maybes, the doubt, the efforts towards contextualising and the search for alternative explanations that are essential parts of nurses' approach to violence. Moreover, its tuning to yield as few false negatives as possible fits with the general idea of security the algorithms is mobilised to achieve. When applied to issues such as violence algorithms' pre-emptive orientations are likely to enhance trajectories of increased securitisation that appear disproportionately punitive towards patients.

Indeed, the punitive logic that we have seen emerging in the final evaluation of the pilot, with risk scores' disproportionate number of "false positives," warrants reflection. As we have seen, integral to the algorithm is an assumption that it should identify risk nurses are not aware of, thus generating new targets for intervention. This is characteristic of pre-emptive approaches that, in the name of security, aim at obliterating threats before they emerge (Amoore 2017; 2020). Punishing (including sedating) patients earlier, and more, appears as the only way that an algorithm can live up to its promise of pre-emption in acute psychiatry. This, as we have shown, represents an attempted shift away from a field of clinical practice where decisions are acknowledged to be situated, tentative and ethically-charged. Other machine-learning approaches to patient and personnel safety in acute psychiatry might be possible. However, as we suggest, they need to take into account, and indeed make central, the doubt at the heart of decision-making in these settings.

In closing, the informal ways through which algorithms with potentially serious implications are slipping from research settings, into pilots, into (potentially) clinical practice, albeit marginal to our analysis here, warrants at least mentioning. Like previous studies of machine learning in clinical practice (Lebovitz 2019; Maiers 2017), our engagement with this pilot has yielded a picture of selective reliance on algorithmic predictions. Nonetheless, it worth taking seriously the work that risk scores such as the ones analysed here might do in reshaping psychiatric practice in the future. If pre-emption is a feature of many applications of machine learning dealing with some sort of clinical risk, particularly in acute settings, machine learning in psychiatry necessarily impinges on behavioural and affective components that make its ethical stakes particularly pronounced. As we suggest, this dynamic, and the ease with which it seeps into practice, warrants careful observation, and perhaps precaution, within the academic community and beyond.

Chapter five
Discussion

Chapter 5

In AI's shadow

I opened this book with the surreal scenes, filmed and commented by Hito Steyerl (2019), of engineers shattering windows in a World War II hangar. Throughout the previous chapters, I have attempted to bring into view the glass-breaking encouraged and presupposed by clinical AI through a meticulous empirical investigation of the objects, subjects, practices, tools and data that AI enacts even before reaching clinical settings. Over and over, we have witnessed tangible reconfigurations of practice being set in motion in the name of an eminently intangible (and, in fact, often absent) technology. Taking these reconfigurations seriously means, as I have endeavoured to do, not only documenting the empirically observable, subtle shifts in professionals' (and non-professionals') practices, but also, crucially, moving beyond the empirical plane to think through the realities such reconfigurations attempt to bring about. In closing this study, then, I turn once again to Steyerl's lessons to think about how to engage critically with a partially absent object.

In her talk *The language of broken glass*, Steyerl (2019) discusses, like I often did in this book, an AI-driven technology that does not yet exist. She does so by humorously speculating on the conditions under which the device her engineers are trying to build would make sense. This thought experiment repurposes a point of Marxist ascendence: in *Capital*, Marx proposed that an archaeological examination of past technologies of production as able to yield insights into the "extinct economic formations of society" (1976: 286; cf. also Pasquinelli 2023). In her provocative thought experiment, Steyerl picks apart AI to deduce the shape of a possible future world in which there might be a market for the particular AI the engineers are developing — a world in which there is a market for a device that can recognise the sound of shattered glass and alert authorities about a break-in. Considering the functionality and perceived necessity of such a device, this would seem to be a world with a lot of burglars and very little police. In turn, this might imply a world of rampant inequality and of defunded state services — one in which, perhaps, private actors (or, in this case, militias) would emerge to fill in the gaps opened up by the gradual withdrawal of the State. In short, Steyerl concludes, the world in which these devices make sense is essentially a kind of "luxury version of a war zone" (2019).

Clearly, this book was not concerned with the politics of privatisation and securitisation that often rely on and usher in AI technologies.[16] However, not unlike Steyerl, by focusing on clinical AI, it has tried to engage with an "emerging, invisible object" (Steyerl 2019). The clinical AI technologies I have traced in my analyses are still (and, as I claim later in this discussion, perhaps will always be) under development. Yet, somehow, they already exist and achieve things in the world: to say it with Steyerl, they "already cast a shadow."

Like Steyerl's talk, my study has begun by registering scenes of *glass-breaking* — instances of sometimes paradoxical organisational change aimed at making clinical practice itself a more suitable ground for training and applying artificial intelligence. Yet, making sense of these scenes, in their often slightly absurd character, has led me increasingly towards a contemplation of *AI's shadow* — that is, the possible organisational futures that contemporary material and discursive changes suggest we are working towards. Observing the matter-of-fact way in which these technologies were repeatedly ushered into clinical practice as the only viable future for a healthcare system in crisis, and regardless of the inconveniences, disruptions, and exclusions glass-breaking inevitably entailed, I have become interested in (and indeed felt it particularly urgent to) trying to speak about *the (future) world that AI technologies, as they are currently mobilised in clinical settings, seem to presuppose*. Pausing to reflect on AI's shadow, thinking through the implications of its imagined workings, has provided me with an opening to critically interrogate the conditions that these technologies would enable preserving or, conversely, might clash with. By embodying imaginaries of privatised security, the sound recognition algorithm that Steyerl picks apart risks participating in reproducing the inequalities and defunding of public services that turn privatisation and securitisation into an obvious necessity. In a similar fashion, this book has sought to piece together the scattered shadows of different clinical AIs to speculate on their potential role in alternatively (re)producing structures, ways of knowing and labour conditions in organisations and healthcare systems at large.

This speculation, in its attempt to make speakable something that not yet (fully) is, might be accused of subscribing to a determinism akin to the one that

[16] Cf. e.g. Burton (2023) and Molnar (2021).

I have identified as a major issue with clinical AI's approaches to future spaces (cf. e.g. chapter four). However, as I hope has emerged from my analyses thus far, I aim to impose no ontologically-closed claim on the future of AI in clinical settings. My attempt to lay out the logics immanent to the (imagined) functioning of the AI technologies I have examined in this book is to be read as a thought experiment that seeks to prove generative. My aim is to make tangible the ways in which organisational futures are being rethought *now*, and the aims towards which technological affordances are being mobilised. The analytical points that I have made throughout this study, and that I will synthetise in this discussion, generally amount to an attempt to resurfacing some of the organisational conditions that are silently reproduced, and the aspects of work that are concealed or obliterated by clinical AI in its current (i.e. expected, or prototyped) forms. The speculative character of this analysis, as I shall elaborate below, is also somewhat moderated by a meticulous engagement with the current organisational and labour realities, which provides it with an empirically-grounded contextualisation.

I am convinced that foregrounding, tracing, and making AI's shadow speakable is crucial if we want to achieve just futures — ones that are different from the at-times bleak ones this book articulates. To this end, in this final chapter, I condense the lessons on clinical AI that have emerged throughout my case studies into answers to my research questions, and lay out a few theoretical, methodological and practical implications stemming from my work.

Answering (slightly adapted) research questions

At the start of this study, I broke down an overarching research question,

How does AI manifest in contemporary clinical practice, and what implications does this bear for the organisation and practice of care provision?

into three sub-questions to address across my cases. Perhaps unsurprisingly, as the study proceeded, and particularly in the wake of first-hand encounters with glass-breaking in clinical settings, some of the questions I set out to answer have needed some slight adjustments in order to better illuminate crucial tensions emerging in practice. Moreover, as explained at the start of this discussion, new questions worth asking have come up. In this section, I thus turn to the onerous task of synthetising the answers this study has provided, and signal any shifts from the original questions' formulation (by bracketing them).

1. How are (good) data produced in clinical settings?

A focus on clinical AI's glass-breaking has, quite unsurprisingly, revealed that AI and datafication are closely entangled phenomena. However, this study has highlighted how what is at stake in organisational reconfigurations set off for and by AI is not simply the production of data, but the production of *good* data. Indeed, AI technologies rely on data to be trained, tested, and run, and that data quality translates into the performance of AI technologies, as testified by quasi-truisms such as "garbage in, garbage out" is essentially a truism. Across these chapters, I have shown how clinical data is always produced at the intersection of bodies and technologies — or, more consistently with my theoretical stance here, results from the intra-action of bodies that are either "fleshy or otherwise" (cf. Garcia 2020). Both in the pathology department and in the ICU, imaging and numerical data are extracted from bodies through sensors (chapter three) or through the very excision and manipulation of flesh (chapter two). This means that producing "good" data is a matter of work, of carefully attending to the encounters of bodies and technologies.

In chapter one, we have seen how the critical approach underpinning much of the sociological literature on the digitalisation of healthcare work tends to foreground questions concerning new technologies' uneven consequences, as they are distributed along the faultiness of professional hierarchy. This unevenness often has to do with the proliferation of extra, hidden and often non-meaningful tasks for some — as the work of others becomes more streamlined. In a similar fashion, this study started off by assuming that a proliferation of similar tasks, grouped under the concept of *data work*, would prove an essential component of glass-breaking, and that what appear as instances of automation need be considered in conjunction with the work they inevitably conceal. The connection between data, the work needed to generate them, and the AI technologies that rely on them was thus already at the very heart of my research design.

The first contribution of this study to current discussions of data work in clinical practice lies in an articulation of how the work of (some) humans directly impacts the *quality* of clinical data, and thus the very possibility of introducing AI into clinical settings. Indeed, as I have proposed across chapters, *the material conditions for data work in the clinic need to be considered in close conjunction with the quality of the data produced*. This bears implications for how we are to think about, first, the materialities of data work and, second, what "good" data means in the clinic. Bringing together findings from across my case studies enables me to explore these dimensions in turn.

First, **data work is sustained by attunement**, which I have described as an embodied, sensory and cognitive engagement that draws on contextual and idiosyncratic knowledge. Data work in clinical settings is about coordinating the multidirectional relations amongst bodies and technologies involved in care provision. The concept of *attunement* aims at foregrounding the embodied dispositions professionals develop in processes of highly datafied care provision. In the ICU examined in chapter three, we have seen how care provision entails an ongoing and simultaneous attuning of not only technologies and patients' bodies, but of professionals' bodies themselves. We have seen how ICU nurses constantly attend to sensors and their tracing, by repositioning the former and scrutinising the latter, to stave off the possibility of artifacts making real time data unreliable; how this reliability also entails constantly readjusting patients' bodies to make them more perceptible to sensors; and how all of this is orchestrated through careful looking, listening and feeling. Chapter two examined pathology, a specialty that is similarly data-intensive, this time not because data are a crucial way to get insight into unstable bodies, but because bodies are virtually absent from the clinic. Here, once again, the production of (good) data relied on an alignment between technologies and (parts of) patients' bodies, and this alignment was achieved through embodied and sensorial practices — from assistant pathologists in the grossing room touching samples to cut them properly, to lab technicians stretching tissues to cut even slices and checking whether they had been evenly mounted on slides and stained correctly. The pathology case also showed us how the attunement scaffolding clinical data work has to do with the contextual knowledge developed as a body in a specific organisational setting: indeed, the scanning secretaries managed to see "better" than scanners because they were aware of local professionals' practices. Thus, instead of being confused by circled pieces of tissues, they understood they indicated areas that deserved more attention, and thus sharper focus. Attunement is thus a matter of situated knowledge as much as of embodiment, and is crucial for the performance of data work.

A second contribution to the conceptualisation of data work, as emerging from this study, is that **data work in clinical practice is, chiefly, a matter of making care practices machine-readable**. Interrogating data work, and what it increasingly tries to achieve, also makes it possible to foreground machine readability as a major driver of AI-related organisational and professional reconfiguration. If Steyerl's engineers shattered windows over and over so that algorithms would have a diverse enough database of glass-breaking sounds to analyse, a huge part of the glass-breaking I have observed in contemporary clinical settings had to do with enabling machines of various kinds to perceive various aspects of clinical practice. Thus, in chapter two, pathology lab technicians had to learn a new way of stretching tissues when mounting them on slides in order for them to be more legible to scanners. In the ICU, nurses' careful, embodied work enables manipulating patients' bodies in a way that makes them perceptible to sensors. In psychiatric wards in chapter four, recording patients' behaviour and possible incidents starts to morph from a question of institutional memory to one of machine readability[17] — that is, providing the algorithm with a better outcome measure. Even for data that was available before clinical AI's glass-breaking, acquiring machine readability as a central aim brings about stricter requirements in terms of what data needs to be.

Now, as mentioned in opening this book, different kinds of data (digital and not) have long been part and parcel of care provision. However, what we see happen in the current moment, and what indeed makes data work a component of glass-breaking, is that clinical data are increasingly been thought of as something that can be mobilised for machine learning. This

[17] Of course, and especially in this last case, a great deal of what patients' bodies do and are necessarily escapes the scope of machine readability. Data, as much as it can be of good quality (Leonelli 2012), is always a representation, a proxy, and as such a reduced, machine-metabolised version of reality (Pink et al. 2018). This needs not be an issue as long as represented reality is not what we base our care provision on. I will return to this point later.

means that machine readability, that is, data's suitedness for machine learning purposes, becomes an increasingly central component in assessing what *good* data means. In other words, *machine readability imposes a specific idea of data quality*. We have seen this with pathology slides, where tissue needs to be cut evenly and mounted much more cleanly and precisely, lest the scanner rejects it or fails to focus on it; we have seen this with the predictive algorithm in psychiatric wards, where the incompleteness of data on incidents translated into a poorly-performing algorithm. Glass-breaking thus disproportionately attempts to produce data that is of good quality *for AI*. This mean data that is as complete and standardised as possible. In agential realist terms, AI is always included in agential cuts as the prime epistemic subject in relation to data as enacted epistemic object. Subsequently, data work can be thought of as *the work necessary to make clinical practice machine-readable*. This conceptualisation resonates with points previously made in the literature. For instance, Hogle (2016) has noticed out the circular rationalities at the heart of datafication: it is necessary to transform healthcare in a way that accommodates data capture in order for data capture to be able to deliver on its promise of transforming healthcare.

Because of the requirements that glass-breaking imposes onto practice, the necessity of this data work is never doubted. In other words, because the advent of clinical AI is considered inevitable, implementing the necessary steps to prepare for this by achieving extensive machine readability of clinical practice is hardly up for discussion. However, *data work is not performed by all bodies equally*: in the fleshy realm, it is relegated to bodies that, because of their position in professional hierarchies, have less autonomy in choosing their tasks, or less freedom to delegate unwanted ones (e.g. Burri 2008; Maslen 2017; Mort et al. 2003). Not incidentally, it is nurses in both the ICU and the psychiatric clinic, lab technicians and secretaries in the pathology lab, who are asked to adapt their practices to prepare the clinic for clinical AI. It is nurses, technicians and secretaries who need to break glass over and over so that AI can start to perceive clinical practice.

Thinking of data work in the clinic as aimed to achieving machine readability also surfaces another point of tension — namely, that data quality in the context of care provision might have a different meaning, and different requirements, than data quality in the context of machine learning. This work has shown how data that are “good” for care provision are data that **enable taking responsibility for the care provided**. In the pathology case, we have seen how digitised data started to be plagued not only by issues of sharpness, but also of reliability (i.e. digital slides’ completeness) when the data work performed by the scanning secretary was automated. This data work targets a different kind of *quality as responsibility-enabling*. In turn, it enables to sustain the representationalism on which datafied knowledge-making (and thus diagnosis and care provision) is predicated. If data are to offer a representation of a specific phenomenon (e.g. a pathogenic process in a patient’s body), responsibility-enabling data work entails taking responsibility for the accuracy and reliability of that representation. As we have seen in the ICU, guaranteeing ongoing accuracy is of utmost importance, especially when data is used to constantly monitor patients’ stability and to flag sudden deteriorations. Without a guarantee of accuracy, professionals struggle to provide a diagnosis (chapter two), leave patients’ bedside (chapter three), or take algorithmic predictions seriously (chapter four). This kind of data quality, crucial for care practices, but less so for AI, is what necessitates a type of data work rooted in attunement. For data to be mobilised in the context of care practices, there needs to be some kind of warranty that, albeit a representation of reality, data is a faithful representation nonetheless. As a third, and final contribution to discussions of data work, this work suggests that different conceptions of data quality (standardisation versus accuracy), tied to different aims of data (machine learning versus care provision) might translate into different types of data work (data compiling versus attunement).

To sum up, data is produced in clinical settings through the encounter of fleshy and non-fleshy bodies. However, *good* data is produced by carefully attending to, and coordinating, such encounters. Producing good data, that is, data that both enables care provision and can be mobilised for AI’s data analytics, requires a professionalisation of data work — that is, data work needs to be turned into an organizationally-valued task, and to which necessary time is allocated.

2. **What objects for organisational and professional intervention are created or reconfigured by clinical AI and its attendant glass-breaking, and with what implications for professional knowledge practices?**

Clinical AI intervenes in professional and organisational practices by enacting new or reconfigured epistemic objects. These objects vary depending on AI's distance from the clinic, as well as on the purpose of the AI technology examined. In the pathology department, where AI was the farthest away from clinical practice, we have encountered the purest cases of glass-breaking, where the clinic is reconfigured to progressively become a space of model training and application. In this case, the main epistemic reconfiguration consisted of the introduction of tangibly new epistemic objects: digital slides. As we have seen, digital slides' epistemic affordances push pathologists' knowledge-making towards the more quantitative aspects of digital images (cells' measurements, amounts). This is in stark contrast with the knowledge-making supported by glass slides, which relies on cellular structures' colours and shapes.

As AI moves closer to the clinic, and we can trace the contours of specific technologies, as we did in the ICU and psychiatry cases, AI-supported categories of patients (we could say, "stable patients" in the ICU case, and "non-aggressive patients" in the psychiatry one) attempt to make their way into professionals' care practices. In both cases, patients are made into flat and quantified objects: patients' health state is replaced by a colour in the ICU, and aggressivity by a number in the psychiatric clinic; in both cases, the quantified output of data analytics replaces patients' bodies and stories, that is, the qualitative aspects that would support knowledge-making in pre-existing practices.

Moving on to the implications of these quantified epistemic objects, it is interesting to notice how the objects enacted by clinical AI technologies (i.e. categories of "(un)stable" and "(un)aggressive" patients) seem to *exist at the intersection of patient care and resource allocation*. Indeed, patients' health data is mobilised to create categories of patients that, at least temporarily, are less in need of care. This is part of an effort in meeting the joint, if somewhat clashing, goals of patient care and efficient utilisation of scarce human resources. However, and in light of the answer to sub-question 1, this imperative of efficiency presents two issues.

First, **efficiency is at odds with data quality**. As we have begun to see, foregrounding issues of *data work* provides an entry point for understanding a crucial tension at the heart of contemporary clinical AI projects. On the one hand, as articulated above, data work constitutes the human and work-related component of glass-breaking: it opens up possibilities for AI in clinical practice by producing reliable data and making care practices machine-readable. On the other hand, data work does, in many cases,[18] constitute extra work introduced for the sake of clinical AI, thus clashing with narratives casting (AI) technology as labour-saving. Finally, in many cases,[19] it is perceived as tedious work by the ones who have to perform it. The fact that *data work is both inefficient and often experienced as tedious* creates a fertile ground for narratives promising to get rid of it, so that plans to automate it often encounter workers' favour (cf. chapter four). The issue here, however, becomes somewhat circular. As discussed in chapter two with the concept of *fauxtomation*, narratives of automation rely precisely on concealing undervalued labour. Tasks around data work are perhaps increasingly *de-professionalised*, as we saw in the case of secretaries, while hopes for a presumed "real" automation of tedious work is continuously displaced to a more or less distant future.

A second problem with the imperative of efficiency is that **contemporary instantiations of clinical AI promise to achieve efficiency through a redirection of professionals' attention**. The two chapters that addressed more concrete instantiations of clinical AI (chapters three and four) suggest that, especially when targeting nurses (who, it should be noted, are the professional group whose scarcity is felt the most after Covid-19), these technologies are aimed at performing significant algorithmic management (Kellogg et al. 2022). In a nutshell, they are supposed to help dealing with workforce shortages by making nurses' work more efficient or automating their tasks. Although this situation might shift in future instantiations of clinical AI, what we observe in the current moment is that nurses' work tends not to be taken over by a machine that performs it in their stead. Rather, it tends to be *redirected away* from tasks that are deemed, in a particular moment, unworthy of their time. Noticeably, then, the question appears to be not so much which tasks can

[18] I.e. when it is not integral to pre-existing care practices, such as, for instance, in the case of ICU nurses adjusting sensors (cf. chapter three).
[19] Again, from the cases presented here it would emerge data quality work becomes tedious, as a minimal conditions, when it is decoupled from patient care (cf. also Høyer 2023). High repetitiveness of tasks involved might also contribute to this (think of the scanning secretary in chapter two, or of the psychiatric nurses refusing to score BVCs and to report all violence incidents in chapter four).

be performed by technologies, as classical automation debates, but indeed which tasks can be left (at least temporarily) unperformed.

The implication here is one of epistemic resources and of epistemic politics: if professionals' attention is redirected following algorithmically-identified targets, we need to ask questions about how, and based on what, those targets are identified, and about the targets that might start falling out of their scope. The efficiency AI technologies are mobilised to achieve can only be realised if they do manage to redirect nurses' attention away from (care) objects that algorithms themselves surface as not worthy of their time (i.e. "stable" ICU patients and "unaggressive" psychiatric patients). As should be clear by now, "worthy" becomes here a function not only of complex calculations, but also, emphatically, of an output signal that has been considered organisationally feasible, and of what has been made legible to algorithms. This is in obvious contrast to the practices through which professionals themselves establish their priorities and structure their work days, in which professional autonomy, as well as broader considerations of patients', colleagues' (and technologies') current and upcoming needs play a central role.

We can thus see that the potential negative implications for professional epistemic practices do not stem so much from AI's enacted epistemic objects per se, but rather from the efficiency AI technologies and their epistemic objects are mobilised to achieve. To reiterate, these implications consist, first, of the likely obliteration of data work and the subsequent plummeting of data quality as required in practices of care provision. A second implication lies in the redirection of professional attention these epistemic objects encourage, which might take resources away from other, less machine-readable, tasks and goals. Whether the epistemic objects enacted by clinical AI are taken up in practice, and whether the knowledge-making practices of professionals are effectively reconfigured is, however, a different question. It is not clear that nurses would stop performing tasks identified as nonurgent, were they to engage with clinical AI technologies, were these technologies first-hand. This leads me to issues of use.

3. What does this tell us about the mechanisms underlying how professionals (do not) embed AI in their work?

In the three cases examined in this book, professionals have managed to circumvent or resist some of the reconfigurations clinical AI and its modes of glass-breaking seemed to require. In the pathology department, despite worrying that the distribution of glass slides may one day cease, pathologists continued to turn to their microscopes, especially when dealing with complex cases. In the psychiatric clinics, nurses and psychiatrists never mobilised AI-generated risk scores as part of their deliberations. Moreover, instead of taking algorithmic scores seriously, they simply ruled them out as wrong whichever prediction deviated from their own judgement. Finally, the ICU nurses in chapter three requested to have the possibility to determine themselves whether their patient was stable. This claim to autonomy would defeat the purpose of even mobilising data analytics, thus indicating some likelihood of resistance to the dashboard's script of attention redirection on their part.

Across cases, we have thus encountered instances of resistance to technological scripts that border more or less strongly on non-use. Granted, these instances of resistance might be temporary, and are always only accessible to some. For instance, although pathologists barely engage with digital slides, technicians and secretaries have their work more or less radically reconfigured in order to produce them. Moreover, it is likely that pathologists' non-use would decline if glass slides were no longer being distributed. Similarly, if the management of the psychiatric organisation in chapter four put more emphasis on the use of algorithmic scores, nurses would, at the very least, have to adapt their registration practices to the algorithm's needs (e.g. reporting incidents more thoroughly). Moreover, and as one of the local psychiatrists was keen to point out, the lack of use of algorithmic risk scores might be a symptom of staff perceiving the scores themselves as temporary due to them being part of a pilot. As such, they might be further incorporated in decision-making in further stages of implementation.

If this point deserves further reflection, for not it will suffice to point out how, across chapters, resistance and non-use have shown considerable analytical potential.[20] They have provided openings to analyse *ethical frictions and disruptions in knowledge-making practices that AI and its modes of glass-breaking (threaten to) produce*. This analytical stance, of course, resonates with research agendas centred on (non-)use, and articulated in both STS (Kiran, Oudshoorn and Verbeek 2015; Oudshoorn 2019; Wyatt 2003; Wyatt

[20] This point bears methodological implications on which I shall return later.

et al. 2002) and in human-computer interaction literature (e.g. Satchell and Dourish 2009). In particular, I was inspired by arguments for foregrounding non-use as an "opportunity to understand the tacit, moral routines threatened by the introduction of new technologies" (Oudshoorn 2019: 171). In my analyses, I have proposed that theses threatened moral routines should be considered in conjunction with epistemic routines. This is an insight I derive from agential realism's points around the necessity to take responsibility for the knowledge and the worlds in the enactment of which one is involved. Moreover, already in the review in chapter one, knowledge-related questions emerged as a prominent theme in the wake of the digitalisation of healthcare work. Especially sociotechnical and critical approaches, typical of STS and sociological literatures, foregrounded the question of what happens to pre-existing knowledge practices when new knowledge-mediating technology is introduced, and how information that falls outside the scope of what these technologies' scripts consider relevant is still made (or, alternatively, fails) to circulate.

In this book, I have thus interrogated non-use as symptomatic of ethical frictions and disruptions in knowledge-making practices. Specifically, I have considered how organisational structures remade or reinstated through the introduction of AI to clinical settings give rise to *exclusions and inclusions in knowledge-making*. As I have argued, we should not only ask what kind of world data analytics enact, but rather what kind of decision-making AI outputs afford. This has to do with the aura of certainty surrounding algorithmic outputs (Amoore 2020). Indeed, all the technologies that appear in this study operate, at some point, an agential cut (Barad 2003) that both *excludes* professionals and aims to present them with some kind of *certainty*: certainty that the relevant features are already visible on a digital pathology slide (chapter two); that the task recommended by a dashboard is the one that most deserve a ICU nurse's attention (chapter three); that a psychiatric patient is or is not aggressive (chapter four).

In a way, the certainty inherent to algorithms' output is a function of the very exclusion of professionals, in that it is predicated on the obliteration of all the alternative outputs excluded by the algorithm in its learning process. However, because of the exclusion from the agential cuts involved in defining objects for intervention (i.e. digital slides, stable or unstable ICU patients, psychiatric patients at low or high risk of violence), as I have argued, epistemic disruptions translate into *ethical issues* for the professionals tasked with intervening on these objects — specifically, issues that have to do with taking responsibility for the care they provide. This is, of course, a point that resonates strongly with agential realism, according to which it is possible, and indeed ethically necessary, to take responsibility for the intra-actions one is *involved in*.

Throughout my analysis, I have identified *doubt* as an essential component of what makes intra-actions more or less suited for ethical knowledge-making. Concretely, the objects I have examined as enacted in the context of glass-breaking are objects the properties of which have already been stabilised in previous intra-actions that did not include professionals themselves. That is, digital slides have already been put in focus and illuminated by a scanner, thus making only some feature of the tissue visible; ICU patients' data have been analysed and translated into a colour indicating "stability" by real-time analytics; and the predictive algorithm in chapter four has decided which words and which events matter in terms of assessing violence risk, and which should be ignored. These epistemic objects, as central as they are expected to be to professionals' decision-making, operate a closure: they do not enable them to probe *what else may be there*, what might have been ignored, which other accounts might be possible. Opposite to algorithmic outputs' aura of certainty, *doubting data* has emerged as crucial in professionals' ethical decision-making: pathologists, guided by the imperative "to be sure" and "not to miss anything" are seldom content with only looking at one layer of a slide under normal light (chapter two); ICU nurses never take for granted the data that they see on screens, but actively investigate and probe the constellations of bodies and machines underpinning their production (chapter three); psychiatric nurses always try to understand whether a higher BVC or algorithmic risk score registered for some of their patients might be justified under the circumstances (chapter four). Registering data is not enough: doubting them in order to understand them, and to probe alternative stories underlying them, is a major way in which professionals relate to it. Crucially, it is not a way of relating that AI supports or even enables.

Returning to resistance and non-use, this study has thus made two contributions. First, it has pointed out the *ethico-epistemic roots of non-use*, showing that a lack of engagement with algorithmic technologies can stem from moral qualms, especially in the context of momentous and ethically-laden decisions. Second, and related, it has begun to postulate the necessity for professionals to be meaningfully included in the agential cuts facilitated by clinical AI in order to be able to take responsibility for the knowledge produced through such cuts. Based on insights emerging from the empirical material collected here, I am now able to specify *the possibility of exerting a doubtful attitude* as a crucial mechanism through which such inclusion is made meaningful in datafied clinical settings. This point deserves further consideration in light of efficiency-geared organisational ambitions that I have registered across cases (cf. also above). This doubtful attitude is inefficient, and clashes with imperative of cost-cutting pervading current healthcare discourse. However, as I suggest, focusing on the ability of exerting doubt gives us a potentially deeper insight into possible underlying causes of what is usually discussed as "resistance" to innovation.

4. (Which organisational and professional future is clinical AI mobilised to achieve?)

At the risk of deviating from academic conventions, I dwell briefly on an additional, unforeseen sub-question that exceeds the initial scope of this study. As I have learned in my engagements with clinical settings, an analysis that stays close to the empirically-observable reconfigurations set in motion by (the expectation of) clinical AI fails to entirely capture its performativity. If we are to fully appreciate the ways in which "AI manifests in contemporary clinical practice," as my overarching question begs, we need to broaden the scope of our analysis to the organisational conditions and professional practices that clinical AI might enable, reinforce, or hinder. What I am getting at here is the interplay between technological affordances and organisational issues that shapes current and future clinical applications of AI. What AI is or will be mobilised to achieve in healthcare organisations is not only a question of what AI *can do*, but also of which issues an organisation considers solvable *and* worthy of investment. I am proposing that, by mobilising AI for a specific aim, healthcare organisations express an *intention to bring about a specific version of the future*, in which certain values and goals prevail over others. Unpacking clinical AI's performativity, thus, becomes crucially a matter of unearthing the organisational structures and labour conditions that it promises to create, with a keen eye for the issues it leaves unaddressed. Going back to both Steyerl (2019) and Marx (1976), what can a technology tell us about the sociotechnical worlds in which it is supposed to function?

Chapters three and four have offered some hints as to these future worlds. Both in the ICUs and in the psychiatric clinics, AI existed in an organisational context of resource scarcity, where investing in cutting-edge technologies was considered more feasible than investing in human resources. Indeed, the AI technologies we examined in those chapters were mobilised more or less explicitly to deal with workforce shortages. However, dealing with workforce shortages through these technologies meant trying to intensify the labour of available professionals, thus increasing ward capacity without expanding the workforce. Indeed, these technologies only make sense in a world in which *efficiency supersedes other values and becomes the main aspect of quality of care*.[21]

In the ICU, this world of efficiency is in tension with data reliability. A world in which ICU nurses work as efficiently as possible thanks to real-time analytics is a world in which they have no time to check for artifacts in displayed data, to make sure that all sensors on patients' bodies are working properly — in other words, a world where the doubt that is so essential to care provision wanes. However, as we have seen, the real-time analytics enabling efficiency in the first place can only function if data are reliable, that is, if nurses have time slow down to perform maintenance on the apparatus producing them. This AI might thus seem to make sense only in a future in which nurses are even more overworked, having to care for even more patients than they could before. In this future, being an ICU nurse would become essentially a matter of intervening on unstable patients. This would arguably leave very little time not only to ensure data reliability, but also to attend to patients' emotional needs — *pace* narratives of automation as bringing back the "human dimension" to healthcare work (cf. chapter four).

In the psychiatric clinic, the world of efficiency AI is mobilised to achieve comes at the cost of increased organisationally-sanctioned and chemically-

[21] Quality of care does include efficiency as one of its core components in many definitions. According to the Institute of Medicine (2001), other crucial components are safety, effectiveness, patient-centredness, timeliness, and equitability.

achieved violence. In this future, the organisation can deal with risk despite not hiring more psychiatric nurses trained to recognise and deal with violence. Moreover, it can end practices of patient isolation by relying on risk scores that suggest intervening at a much lower risk threshold. As we have seen in chapter four, effectively using these risk scores would mandate conversations about increasing medication dosages for patients that nurses currently do not consider high-risk. Moving away from current wait-and-see attitudes, which still leave open some possibilities for escalation, the only way to pre-empt all risk of violence (which still, of course, leaves open other types of risk) would be to make all patients effectively inoffensive through increased, and earlier, sedation.

Addressing the main research question: Ontologies of partial absence

In returning to my overarching research question, and in starting to move towards the theoretical take-aways of this book, I propose to think about **AI as displaying ontologies of partial absence**. This notion seeks to get to the ways in which AI manifests as never fully absent when it is not yet in clinical settings, but also as never fully present even when algorithms are introduced in organisations. I flesh out these ontologies below.

I offer the notion of **ontology of *not-yet*** to describe instances where AI appears to loom in a more or less distance future, yet it justifies practical readjustments in its expectation. In its ontology of not-yet, AI's absence is considered a matter of time, and its presence is restricted to practices of preparation. We have seen this in the pathology lab, where all we could see was AI's glass-breaking. Here, AI was a necessary logical proposition without which the reduced quality of digital slides would make no sense. AI was, consistently, *not-there-yet*, and yet it already manifested in the lab technicians' adjusted methods of mounting slides so as not to confuse scanners with folds in the tissue; it was in the scanning secretary being moved to a small, windowless room and attending to buzzing machines 8 hours a day; it was in the adjusted workflow and in the scanners' digitised slides; it was in the pathologists trying and often failing to find advantages to the new way of working being pushed upon them; it was, more obviously, in the tendering of always newer and more sophisticated scanners ordered all the way from Japan, and in the frequent presentations of AI-assisted image analysis given to the department's digitalisation team.

I suggest that, in instances where clinical AI is being designed and its implementation is being planned, AI morphs into an ontology of *elsewhere* — at least when looked at from the clinic. In these instances, AI's differential absence or presence eminently a matter of politics, of hierarchy-driven inclusions and exclusions. As we saw in chapter three, AI was absent from ICU rooms. It was present, at least partially, in a different building, where small lectures about technological fixes were given to nurses, and where PhDs would sit in a fancily decorated office training algorithms and writing papers about them. Though certainly not a surprising point, we found AI to be the rooms where consortia are formed to apply for grants to bring AI to the clinic. These are rooms to which nurses, who spend more time in the ICU, touching patients and looking at their vital signs' monitors, were rarely invited. And, in the rare cases in which they were invited, they rarely felt to be in the position to make meaningful contributions. The distance between the work they perform on bodies and on data and the AI technologies that would be unimaginable without this very work was left unbridged. Chapter three has shown how glass-breaking is unevenly distributed, with some involved in planning it, others facing the prospects of merely having to execute it.

Finally, I propose thinking about instances where AI is initially introduced in the clinic, but in the form of pilots and experiments, as informed by AI's ontology of *for-now.* In these instances, AI is present and actively working in the clinic, yet the seemingly temporary nature of its introduction, as well as its continuous operating out of sight, bears implications for how other actors relate to it. In these instances, AI's absence is a matter of interstitiality and potential imminent disappearance, while its presence is a matter of testing a prototype. The acute psychiatry clinics in chapter four are the only place where clinical AI technically was "in practice." Even there, however, it was present in the context of a pilot, as a prototype being tested in its technical and organisational performance. Although the algorithmically-generated scores would appear every morning on the clinics' mailing lists, their presence was tenuous by virtue of their being "tested out." They were always, as it were, on the verge of disappearing. Professionals often forgot that these scores even existed until the ethnographer's or the data scientists' presence made them into a matter of concern again. When professionals did think about them, their existence in a temporary and experimental space sheltered them from sustained critical scrutiny: they were but a prototype amenable to improvement, thus their getting things wrong should not alarm anyone. However, as chapter four showed us how, even when relegated to an interstitial existence, out of sight and discussions, clinical AI exists nonetheless: the algorithm continuously and untiringly analysed nurses' and doctors' clinical notes, producing risk scores for each patient. As often noted by both the data scientist and the professionals involved in the pilot, these risk scores could be mobilised, for instance, as a source of external accountability. This last ontology of partial absence bears relevant theoretical implication, which I turn to in the next section.

Theoretical implications and avenues for research

Experimental world-making: Absences and glass-breaking

The first implication of my analysis in this book is that, in studying AI's ingression into clinical practice, it is necessary to take experiments seriously in their world-making capacity. Future research endeavours should thus **problematise and fully flesh out current experimental modes of AI governance within organisations**. This means, in my view, two things. First, that given the *experimental nature of AI* as a technology, we need to rethink the notion of pilot both as offering naturalistic "real world" settings for testing technology, and as clearly temporarily delimited. Second, that we should interrogate *experiments as a mode of governance for clinical AI* that introduces a tension between innovation and quality and safety of care. I will articulate these points in turn.

To my first point, the research presented in this book confirms the current tendency in clinical settings towards experimental modes of technology introduction — previously referred to as "pilotitis" (Egermark et al. 2022). In the case of AI, these experiments are aimed at finding out whether a specific AI technology performs satisfactorily in clinical settings (commonly referred to as "real world settings"). Given the kinds of reconfigurations I have examined in this book as glass-breaking, it is worth it to dwell on what this "real world" actually entails in this context. I am supported here by Noortje Marres's recent discussion of public tests of self-driving cars, which she described as reliant on the staging of "highly artificial situation[s]" (Marres and Sormani 2023). As Marres argues, "the introduction of "AI" into society entails modifications of environment in society, modifications which … trouble and to a degree undermine the very distinction between "artefact" and "environment."" In other words, as I have argued at length, a test, or a pilot, requires preparation, and this preparation entails reconfiguration — glass-breaking, as I have called it, or, with Marres, environmental modification. This has implications both for how "real" clinical settings can be considered when they are turned into a testbed for AI, and for what we are to consider the start of the pilot. Arguably, testing is already happening before the technology is introduced, and it should be considered to span the glass-breaking "(e.g. materialities, infrastructures, but also sensing and care practices) that is necessary to accomplish 'artificial "intelligence"'" (Marres and Sormani 2023).

Decreeing the end of a pilot is as (if not more) problematic as establishing its beginning. This has to do with the experimental nature of AI technologies themselves. As Amoore (2020) argues, AI technologies "engage experimentally with the world" (12) in that they rely on an ongoing, productive incorporation of doubt and error as a means to refine their outputs (cf. also Parisi 2019). AI's parameters and weightings are amenable to continuous optimisation in order to produce outputs that are increasingly more accurate. AI technologies are open to change potentially *ad infinituum*, since emerging instance of "real world" might provide occasions for refinement that approximate the variety of reality itself. Therefore, it is worth asking whether AI itself could ever be anything else than a prototype. So-called smart technologies, as Halpern and Mitchell (2022) have shown, are embedded, and indeed call into existence, zones of experimentation that cast the present as an infinitely optimisable "test" of a future always to be deferred. Pilots of clinical AI technology, with their assumption that possible failure can be constantly incorporated as a means of continuous refinement, are no different. The notion of AI's *ontology of for-now*, which I have offered above, gets to the core of the very unstable nature of this technology. This has consequences, of course, for how we think of pilots, and of the very possibility for a pilot's end. Thinking through AI's ontology of for-now, we can begin to see how AI's architecture is entangled with experimental modes of technology governance such as pilots. This begs the question of whether non-experimental modes of AI's presence in clinical settings might even be possible and, if not, what new modes of relating to pilots might be necessary in organisations.

Finally, it is worth stressing how the experimental spaces through which AI is entering clinical settings (such as pilots) are likely to introduce, or intensify, tensions between emerging logics of innovation and pre-existing mandates around quality and safety in healthcare. Both AI technologies' needs and emerging cultural dynamics (Halpern and Mitchell 2022) increasingly cast innovation as an inherently unstable space of experimentation in which rules can be suspended and failures can be harnessed as generative. Granted, in the realm of healthcare quality and safety, a similar movement towards ongoing learning has animated the Safety-I versus Safety-II discussion. In this context, Safety-II perspectives already exemplify an ambition for learning from errors and ongoing optimisation that arguably resembles current AI epistemologies. Regardless, in intervening in decision-making, AI technologies introduce risks to safety (Challen et al. 2019) and quality of care,[22] as I have shown in this book. In response to this, experimentation in clinical and other settings is increasingly being institutionalised, for instance through the recent introduction of "regulatory sandboxes" for AI development in the EU (Ranchordas 2021). Foregrounding the institutionalisation of experimental modes of technology governance, as well as the tensions with quality and safety of care this dynamic introduces, promises to be a fruitful avenue for future research (e.g. Taylor 2024; Van de Sande, personal communication).

Doubt and responsibility: Agential realism's implications for AI ethics

A second implication of my work is that **current discussions around AI ethics should consider ethical and epistemic issues jointly**. Though I derive this point from agential realism, it is also true that agential realism does not bring forward a clear conceptualisation of its ethical stance. The ethics aspect of ethico-onto-epistemology is mostly discussed normatively, and in terms of "taking responsibility" for knowledge produced as the result of intra-actions. Based on my analysis, and as already introduced above, I propose that a focus

[22] Alongside many other risks and ethical concerns, cf. Mittelstadt et al. 2016.

on *doubt* can function to operationalise the ethical component of agential realism in datafied clinical settings. This, in turn, brings forth a *de-centring of AI* that might provide a fresh perspective to discussions of AI ethics that are often removed from the field, with significant practical implications. I will tackle these points in turn, starting with the latter.

Guidelines for AI development and implementation, whether developed in philosophical or legal circles, have by now been thoroughly criticised for their lack of impact and for the window-dressing they enable on the part of science and industry (e.g. Hagendorff 2020). As a response, frameworks and roadmaps are increasingly being developed to help developers, managers and practitioners think through ethical challenges arising at different stages of development and introduction of AI technologies (e.g. Mökander and Floridi 2021). These frameworks, and the discussions underpinning them, mostly focus on two dimensions: AI's epistemology (i.e. the processes by which outputs are generated) and AI ethics (defined narrowly in terms of fairness of and bias in algorithms' outputs; cf. Russo, Schliesser and Wegemans 2023). Even critiques of this bifurcation underpinning current debates often result in a set of propositions that, albeit far-ranging and sophisticated, remain abstract and end up relying on buzzwords such as opacity, transparency, explainability (e.g. Russo, Schliesser and Wegemans 2023). Arguably, this does little in the way of pre-empting future ethical window-dressing.

Based on my work, I argue that agential realism offers us a way to harness the intuitions inherent to many critical reflections around AI ethics. An agential realist approach to AI ethics would consider ethical and epistemic issues as undivorceable from one another and, simultaneously, would operate a post-humanist *de-centring of AI as the locus of ethics and epistemology*. In other words, AI ethics should not be considered as a question limited to the (admittedly hard to trace) boundaries of the technology itself. Instead, it should be conceptualised as an issue spanning a plethora of other actors, humans and non-. This would entail a different conception of epistemic issues: they should not only be about interrogating how algorithms reach a specific output but should also include, for instance, the material conditions, and the lived reality, of the data work that accompanies and makes possible processes of machine learning.

When it comes to specifying the strictly ethical issues that an agential-realism-inspired AI ethics should address, agential realism leaves us relatively unequipped. Barad's treatment of ethics is normative (subjects emerging from agential cuts *need to* take responsibility for the enacted realities), but it tells us little about the conditions and mechanics of such responsibility. Building on Amoore (2020) and on my own empirically-rooted analysis of ethical issues emerging in datafied clinical practice, I point here to the potential of *doubt as a concept to think through the mechanics of responsibility in intra-actions with AI*. Indeed, one of the foundational points of Amoore's cloud ethics is the necessity to restore the doubt inherent to each bifurcation in an algorithm's decision tree, against algorithmically-driven certainty (2023). In a similar fashion, the ability to doubt data, to consider the possibility of the otherwise, emerged in my analyses as a necessary condition for professionals' decision-making in datafied and increasingly automated clinical settings.

Noticeably, this notion of doubt might, at first sight, present some similarities with discussions of transparency and opacity in AI ethics debates. However, these discussions tend to be framed in cognitive terms, as the ability to *understand* a given technology and its workings (Khalili 2023; Russo, Schliesser and Wegemans 2023). Conversely, the idea of doubt I work with, building on Amoore, foregrounds precisely the *incompleteness of any form of knowing* and the ultimate contingency of reality as a basis for ethics. It is exactly because anyone's (including an AI's) understanding is necessarily partial that following doubt leads to exploring alternative configurations of what might be real, what might be the case. This, incidentally, proposes a radical different perspective on questions of responsibility and decision-making in the wake of AI. Namely, it raises questions about not who should be held accountable when things go wrong, but rather around how as-good-as-possible conditions can be created for professional to perform responsible decision-making, that is, in my view, to exert doubt *with* AI technologies.

To sum up, it is worth reiterating how future research should harness the potential of post-humanist approaches to AI ethics. This would entail an empirically-rooted consideration of ethical and epistemic issues that de-centres technology, reframing questions of knowledge, doubt and responsibility at the intersection of technology and humans.

Unpacking performativity: Clinical AI as an apparatus

A third implication of my work is the necessity of **considering AI's performativity at the care- and organisational level simultaneously**. My empirical analyses have demonstrated that care-related aims of clinical AI (i.e. the dimensions that tend to be addressed by analytics and algorithmic output signals) are, in practice, hard to disentangle from organisational aims (generally, the efficient use of scarce resources, disproportionately human resources). In other words, patients' clinical data is mobilised to allow a scarce workforce to function under duress. Turning patients into more manageable, less doubt-full, epistemic objects serves the goal of achieving care that it as efficient as possible in a context of scarcity.

As already brought up in chapter two, I propose the Baradian concept of *apparatus* to unpack the performativity of clinical AI at the care and organisational level simultaneously. Here, it is worth pausing briefly to consider how the notion of AI as an apparatus relates to other concepts in the literature. For instance, Louise Amoore (2020) has proposed the idea of algorithms as aperture instruments: by "generat[ing] what is of interest in the data environment," algorithms function as an aperture, that is, "an opening that is simultaneously a narrowing, a closure, and an opening onto a scene" (16). Building on Amoore, I also have proposed to consider clinical AI in its processes of exclusion of alternatives, of reduction of multiplicities to a singular output. I have found the concept of aperture instruments generative in unearthing the performativity of clinical AI at the level of enacted care objects. Thinking of algorithms as aperture instruments enables us to describe how output signals remake various categories of patients into graspable and unidimensional objects for professional intervention. However, in my view, relying on this concept alone leaves unattended the organisational dimension that emerges as a central target of clinical AI interventions.

I thus propose mobilising the Baradian notion of apparatus. Like Amoore, Barad emphasises apparatus's role in enacting "exclusionary boundaries" (2003: 816). Like apertures, apparatuses are implicated in restricting a vast field of possibilities contained within intra-acting phenomena, extracting from them a limited range of what can be. Unlike aperture instruments, apparatuses let us think through clinical AI's organisational performativity. The concept of apparatus broadens our focus from questions of perception (i.e. what is made perceivable through data analytics) to the concrete reconfigurations, as well as the inclusion and exclusions in knowledge-making practices, that I have described in this book as glass-breaking. As I have shown, thinking through the concept of apparatus enables us to account for materiality and affordances of machine-legible objects, as well as for the organisational reconfigurations (i.e. the roles and labour) required to achieve machine legibility.

The concept of apparatus enables us to account for the local specificity of clinical AI, casting it as an unstable phenomenon itself open to reconfiguration, and which can variate across organisational settings. Moreover, it pushes us to look to specific organisational and socio-historical contingencies that, themselves, shape the (unstable) phenomenon of clinical AI, restricting the possibilities for what clinical AI itself might be. Thinking through clinical AI as an apparatus lets us account for the role of *efficiency* in bringing about the specific version of clinical AI that we observe in practice today. This means that the reconfigurations AI generates in clinical practice (e.g. how it reshapes practices of attention and care provision, or how it reconfigures patients as objects for intervention) can be appraised as informed and constrained by a context of scarcity and an imperative of efficiency widespread in healthcare organisations. Finally, the notion of apparatus offers a way out of deterministic views of technologies. An apparatus itself emerges out of repeated intra-actions, and immanent in those repetitions is a core of instability and an inherent potential for reconfiguration. This offers a hopeful view of clinical AI: there is no intrinsic necessity for its potentially oppressive qualities, as I have described them here, and there are openings for changing clinical AI into a different kind of technology.

To sum up, the notion of clinical AI as an apparatus enables to surface its processes of exclusion of alternatives and reduction of multiplicities to a singular output, and to de-naturalise the reconfigurations it brings to clinical practice by tracing them back to the context of scarcity and to the efficiency imperatives that necessitate them.

Methodological implications: On studying absent technologies

The main methodological implication emerging from my work is about finding strategies to study a technology that, like clinical AI, is characterised by ontologies of absence — that is, a technology that manages to be present, and does things (thus *de facto* displaying a form of agency), all the while not being, technically speaking, embedded in practice[23]. The strategy I have proposed throughout this book consists of two components: first, an empirically-anchored study of the reconfigurations set in motion by the *expectation* of said technology; second, a speculatively-inflected, but organisationally situated, analysis of the *technological script*.

Studying organisational and professional reconfigurations driven by the expectation of clinical AI means, in a nutshell, foregrounding all the work that is being performed to bring about AI in various ways. This means, for instance, paying attention to new roles, tasks, objects, routines, and spatial reconfigurations that emerge in connection with AI technologies. In this book, I have proposed identifying such "connection" *emically* — that is, taking actors in the field seriously when they justify new practices and organisational structures as necessitated by (incipient) technological change. In other words, if actors feel they need to do something because of AI, this something should be subsumed under AI's performativity. This is a first, crucial component of studying absent technology: foregrounding the practical reconfigurations carried out to *prepare* clinical settings for AI's introduction.

The second component requires moving one step further away from the empirically observable, turning to the organisational and futures the

[23] Empirical studies of clinical technologies, both in STS and in (medical) sociology, tend to build on sociotechnical approaches that avoid both social and technological determinism by considering social and technical change as processes inextricable from one another. One such approaches, and one that has received considerable attention in the literature, as I articulated in this book's introduction, is Timmermans & Berg's (2003) ANT-inspired 'technology-in-practice.' This approach builds on the assumption, derived from ANT, that what technology does to clinical practice is a matter of what I have been calling intra-action with other actors, thus turning technological agency into an open question to be researched empirically in its specificity. At the methodological level, this presupposes two things: one, that the technology in question is (in the process of becoming) embedded in a network of practice; two, that it is possible to ethnographically observe the shifts, and indeed the various actants' achievements, that emerge in such process of embedding. Since these two conditions do not materialise in the case of clinical AI, we are pressed to find alternative methodological approaches to its study.

technology is mobilised to achieve. To do so, I propose revisiting the methodological implications of a classic STS concept: technological scripts (Akrich 1992). The concept builds on the insight that a fundamental aspect of technology consists of making hypotheses about how and where that technology will be used, as well as wagers on how it could change pre-existing practices. These hypotheses, of course, are reflected in the specific ways in which a technology is designed to function, and thus give rise to specific material affordances.

If the notion of scripts provides us with a conceptual basis to think through how technologies' materiality enables and restricts change in practice, it also avoids technological determinism through focus on *users' response to scripts*. Methodologically, then, the concept lends itself mostly to retrospective analyses, and does little in the way of analysing absent technologies.

To repurpose the notion of technological scripts for technologies that are characterised by *ontologies of absence* and achieve things while *not being used*, in this book I have proposed putting a speculative spin on it. This means, first, reverse-engineering technologies' expected uses, their users, and their workings (including data- and other infrastructures). In this reverse-engineering, we attempt to tease out the ways of thinking about the future guiding its design, and specifically what is (not) problematised in the future scenarios within which the technology is expected to act. In short, this in an exercise in unpacking which questions and issues AI is mobilised to address, which ones it disregards or sidelines, and the modalities of this addressing (for a similar example, cf. Delfanti and Frey 2020).

In the absence of empirically-observable instances of use, we need to find ways to contextualise this reverse-engineering withing an empirically-grounded study of current professional practices. In this step, it is necessary to identify aspects of professional practices that would be addressed and remade by the algorithmic output (e.g. assessing violence; ensuring a patient is not deteriorating). Close ethnographic attention can then be paid to how (i.e. using which technologies and which embodied techniques) such practices currently unfold. Producing a detailed ethnographic account of these practices enables the researcher to then compare them with the technology's working and underlying assumptions. This comparison inevitably entails some degree of speculation, since it leaves it to the analyst to imagine how the practices that they observe might reasonably shift were the technology in question to be introduced and used. However, the determinism such an exercise might be accused of can be moderated by involving professionals themselves in the reflection on how their practices might change.

It might, undoubtedly, be possible to push this speculative script analysis further than I have had the chance to do in this book. Indeed, my intensive ethnographic engagement meant that, because of time constraints, my fieldwork was not only strongly anchored in, but also limited to, the clinic. This, on the one hand, provided me with a solid engagement with local realities and actors. On the other hand, however, it did limit what I could observe in terms of the broader political economy within which these contemporary glass-breaking takes place. It also limited my analysis to the chronological snippets that I was able to piece together throughout my ethnographies, meaning that the way I cut the network to surface units of analysis of this study was reductionist not only in terms of spatial scales, but also of temporal ones (e.g. historical) (cf. Suchman 2007: 284, on this methodological consideration). If this is somewhat moderated through my engagement with prior literature addressing similar issues, it does still translate into a certain historical myopia. In this sense, anchoring speculative script analyses more strongly in a historical and political-economic examination of patterns of technological change in and beyond organisational contexts might help further sharpening the analytical power of a methodology that is, by necessity, rooted in speculation.

Practical implications and concluding remarks

In this book, I have endeavoured to open up for reflection the various ways in which AI's performativity materialises in current clinical practice. I have shown that this performativity should be considered to span both the adjustments clinical settings undergo to prepare for the advent of AI, and the professional and organisational realities that AI technologies promise to alternatively disrupt or reinforce. In this closing section, I attempt to condense what has often been a fairly abstract reflection into a few points of attention for practice.

First, policymakers, managers and professionals should be mindful that current instantiations of clinical AI fall short of the promise mostly upheld across innovation projects analysed in this thesis: tackling workforce shortages.

Presenting clinical AI as a solution to workforce shortages, as by consultancies have long been doing, and as policymakers and practitioners are increasingly starting to (Agyeman-Manu et al. 2023), is founded on a techno-solutionism that has shaky empirical foundations. Ethnographic engagement with clinical practice shows, in the first place, the *importance of data work*, that is, the work that goes into producing good (i.e. reliable and complete) clinical data. The work of producing good data proves cognitively and physically taxing, yet, as much as it is often brushed off as "administrative burden," it is increasingly inextricable from, and necessary for, care provision. Given the importance of data work, it should be an organisational ambition to professionalise it, rather than to automate it.

Although this book did not explicitly engage with this additional promissory dimension, one could venture that the reflection around data quality presented here also bears implications for the promise of enhanced quality of diagnosis that is often weaved into AI narratives. AI is often expected to improve the accuracy of diagnosis, "getting things right" more often than professionals thanks to its ability to access to large datasets. However, it is important to realise how the very possibility of improved diagnosis necessarily entails feeding high-quality data to AI systems. In turn, this necessitates intensified data work both in terms of data production and dataset curation. With the caveat that new techniques of data generation and data quality improvement might emerge in future years, it behoves managers and practitioners to be mindful of the potential tension between increased efficiency and enhanced quality as two central aspects of AI's future promise.

A second reason for caution around casting AI as a solution to workforce shortages has to do with the fact that *clinical AI operates a redirection of professionals' attention, rather than taking over their work*, as is often suggested in managerial and consultancy discourse. Unlike for other technologies of automation, AI's attention redirection is not a question of making some occupations obsolete, or even of job polarisation, where the tasks and job quality of non-professionalised workers worsen progressively (e.g. Autor, Katz and Kearny 2006; Good, Manning and Salomons 2009). Rather, we are witnessing a qualitative shift at the level of professional autonomy, in which AI-systems are actively implicated in defining the conditions of what counts as a "worthy" task.

Second, managers and designers need to think carefully about what **meaningful inclusion in decision-making** means. This entails opening up changing epistemic dynamics in a way that goes beyond simply inserting humans in the loop.[24] In the first place, there needs to be a keen eye for the knowledge-related disruptions that are generated by glass-breaking (i.e. the organisational reconfigurations prior and in the wake of AI introduction). Humans in the loop may have relatively little to contribute if these reconfigurations prevent them from verifying, for instance, the reliability of the data they are working with. Moreover, and emphatically, these reconfigurations should be guided by a sensitivity to what it means, in a specific organisational setting, for a human to meaningfully contribute to knowledge-making and, relatedly, to take responsibility for this knowledge. Based on my analysis, creating spaces and strategies for professionals to exert a doubtful attitude towards the data underpinning clinical AI's outputs might be necessary to this end. If healthcare organisations usually have structures in place to critically evaluate data and related professional decisions (e.g. interdisciplinary meetings, mobilisation of alternative sources), it is vital that these structures and practices remain central as data production becomes increasingly automated, and algorithmic output signals start to be included in decision-making.

Third, and related, professionals should **be wary of the aura of certainty projected by clinical AI's outputs**, and preserve spaces in which they can maintain a doubtful attitude towards data and knowledge. This means, in concrete terms, refusing to rely solely on AI's suggestions, especially when they impinge on issues that professionals themselves consider ethically laden. Recognising the centrality of an ongoing doubtful attitude when it comes to ethically engaging with data and machine learning outputs should is

[24] Most clinical AI technologies build on human-in-the-loop (HITL) frameworks. HITL refers to the design of a system in which humans are crucially embedded in automated systems, "handling challenging tasks of supervision, exception control, optimisation and maintenance" (Rahwan 2018: 6). In the context of AI, examples of HITL approaches are, for instance, mobilising humans to label data on which algorithms are trained, or integrating feedback from users through interactive machine learning. HITL is often proposed as a solution to enhance the outcomes of machine learning, but also to increase transparency and accountability in AI systems, thus tackling some of the ethical issues that emerge in the development and implementation of AI technologies. HITL aims at providing a way to embed accountability in automated decision-making and classification systems by introducing humans as actors that can be accountable for the technology, as well as correct machinic misbehaviour and to explain the reasoning behind their decisions (cf. Floridi & Sanders 2004; Rahwan 2018).

paramount. This doubting should be read as a questioning of the knowledge produced through these technologies that is not a dismissal or a lack of trust, but a constant probing of alternative accounts. Moreover, professionals should be proactive in deciding which part of clinical practice cannot (due, for instance, to concerns around reductionism), or should not (due to attendant ethical concerns, be made machine-readable. Selecting viable and desirable output signals for clinical AI should be object of joint consideration for professionals, designers and managers. Strategies for including patients in these deliberations should also be implemented.

Fourth, policymakers and regulators should **pay attention to the experimental modes of AI's ingression into clinical practice**. Indeed, despite regulations, there appears to be large underbelly of AI technologies experimentally entering clinical settings, and potentially staying. If pilots of these technologies, due to the very "learning" nature of AI, cannot have a clear endpoint, effective safeguards should be implemented in experimental settings, especially when vulnerable populations are concerned. At the same time, given the difficulty legislation shows in reaching clinical practice, professionals themselves need to be increasingly aware of the ethical challenges. Ethical debates should be taken closer to their lived realities, starting for instance from the negotiations and changes in decision-making that AI introduces in their own daily work, rather than from abstract ethical principles. This should be an ongoing process of evaluation, and could be achieved harnessing ethnographic engagements with clinical practice.

Fifth, and stemming from this last point, I invite managers, designers and action-oriented researchers continue to draw upon ethnographic participative methodologies, now inflected in a speculative direction to better address the current moment and the nature of the technology at hand. Ethnographic research enables us to shed light on the generative mechanisms of organisations and practices, and provides a basis for professionals' participation in organisational and technological changes by bringing abstract discussions closer to the lived reality of their daily work.

Initiatives around "inclusive AI" (e.g. Shams, Zowghi and Bano 2023), aimed at including professionals' perspectives and values in technology design, are mushrooming in healthcare organisations. However, in practice, such inclusion proves difficult, since professionals themselves (and, often, nurses in particular) feel like they do not have sufficient competencies to participate meaningfully in innovation projects. Moreover, professionals are pressed for time, meaning that inclusion is often performed in a hasty fashion. A speculative, ethnographic, participative methodology would materialise in workshops that mobilise ethnographic observations of concrete practices to help professionals articulate stakes and values inherent to their tasks. It would combine these conversations with discussions that support professionals in picking apart AI technologies' scripts, thinking through the potential implications of new geographies of responsibility in their concrete organisational settings. A corollary of this, of course, is that the necessary space and time for practicing inclusive innovation must be provided. Moreover, it behoves us to consider to what extent innovation should be termed "inclusive" in instances where professionals might have limited opportunities to ultimately refuse the technology being proposed.

Sixth, and final, policymakers, managers and professionals should be mindful that **a focus on clinical AI should not obscure other politically relevant issues**, such as labour conditions. Clinical AI should not be considered a necessity, and the reconfigurations it supposedly mandates should not be cast as inevitable. It is vital to keep in mind that issues emerging in the healthcare sector might deserve other, non-technological, approaches, and that opting for investments in cutting-edge technologies (as opposed to other types of organizational change) represents a political choice.

If the shadow of clinical AI I have traced in this book might appear bleak, this should not lead to moral panics or discouragement, but rather should be a spur for finding new strategies to achieve better futures *alongside*, rather than *through*, technology. I return, in closing, to Hito Steyerl's reflections on broken glass and AI's shadows. At the end of her talk, Steyerl states:

> the object casting the shadow defines the shadow, but if the object isn't present, then, by changing the shadow, the object itself will change. If … the shadow starts to be healed, then the object which is supposed to cause it can retroactively be changed as well. (2019)

Steyerl's insight that changing a shadow might retroactively do something to the not-there-yet object casting it has much to offer to policy and managerial approaches to clinical AI. Rather than resigning to an oppressive present, and awaiting utopic futures brought about by technology alone, we need to let AI's shadow point us to what needs to be changed. Achieving a clinical AI that is ethically, professionally, and organisationally desirable implies foregrounding, and taking seriously in their irreducibility, the ethical, professional, and organisational issues of the present.

References

Abbott A (1988) *The system of professions: An essay on the expert division of labor*. Chicago: Chicago University Press.

Abrishami P, Boer A and Horstman, K (2014) Understanding the adoption dynamics of medical innovations: Affordances of the da Vinci robot in the Netherlands. *Social Science & Medicine* 117: 125–133.

Abrishami P, Boer A and Horstman K (2017) Value in co-creation: Subjecting innovative in-hospital technologies to multi-stakeholder appraisal. *International Journal of Hospital Based Health Technology Assessment* 1: 12–30.

Agyeman-Manu K, Ghebreyesus TA, Maait M, et al. (2023) Prioritising the health and care workforce shortage: Protect, invest, together. *The Lancet Global Health* 11(8): e1162–e1164.

Ahmed S (2010) Happy objects. In: Gregg M and Segworth GJ (eds) *The affect theory reader*. Durham and London: Duke University Press, pp. 29-51.

Ahmed S (2014) Atmospheric walls. Blog post. Available at: https://feministkilljoys.com/2014/09/15/atmospheric-walls/ (accessed 4 April 2023).

Ahmed S (2019) *What's the use? On the uses of use*. Durham and London: Duke University Press.

AI HLEG (2019) Ethics guidelines for trustworthy AI. European Commission: Brussels.

Akrich M (1992) The de-scription of technical objects. In: Bijker WE and Law J (eds) *Shaping technology / building society: Studies in sociotechnical change*. Cambridge, MA: MIT Press, pp. 205-224.

Almvik R, Woods P, and Rasmussen K (2000) The Brøset Violence Checklist: Sensitivity, specificity, and interrater reliability. *Journal of Interpersonal Violence* 15(12): 1284–1296.

Amoore L (2013) *The politics of possibility: Risk and security beyond probability*. Durham and London: Duke University Press.

Amoore L (2020) *Cloud ethics: Algorithms and the attributes of ourselves and others*. Durham and London: Duke University Press.

Amoore L (2023) Machine learning political orders. *Review of International Studies* 49(1): 20-36.

Anderson B (2010) Preemption, precaution, preparedness: Anticipatory action and future geographies. *Progress in Human Geography* 34: 777–798.

Anderson C (2008) The end of theory: The data deluge makes the scientific method obsolete. *Wired*, June 23: 1– 3.

Andreassen HK, Kjekshus LE, and Tjora A (2015) Survival of the project: A case study of ICT innovation in health care. *Social Science & Medicine 132:* 62–69.

Apfeldorfer C, Ulrich K, Jones G, et al. (2008) Object orientated automated image analysis: Quantitative and qualitative estimation of inflammation in mouse lung. *Diagnostic Pathology* 3(Suppl 1).

Archibald M, Wiebe S, Rieger K, Linton J, and Woodgate R (2021) Protocol for a systematic review of living labs in healthcare. *BMJ open*, 11(2): e039246.

Autor D (2015) Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives* 29(3): 3–30.

Autor D, Katz L and Kearny M (2006) The polarization of the American labor market. *American Economic Review* 96(2): 189-194.

Avnoon N and Oliver AL (2023) Nothing new under the sun: Medical professional maintenance in the face of artificial intelligence's disruption. *Big Data & Society*, 10(2).

Bae H (2021) Intensive care nurse staffing and nurse outcomes: A systematic review. *Nursing in Critical Care* 26(6): 457-466.

Bailey S, Pierides D, Brisley A, Weisshaar C, and Blakeman T (2020) Dismembering organisation: The coordination of algorithmic work in healthcare. *Current Sociology 68*(4): 546–571.

Bar-Lev S (2015) The politics of healthcare informatics: Knowledge management using an electronic medical record system. *Sociology of Health & Illness 37*(3): 404–421.

Barad K (2003) Posthumanist performativity: Toward an understanding of how matter comes to matter. *Signs: Journal of Women in Culture and Society* 28(3): 801-31.

Barad K (2007) *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Durham and London: Duke University Press.

Barrett M, Oborn E, Orlikowski WJ and Yates J (2012) Reconfiguring boundary relations: Robotic innovations in pharmacy work. *Organisation Science* 23(5): 1448–1466.

Bareis J and Katzenbach C (2022) Talking AI into being: The narratives and imaginaries of national ai strategies and their performative politics. *Science, Technology, & Human Values* 47(5): 855-881.

Barnett H (2014) What humans can learn from semi-intelligent slime. Available at: https://www.youtube.com/watch?v=2UxGrde1NDA (last accessed 20 October 2020).

Barrett M, Oborn E, Orlikowski WJ, and Yates J (2012) Reconfiguring boundary relations: robotic innovations in pharmacy work. *Organisation Science 23*(5): 1448–1466.

Beaulieu A (2002) Images are not the (only) truth: Brain mapping, visual knowledge, and iconoclasm. *Science, Technology, & Human Values* 27(1): 53–86.

Beaulieu A, Scharnhorst A, and Wouters P (2007) Not another case study: A middle-range interrogation of ethnographic case studies in the exploration of e-science. *Science, Technology, & Human Values* 32(6): 672–692.

Berg M and Goorman E (1999) The contextual nature of medical information. *International Journal of Medical Informatics* 56(1-3): 51-60.

Bijsterveld K (2019) *Sonic skills:Llistening for knowledge in science, medicine and engineering (1920s-present).* Basingstoke: Palgrave Macmillan.

Biruk C (2018) *Cooking data: Culture and politics in an African research world*. Durham and London: Duke University Press.

Blease C, Kaptchuk TJ, Bernstein MH, Mandl KD, Halamka JD, and DesRoches CM (2019) Artificial intelligence and the future of primary care: Exploratory qualitative study of UK general practitioners' views. *Journal of Medical Internet Research* 21(3): e12802.

boyd d and Crawford C (2012) Critical questions for Big Data. *Information, Communication & Society* 15(5): 662-679.

de Bont A, van Exel J, Coretti S, Ökem SG, Janssen M, et al. (2016) Reconfiguring health workforce: A case-based comparative study explaining the increasingly diverse professional roles in Europe. *BMC Health Services Research* 16: 637.

Bourla A, Ferreri F, Ogorzelec L, Peretti CS, Guinchard C, and Mouchabac S (2018) Psychiatrists' attitudes toward disruptive new technologies: Mixed-methods study. *JMIR Mental Health* 5(4): e10240.

Borger T, Mosteiro P, Kaya H, Rijcken E, Salah AA, Scheepers F, and Spruit M (2022) Federated learning for violence incident prediction in a simulated cross-institutional psychiatric setting. *Expert Systems with Applications* 199: 116720.

Borgman CL (2015) *Big data, little data, no data: Scholarship in the networked world*. Cambridge: MIT Press.

Bossen C (2019) The emergence of new data work occupations in healthcare: The case of medical scribes. *International Journal of Medical Informatics* 123: 76–83.

Bossen C, Pine KH, Cabitza F, Ellingsen G, and Piras EM (2019). Data work in healthcare: An introduction. *Health Informatics Journal* 25(3): 465–74.

Boyko JA, Lavis JN, Abelson J, Dobbins M, and Carter N (2012) Deliberative dialogues as a mechanism for knowledge translation and exchange in health systems decision-making. *Social Science & Medicine* 75: 1938-1945.

van den Broek D (2017) Perforated body work: The case of tele-nursing. *Work, Employment and Society* 31(6): 904–920.

Bucher T (2012) A technicity of attention: How software 'makes sense.' *Culture Machine* 13: 1-23.

Bucher T and Helmond A (2018) The affordances of social media platforms. In: Burgess J, Poell T and Marwick A (eds), *The SAGE Handbook of Social Media*. London: SAGE Publications Ltd, pp. 233-253.

Bullock JB (2019) Artificial Intelligence, discretion, and bureaucracy. *The American Review of Public Administration* 49(7): 751-761.

Burri RV (2008) Doing distinctions: Boundary work and symbolic capital in radiology. *Social Studies of Science* 38(1): 35–62.

Burton J (2023) Algorithmic extremism? The securitization of artificial intelligence (AI) and its impact on radicalism, polarization and political violence. *Technology in Society* 75: 102262.

Butler J (2011 [1990]) *Gender trouble: Feminism and the subversion of identity*. London: Routledge.

Butler J (1993) *Bodies that matter: On the discursive limits of sex*. London: Routledge.

Callen JL, Westbrook JI, and Braithwaite, J. (2006). The effect of physicians' long-term use of CPOE on their test management work practices. *Journal of the American Medical Informatics Association*: 13(6): 643–652.

Campbell ML and Rankin JM (2017) Nurses and electronic health records in a Canadian hospital: Examining the social organisation and programmed use of digitised nursing knowledge. *Sociology of Health & Illness* 39(3): 365–379.

Campolo A and Crawford K (2020) Enchanted determinism: Power without responsibility in artificial intelligence. *Engaging Science, Technology, and Society* 6: 1-19.

Cartwright L (1995) *Screening the body: Tracing medicine's visual culture*. Minneapolis and London: University of Minnesota Press.

Chan S, Siegel EL (2019) Will machine learning end the viability of radiology as a thriving medical specialty? *British Journal of Radiology* 91: 20180416.

Challen R, Denny J, Pitt M, et al. (2019) Artificial intelligence, bias and clinical safety. *BMJ Quality & Safety* 28: 231-237.

Clarke A (2016) Anticipation work: Abduction, simplification, hope. In: Bowker GC, Timmermans S, Clarke AE Balka E (eds) *Boundary objects and beyond: Working with Leigh Star*. Cambridge: MIT Press, pp. 85-119.

Crary J (2001) *Suspensions of perception: Attention, spectacle, and modern culture*. Cambridge: MIT Press.

Cresswell K, Cunningham-Burley S, and Sheikh A (2018) Health care robotics: Qualitative exploration of key challenges and future directions. *Journal of Medical Internet Research* 20(7): e10410.

Crogan P and Kingsley S (2012) Paying attention: Towards a critique of the attention economy. *Culture Machine* 13: 1-29.

D'Hotman D and Loh E (2020) AI enabled suicide prediction tools: A qualitative narrative review. *BMJ Health & Care* Informatics 27(3).

D'Ignazio C and Klein LF (2020) *Data feminism*. Cambridge: MIT Press.

Danesi G, Pralong M, Panese F, Burnand B, and Grossen M (2020) Techno-social reconfigurations in diabetes (self-) care. *Social Studies of Science* 50(2): 198–220.

Davenport TH and Beck JC (2001) *The attention economy: Understanding the new currency of business*. Cambridge, MA: Harvard Business Press.

David GC, Garcia AC, Rawls AW, and Chand D (2009) Listening to what is said - transcribing what is heard: The impact of speech recognition technology (SRT) on the practice of medical transcription (MT). *Sociology of Health & Illness* 31(6): 924–938.

Delfanti A and Frey B (2020) Humanly extended automation or the future of work seen through amazon patents. *Science, Technology, & Human Values* 46(3): 655-682.

Deloitte (2023) 2024 Global health care sector outlook: Navigating transformation. Available at: https://www2.deloitte.com/content/dam/Deloitte/ch/Documents/public-sector/deloitte-ch-en-shortfall-chapter.pdf (last accessed: 9 February 2024).

Denton E, Hanna A, Amironesei R, Smart A, and Nicole H (2021) On the genealogy of machine learning datasets: A critical history of ImageNet. *Big Data & Society* 8(2).

Derrida J (2001) Cogito and the history of madness. In: *Writing and difference*, translated by A. Bass. London: Routledge, pp. 31-63.

Diana S. (2022). On the barricades of the incalculable: Against algorithm addiction. *INC Long Form*. Available at: https://networkcultures.org/longform/2019/03/22/on-the-barricades-of-the-incalculable-against-algorithm-addiction/ (last accessed: 28 February 2024).

De La Fabián R, Jiménez-Molina Á, and Pizarro Obaid F (2023) A critical analysis of digital phenotyping and the neuro-digital complex in psychiatry. *Big Data & Society* 10(1).

Dixon-Woods M, Cavers D, Agarwal S, et al. (2006) Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Medical Research Methodology* 6(35).

Donaldson M (2013). Stage Manager. Available online: https://www.youtube.com/watch?v=IxaXb_16cSY&t=362s (last accessed 19 April 2024).

Dowling E (2019) *The care crisis: What caused it and how can we end it?* London: Verso.

Dumit J (2012) *Drugs for life: How pharmaceutical companies define our health*. Durham and London: Duke University Press.

Dupret K (2017) Working around technologies-invisible professionalism? *New Technology, Work and Employment* 32(2): 174–187.

Ecks S (2005) Pharmaceutical citizenship: Antidepressant marketing and the promise of demarginalization in India. *Anthropology & Medicine* 12(3): 239-254.

EIT Health & McKinsey & Company (2020). Transforming healthcare with AI: The impact on the workforce and organisations. Available online: https://eithealth.eu/our-impact/our-reports/report-transforming-healthcare-with-ai/ (last accessed: 19 April 2024).

Ekbia HR and Nardi BA (2017) *Heteromation, and other stories of computing and capitalism*. Cambridge, MA: The MIT Press.

Elemento O, Leslie C, Lundin J, and Tourassi G (2021) Artificial intelligence in cancer research, diagnosis and therapy. *Nature Reviews Cancer* 21(12): 747-752.

Elish MC and Watkins EA (2020) Repairing innovation: A study of integrating ai in clinical care. Report. Data & Society.

Egermark M, Blasiak A, Remus A, Sapanel Y, and Ho D (2022) Overcoming pilotitis in digital medicine at the intersection of data, clinical evidence, and adoption. *Advanced Intelligent Systems* 4(9): 2200056.

Engesmo J and Tjora AH (2006) Documenting for whom? A symbolic interactionist analysis of technologically induced changes of nursing handovers. *New Technology, Work and Employment* 21(2): 176–189.

Espeland WN and Stevens ML (2008) A sociology of quantification. *European Journal of Sociology* 49: 401–436.

Fältholm Y and Jansson A (2008) Telephone advisory services-nursing between organisational and occupational professionalism. *New Technology, Work and Employment* 23(1–2): 17–29.

Felder M, Schuurmans J, van Pijkeren N, Kuijper S, Bal R, and Wallenburg I (2023) Bedside politics and precarious care: New directions of inquiry in critical nursing studies. *Advances in Nursing Science*.

Felt U (2014). Sociotechnical imaginaries of “the internet,” digital health information and the making of citizen-patients. In: Hilgartner S, Miller C and Hagendijk R (eds) *Science and democracy: Making knowledge and making power in the biosciences and beyond*. London: Routledge, pp. 176-197.

Fernandes BS, Williams LM, Steiner J et al. (2017) The new field of 'precision psychiatry'. *BMC Med* 15(80).

Findlay P, Lindsay C, McQuarrie J, et al. (2017) Employer choice and job quality: Workplace innovation, work redesign, and employee perceptions of job quality in a complex health-care setting. *Work and Occupations* 44(1): 113–136.

Fisher Wilson J (2009). Making electronic health records meaningful. *Annals of Internal Medicine* 151(4): 293-296.

Flemming K (2009) Synthesis of quantitative and qualitative research: An example using Critical Interpretive Synthesis. *Journal of Advanced Nursing* 66(1): 201-217.

Floridi L and Sanders J (2004) On the morality of artificial agents. *Minds and Machines* 14: 349–379.

Flyverbom M and Garsten C (2021) Anticipation and organisation: Seeing, knowing and governing futures. *Organisation Theory* 2(3): 26317877211020325.

Ford M (2021) Nurse leader sounds alarm over pandemic ICU ratios 'becoming routine.' Available at: https://www.nursingtimes.net/news/workforce/nurse-leader-sounds-alarm-over-pandemic-icu-ratios-becoming-routine-29-09-2021/ (last accessed: 11 May 2023).

Foster J, McLeod J, Nolin J, and Greifeneder E (2018) Data work in context: Value, risks, and governance. *Journal of the Association for Information Science and Technology* 69(12): 1414–27.

Foucault M (1980). *Power/Knowledge: Selected Interviews and Other Writings, 1972–1977*. New York: Pantheon Books.

Friedrich K (2010) 'Sehkollektiv': Sight styles in diagnostic Computed Tomography. *Medicine Studies* 2: 185–195.

Friis JKBO (2021) Radiology as skillful coping and enactive hermeneutics. In: Fried SJ and Rosenberger R (eds) *Postphenomenology and imaging: How to read technology*. Washington: Rowman & Littlefield/ Lexington Books, pp. 127-148.

Galetsi P, Katsaliaki K, and Kumar S (2019) Values, challenges and future directions of big data analytics in healthcare: A systematic review. *Social Science & Medicine* 241: 112533.

Gaube S, Suresh H, Raue M, et al. (2022) Do as AI say: susceptibility in deployment of clinical decision-aids. *npj Digital Medicine* 4(31).

Garcia LM (2020) Feeling the vibe: Sound, vibration, and affective attunement in electronic dance music scenes. *Ethnomusicology Forum* 29(1): 21-39.

Gardner, J. & Williams, C. (2015). Corporal diagnostic work and diagnostic spaces: Clinicians' use of space and bodies during diagnosis. *Sociology of Health & Illness*, 37, 765– 781.

Ghaznavi F, Evans A, Madabhushi A and Feldman M. (2013) Digital imaging in pathology: Whole-slide imaging and beyond. *Annual Review of Pathology* 8: 331–359.

Gibbs A (2010) After affect: Sympathy, synchrony, and mimetic communication. In: Gregg M and Segworth GJ (eds) *The affect theory reade*r. Durham and London: Duke University Press, pp. 186–205.

Giest S and Raaphorst N (2018) Unraveling the hindering factors of digital public service delivery at street-level: the case of electronic health records. *Policy Design and Practice* 1(2): 141-154.

Giraud EH (2019) *What comes after entanglement? Activism, anthropocentrism, and an ethics of exclusion*. Durham and London: Duke University Press.

Gitelman L and Jackson V (2013) Introduction. In: Gitelman L (ed) *"Raw data" is an oxymoron*. Cambridge: MIT Press, pp. 1-14.

Glaser VL, Pollock N and D'Adderio L (2021) The Biography of an Algorithm: Performing algorithmic technologies in organizations. *Organization Theory* 2(2).

Goffman E (1959) *The presentation of self in everyday life*. Garden City, NY: Doubleday & Company.

Goldhaber M (1997) The Attention Economy and the Net. *First Monday* (2)4. Available at: http://firstmonday.org/article/view/519/440 (last accessed 11-05-2023).

Goodwin D (2010) Sensing the way: Embodied dimensions of diagnostic work. In: Büscher M, Goodwin D and Mesman J (eds) *Ethnographies of diagnostic work. Dimensions of transformative practice*. London: Palgrave Macmillan, pp. 73-92.

Goos M, Manning A and Salomons A (2009) Explaining job polarization. *American Economic Review* 104(8): 2509-2526

Grassl N, Nees J, Schramm K, et al. (2018) A web-based survey assessing the attitudes of health care professionals in Germany toward the use of telemedicine in pregnancy monitoring: Cross-sectional study. *JMIR MHealth and UHealth* 6(8): e10063.

Gray ML and Suri S (2019) *Ghost work: How to stop Silicon Valley from building a new global underclass*. Boston: Houghton Mifflin Harcourt.

Greenhalgh T and Stones R (2010) Theorising big IT programmes in healthcare: Strong structuration theory meets actor-network theory. *Social Science & Medicine* 70(9): 1285–1294.

Greenhalgh T, Wherton J, Shaw S, Papoutsi C, Vijayaraghavan S, Stones R (2019) Infrastructure revisited: An ethnographic case study of how health information infrastructure shapes and constraints technological innovation. *Journal of Medical Internet Research* 21(12): e16093.

Grote T and Berens P (2019) On the ethics of algorithmic decision-making in healthcare. *J Med Ethics* 46: 205-211.

Grünloh C, Cajander Å, and Myreteg G (2016) "The record is our work tool!"—Physicians' framing of a patient portal in Sweden. *Journal of Medical Internet Research* 18(6): e167.

Gupta Strategists & Edwards Lifesciences (2021). Prevent the next wave: Why labour-saving innovations are needed in healthcare and how an innovative ecosystem enables progression. Available at: https://gupta-strategists.nl/storage/files/Prevent-the-next-wave.pdf (last accessed: 28 February 2024).

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

Hagendorff T (2020) The ethics of AI ethics: An evaluation of guidelines. *Mind Machine* 30(1):99–120.

Hains IM, Georgiou A, and Westbrook JI (2012) The impact of PACS on clinician work practices in the intensive care unit: A systematic review of the literature. *Journal of the American Medical Informatics Association* 19(4): 506–513.

Håland E (2012). Introducing the electronic patient record (EPR) in a hospital setting: Boundary work and shifting constructions of professional

identities. *Sociology of Health & Illness* 34(5): 761–775.
Håland E and Melby L (2015) Negotiating technology-mediated interaction in health care. *Social Theory & Health* 13(1): 78–98.
Halford S, Obstfelder A, and Lotherington AT (2010) Changing the record: The inter-professional, subjective and embodied effects of electronic patient records. *New Technology, Work and Employment* 25(3): 210–222.
Halpern O (2014) *Beautiful data: A history of vision and reason since 1945*. Durham and London: Duke University Press.
Halpern O and Mitchell R (2022) *The smartness mandate*. Cambridge: MIT Press.
Haraway DJ (1989) Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies* 14(3): 575-599.
Harris A (2011) In a moment of mismatch: Overseas doctors' adjustments in new hospital environments. *Sociology of Health & Illness* 33(2): 308-320.
Harris A (2021) *A sensory education*. London: Routledge.
Harris A (2023) Making Measuring Bodies. *Science, Technology, & Human Values* 48(1): 115-138.
Heath C, Luff P, and Svensson MS (2003) Technology and medical practice: Technology and medical Practice. *Sociology of Health & Illness 25*(3): 75–96.
Heimstädt M, Egbert S and Esposito E (2020) A pandemic of prediction: On the circulation of contagion models between public health and public safety. *Sociologica* 14(3): 1–24.
Henriksen A and Bechmann A (2020) Building truths in AI: Making predictive algorithms doable in healthcare. *Information, Communication & Society* 23(6): 802-816.
Henriksen A and Blond L (2023) Executive-centered AI? Designing predictive systems for the public sector. *Social Studies of Science* 53(5): 738-760.
Højlund H and Villadsen K (2020) Technologies in caregiving: Professionals' strategies for engaging with new technology. *New Technology, Work and Employment 35*(2): 178–194.

Høyer K (2023) *Data paradoxes: The politics of intensified data sourcing in contemporary healthcare*. Cambridge, MA: The MIT Press.

Høyer K and Wadmann S (2020) 'Meaningless work': How the datafication of health reconfigures knowledge about work and erodes professional judgement. *Economy and Society* 49(3): 433–454.

Hogle LF (2016) Data- intensive resourcing in healthcare. *BioSocieties* 11: 372–393.

Holford WD (2022) An ethical inquiry of the effect of cockpit automation on the responsibilities of airline pilots: Dissonance or meaningful control? *Journal of Business Ethics* 176: 141–157.

Hollin G, Forsyth I, Giraud E and Potts T (2017) (Dis)entangling Barad: Materialisms and ethics. *Social Studies of Science* 47(6): 918–941.

Hwang J and Christensen CM (2008) Disruptive innovation in health care delivery: A framework for business-model innovation. *Health Affairs* 27(5): 1329-1335.

Hunt LM, Bell HS, Baker AM and Howard HA (2017) Electronic health records and the disappearing patient. *Medical Anthropology Quarterly* 31: 403– 421.

Hutchins E (1995) *Cognition in the wild*. Cambridge, MA: The MIT Press.

Ibrahim A, Lashen AG, Katayama A, et al. (2021) Defining the area of mitoses counting in invasive breast cancer using whole slide image. *Mod Pathol*.

Ihde D (2016) *Husserl's missing technologies.* New York: Fordham University Press.

Institute of Medicine (2001) *Crossing the quality chasm: A new health system for the 21st century*. Washington: National Academy Press.

Islam G (2022) Business ethics and quantification: Towards an ethics of numbers. *Journal of Business Ethics* 176(2): 195-211.

Ismail L, Materwala H, Karduck AP and Adem A (2020) Requirements of health data management systems for biomedical care and research: Scoping review. *Journal of Medical Internet Research* 22(7): e17508.

Jablonski R, Karppi T and Seaver N (2022) Introduction: Shifting Attention. *Science, Technology, & Human Values* 47(2): 235-242.

Jabr F (2012) How brainless slime moulds redefine intelligence. *Nature.*

Jarrahi MH, Newlands G, Lee MK, Wolf CT, Kinder E, et al. (2021) Algorithmic

management in a work context. *Big Data & Society* 8(2).
Jaton, F. & Sormani, P. (2023). Enabling 'AI'? The situated production of commensurabilities. Social Studies of Science, 53(5), 625-634.
Jensen DC and Pedersen LB (2017) The impact of empathy: Explaining diversity in street-level decision-making. *Journal of Public Administration Research and Theory* 27(3): 433–449.
Kellogg KC, Valentine MA and Christin A (2020) Algorithms at work: The new contested terrain of control. *Academy of Management Annals* 14(1): 366-410.
Kelly L (2023) Re-politicising the future of work: Automation anxieties, universal basic income, and the end of techno-optimism. *Journal of Sociology* 59(4): 828-843.
Keyes O (2021) Disability, discourse, and Artificial Intelligence. *Journal of Sociotechnical Critique* 1(1): 1-31.
Kiran AH, Oudshoorn N, and Verbeek, PP (2015) Beyond checklists: Toward an ethical-constructive technology assessment. *Journal of Responsible Innovation* 2: 5–19.
Kitchin R (2014) *The data revolution: Big Data, open data, data infrastructures and their consequences*. London: Sage.
Knorr Cetina K (1999) *Epistemic cultures: How the sciences make knowledge*. Cambridge: Harvard University Press.
Knorr Cetina K (2001) Objectual practice. In: Schatzki T, Knorr Cetina K and von Savigny E (eds), *The practice turn in contemporary theory*. London: Routledge, pp. 175–88.
Knorr Cetina K and Amann K (1990) Image dissection in natural scientific inquiry. *Science, Technology, and Human Values* 15(3): 259–283.
Knorr Cetina K and Reichmann W (2015) Epistemic Cultures. In: Wright JD (ed) *International Encyclopedia of the Social and Behavioral Sciences*. Amsterdam: Elsevier, pp. 873-880.
Kosofsky Sedgwick E (1993) Queer performativity: Henry James's *The art of the novel*. *GLQ* 1: 1-16.
Koutsouleris N, Hauser TU, Skvortsova V, and De Choudhury M (2022) From promise to practice: Towards the realisation of AI-informed mental health care. *Lancet Digital Health* 4: e829–40.
Lærum H, Karlsen TH, and Faxvaag A (2003) Effects of Scanning and

Eliminating Paper-based Medical Records on Hospital Physicians' Clinical Work Practice. *Journal of the American Medical Informatics Association* 10(6): 588–595.

Lakoff G and Johnson M (1980) *Metaphors we live by*. Chicago: University of Chicago Press.

Latour B (2004) Whose cosmos, which cosmopolitics? Comments on the peace terms of Ulrich Beck. *Common Knowledge* 10(3): 450–462.

Latour B (2005) *Reassembling the social: An introduction to Actor-Network-Theory*. New York: Oxford University Press.

Lebovit S, Lifshitz-Assaf H, and Levina N (2022) To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. *Organization Science* 33(1): 126-148.

Lehoux P, Poland B, Daudelin G, Holmes D, and Andrews G (2008) Displacement and emplacement of health technology: Making satellite and mobile dialysis units closer to patients? *Science, Technology, & Human Values* 33(3): 364–392.

Leonelli S (2012) Classificatory theory in data- intensive science: The case of open biomedical ontologies. *International Studies in the Philosophy of Science* 26: 47– 65.

Leonelli S (2016) *Data-centric biology: A philosophical study*. Chicago: University of Chicago Press.

Li RC, Wang JK, Sharp C, and Chen JH (2019) When order sets do not align with clinician workflow: Assessing practice patterns in the electronic health record. *BMJ Quality & Safety*.

Liboiron M (2021) *Pollution is colonialism*. Durham and London: Duke University Press.

Lie M and Sørensen K (eds) (1996) *Making technology our own? Domesticating technology into everyday life*. Oslo: Scandinavian University Press.

Linaker OM and Busch-Iversen H (1995) Predictors of imminent violence in psychiatric inpatients. *Acta Psychiatrica Scandinavica* 92(4): 250–254.

Lorkowski J and Jugowicz A (2020) Shortage of physicians: A critical review. In: Pokorski M (ed) *Medical research and innovation. Advances in experimental medicine and biology*. Springer, Cham.

Lockertsen Ø, Varvin S, Færden A and Vatnar SKB (2021) Short-term risk assessments in an acute psychiatric inpatient setting: A re-examination of the Brøset Violence Checklist using repeated measurements — Differentiating violence characteristics and gender. *Archives of Psychiatric Nursing* 35(1): 17 - 26

Loukissas YA (2019) *All data are local: Thinking critically in a data-driven society*. Cambridge: MIT Press.

Lupton D (2016) *The quantified self: A sociology of self-tracking*. Hoboken: John Wiley & Sons.

Lupton D and Maslen S (2017) Telemedicine and the senses: A review. *Sociology of Health & Illness* 39(8): 1557–1571.

Maiers C (2017) Analytics in action: Users and predictive data in the neonatal intensive care unit. *Information, Communication & Society* 20(6): 915–929.

Mao Y, Wang D, Muller M, Varshney KR, Baldini I, et al. (2019) How data scientists work together with domain experts in scientific collaborations: To find the right answer or to ask the right question? *Proceedings of the ACM on Human-Computer Interaction* 3: 1–23.

Marres N and Sormani P (2023) Testing AI: Do we have a situation? *CRC Media of Cooperation Working Paper Series* 28.

Marx K (1976 [1867]) *Capital, vol. 1*. London: Penguin.

Maslen S (2017) Layers of sense: The sensory work of diagnostic sensemaking in digital health. *Digital Health 3*.

Maslen S and Harris A (2021) Becoming a diagnostic agent: A collated ethnography of digital-sensory work in caregiving intra-actions. *Social Science and Medicine* 277: 113927.

May C, Gask L, Atkinson T, Ellis N, Mair F, and Esmail A (2001) Resisting and promoting new technologies in clinical practice: The case of telepsychiatry. *Social Science & Medicine* 52(12): 1889–1901.

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

Mesman J (2014) Relocation of Vulnerability in Neonatal Intensive Care Medicine. In: Hommels A, Mesman J, Bijker WE (eds) *Vulnerability in technological cultures*. Cambridge: The MIT Press, pp. 71-88.

Meyer J and Paré G (2018) The influence of telepathology on coordination practices. *Telemedicine and e-health* 24(9): 684-690.

Mittelstadt BD, Allo P, Taddeo M, Wachter S, and Floridi L (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society* 3(2).

Mittelstadt BT and Floridi L (2016) The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. *Sci Eng Ethics* 22: 303-341.

Møller NH, Bossen C, Pine KH, Nielsen TR, and Neff G (2020) Who does the work of data? *Interactions* 27(3): 52–55.

Mökander J, Floridi L (2021) Ethics-based auditing to develop trustworthy AI. *Mind Machine* 31(2): 323–327.

Mol A (2002) *The body multiple: Ontology in medical practice*. Durham and London: Duke University Press.

Mol A (2008) *The logic of care: Health and the problem of patient choice*. London: Routledge.

Molnar P (2021) Surveillance sovereignty: Migration management technologies and the politics of privatization. In: Hudson G and & Atak I (eds) *Migration, security, and resistance: Global and local perspectives*. London: Routledge, pp. 66-82.

Moreira T (2019) Devicing future populations: Problematising the relationship between quantity and quality of life. *Social Studies of Science* 49: 118–137.

Mossfeldt Nickelsen NC (2019) The infrastructure of telecare: Implications for nursing tasks and the nurse-doctor relationship. *Sociology of Health & Illness* 41(1): 67–80.

Mort M, May CR, and Williams T (2003) Remote doctors and absent patients: Acting at a distance in telemedicine? *Science, Technology, & Human Values* 28(2): 274–295.

Mueller F, Valsecchi R, Smith C, Gabe J, and Elston MA (2008) 'We are nurses, we are supposed to care for people': Professional values among nurses in NHS Direct call centres. *New Technology, Work and Employment* 23(1–2): 2–16.

Mulinari S (2023) Short-circuiting biology: Digital phenotypes, digital biomarkers, and shifting gazes in psychiatry. *Big Data & Society* 10(1).

Nagele J and Thamm A (2022) How AI can help avoid catastrophic overload of healthcare system in times of a worldwide pandemic. In: Schönbohm A, von Horsten HH and Plugmann P (eds) *Life science management:*

Management for professionals. Springer, Cham.

Nerland M and Jensen K (2012) Epistemic practices and object relations in professional work. *Journal of Education and Work* 25(1): 101-120.

Nettleton S, Burrows R, and Watt I (2008) Regulating medical bodies? The consequences of the 'modernisation' of the NHS and the disembodiment of clinical knowledge. *Sociology of Health & Illness* 30: 333– 48.

Nicolini D (2006) The work to make telemedicine work: A social and articulative view. *Social Science & Medicine* 62(11): 2754–2767.

Niemiec E (2022) Will the EU Medical Device Regulation help to improve the safety and performance of medical AI devices? *Digital Health* 8.

O'Malley AS, Draper K, Gourevitch R, Cross DA, and Scholle S H (2015) Electronic health records and support for primary care teamwork. *Journal of the American Medical Informatics Association* 22(2): 426–434.

Ọnụọha M (2022) 'If Then Maybe.' In Kahn NM (ed) *Mirror Stage: Between computability and its opposite (HOLO 3)*: 89-94.

Oudshoorn N (2008) Diagnosis at a distance: The invisible work of patients and healthcare professionals in cardiac telemonitoring technology. *Sociology of Health & Illness* 30(2): 272–288.

Oudshoorn N (2011) How places matter: Telecare technologies and the changing spatial dimensions of healthcare. *Social Studies of Science* 42(1): 121–142.

Ozalp H, Ozcan P, Dinckol D, Zachariadis M, and Gawer A (2022) "Digital colonisation" of highly regulated industries: An analysis of big tech platforms' entry into health care and education. *California Management Review* 64(4): 78-107.

Pantanowitz L, Quiroga-Garza GM, Bien L, Heled R, et al. (2020) An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: A blinded clinical validation and deployment study. *The Lancet Digital health* 2(8): e407–e416.

Papadopoulos D (2011) Alter-ontologies: Towards a constituent politics in technoscience. *Social Studies of Science* 41(2): 177–201.

Parisi L (2019) Critical computation: Digital automata and general artificial thinking. *Theory, Culture & Society* 36(2): 89-121.

Parwani AV (2019) Next generation diagnostic pathology: Use of digital pathology and artificial intelligence tools to augment a pathological diagnosis. *Diagnostic Pathology* 14: 138.

Pasquinelli M (2019) How a machine learns and fails — A grammar of error for artificial intelligence. *spheres: Journal for Digital Cultures* (5): 1–17.

Pasquinelli M (2023) *The eye of the master: A social history of Artificial Intelligence*. London: Verso.

Pedersen AM and Bossen C (2021)Data work in healthcare: An ethnography of a BI unit. i *InfraHealth 2021: Proceedings of the 8th International Conference on Infrastructures for Healthcare*. Kristiansand, Norway, 23/09/2021.

Pedersen MA Albris K and Seaver N (2021) The political economy of attention. *Annual Review of Anthropology* 50: 309-325.

Peiris D, Usherwood T, Weeramanthri T, Cass A, and Patel A (2011) New tools for an old trade: A socio-technical appraisal of how electronic decision support is used by primary care practitioners: Use of electronic decision support in primary care. *Sociology of Health & Illness 33*(7): 1002–1018.

Petersson L and Erlingsdóttir G (2018) Open notes in Swedish psychiatric care (part 2): Survey among psychiatric care professionals. *JMIR Mental Health* 5(2): e10521.

Petrakaki D, Barber N, and Waring J (2012) The possibilities of technology in shaping healthcare professionals: (Re/De-)professionalisation of pharmacists in England. *Social Science & Medicine 75*(2): 429–437.

Petrakaki D, Klecun E, and Cornford T (2016) Changes in healthcare professional work afforded by technology: The introduction of a national electronic patient record in an English hospital. *Organisation* 23(2): 206–226.

Petrakaki D and Kornelakis A (2016) 'We can only request what's in our protocol': Technology and work autonomy in healthcare. *New Technology, Work and Employment* 31(3): 223–237.

Pickersgill M (2018) Digitising psychiatry? Sociotechnical expectations, performative nominalism and biomedical virtue in (digital) psychiatric praxis. *Sociology of Health & Illness* 41: 16-30.

Pink S, Sumartojo S, Lupton D and Heyes La Bond C (2017) Mundane data: The routines, contingencies and accomplishments of digital living. *Big Data & Society* 4(1).

Pink S, Ruckenstein M, Willim R, and Duque M (2018) Broken data: Conceptualising data in an emerging world. *Big Data & Society* 5(1).

Pols J (2010) The heart of the matter: About good nursing and telecare. *Health Care Analysis* 18(4): 374–388.

Pols J (2011) Wonderful webcams: About active gazes and invisible technologies. *Science, Technology, & Human Values* 36(4): 451–473.

Popp Berman E & Hirschman D (2018). The sociology of quantification: Where are we now? *Contemporary Sociology* 47(3): 257-266.

Pringle A (2019) Mycologist explains how a slime mould can solve mazes. *Wired*. Available at: https://www.wired.com/video/watch/mycologist-explains-how-a-slime-mold-can-solve-mazes (last accessed: 19 April 2024).

Puschmann C and Burgess J (2014) Big Data, Big Questions| Metaphors of Big Data *International Journal of Communication 20(8)*.

Radder H (1998) The politics of STS. *Social Studies of Science* 28(2): 325-331.

Rahwan I (2018) Society-in-the-loop: Programming the algorithmic social contract. *Ethics and Information Technology* 20(1): 5–14.

Rajpurkar P, Chen E, Banerjee O, et al. (2022) AI in health and medicine. *Nature Medicine* 28: 31–38.

Raley R and Rhee J (2023) Critical AI: A field in formation. *American Literature* 95(2): 185–204.

Ranchordas S (2021) Experimental lawmaking in the EU: Regulatory Sandboxes. *EU Law Live*.

Ratner HF and Elmholdt K (2023) Algorithmic constructions of risk: Anticipating uncertain futures in child protection services. *Big Data & Society* 10(2).

Reed K, Kochetkova I, and Molyneux-Hodgson S (2016) 'You're looking for different parts in a jigsaw': Foetal MRI (magnetic resonance imaging) as an emerging technology in professional practice. *Sociology of Health & Illness* 38(5): 736–752.

Reich A (2012) Disciplined doctors: The electronic medical record and physicians' changing relationship to medical knowledge. *Social*

Science & Medicine 74(7): 1021–1028.

Rothschild A, Meng A, DiSalvo C, et al. (2022). Interrogating data work as a community of practice. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2), 1–28.

Ruckenstein M and Schüll ND (2017) The datafication of health. *Annual Review of Anthropology* 46: 261-278.

Satchell C and Dourish P (2009). Beyond the user: Use and non-use in HCI. In: *Proceedings of the 21st annual conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*. New York, NY: Association for Computing Machinery, pp. 9-16.

Saugel B, Kouz K, Meidert AS, et al. (2020) How to measure blood pressure using an arterial catheter: A systematic 5-step approach. *Critical Care* 24: 172.

Scheepers FE, Menger V and Hagoort K (2018) Datascience in de psychiatrie. *Tijdschrift Voor Psychiatrie* 60(3): 205-209.

Scheuerman MK, Pape M and Hanna A (2021) Auto-essentialization: Gender in automated facial analysis as extended colonial project. *Big Data & Society* 8(2).

Schubert C (2011) Making sure: A comparative micro-analysis of diagnostic instruments in medical practice. *Social Science and Medicine* 73: 851–857.

Schuster R, Pokorny R, Berger T, Topooco N, and Laireiter AR (2018) The advantages and disadvantages of online and blended therapy: Survey study amongst licensed psychotherapists in Austria. *Journal of Medical Internet Research 20*(12): e11007.

Schwennesen N (2019) Algorithmic assemblages of care: Imaginaries, epistemologies and repair work. *Sociology of Health & Illness* 41(S1): 176–192.

Seaver N (2017) Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society* 4(2).

Semel BM (2021) Listening like a computer: Attentional tensions and mechanised care in psychiatric digital phenotyping. Science, *Technology, & Human Values* 47(2): 266-290.

Seoane J, Varela-Centelles PI, Ramírez JR, Cameselle-Teijeiro J and Romero MA (2004) Artefacts in oral incisional biopsies in general dental practice: A pathology audit. *Oral Dis.* 10: 113–7.

Shams RA, Zowghi D, and Bano M (2023) AI and the quest for diversity and inclusion: a systematic literature review. *AI Ethics*.

Sharon T (2016) The Googlization of health research: From disruptive innovation to disruptive ethics. *Personalized Medicine* 13(6): 563-574.

Sharon T (2017) Self-tracking for health and the quantified self: Re-articulating autonomy, solidarity, and authenticity in an age of personalized healthcare. *Philosophy & Technology* 30(1): 93-121.

Sharon T (2021) From hostile worlds to multiple spheres: Towards a normative pragmatics of justice for the Googlization of health. *Medicine, Health Care and Philosophy* 24(3): 315-327.

Shastry KA and Sanjay HA (2022) Cancer diagnosis using artificial intelligence: A review. *Artificial Intelligence Review* 55: 2641–2673.

Sieja A, Markley K, Pell J, et al. (2019). Optimisation sprints: Improving clinician satisfaction and teamwork by rapidly reducing electronic health record burden. *Mayo Clinic Proceedings* 94(5): 793–802.

Simon HA (1971) *Designing organizations for an information-rich world*. Baltimore: Johns Hopkins University Press.

Smith GJD and Vonthethoff B (2017) Health by numbers? Exploring the practice and experience of datafied health. *Health Sociology Review* 26(1): 6-21.

Smith Glasgow ME, Colbert A, Viator J, and Cavanagh S (2018) The nurse-engineer: A new role to improve nurse technology interface and patient care device innovations. *Journal of Nursing Scholarship* 50(6): 601–611.

Spatar D, Kok O, Basoglu N, and Daim T (2019) Adoption factors of electronic health record systems. *Technology in Society* 58: 101144.

Star SL (1991) "The sociology of the invisible: The primacy of work in the writings of Anselm Strauss.' In: Maines DR (ed) *Social organisation and social process: Essays in honor of Anslem Strauss*. Hawthorne, NY: Aldine de Gruyter, pp. 265-283.

Star SL and Strauss A (1999) Layers of silence, arenas of voice: The ecology of visible and invisible work. *Computer Supported Cooperative Work*

8: 9–30.

Stengers I (2011) *Cosmopolitics II*. Minneapolis: University of Minnesota Press.

Stevens M, Wehrens R, and de Bont A (2018) Conceptualisations of Big Data and their epistemological claims in healthcare: A discourse analysis. *Big Data & Society 5*(2).

Stevens M, Wehrens R, and de Bont A (2020) Epistemic virtues and data-driven dreams: On sameness and difference in the epistemic cultures of data science and psychiatry. *Social Science & Medicine* 258: 113116.

Stevens M, Wehrens R, Kostenzer J, Weggelaar-Jansen AM and de Bont A (2022). Why personal dreams matter: How professionals affectively engage with the promises surrounding data-driven healthcare in Europe. *Big Data & Society* 9(1).

Steyerl H (2019). The language of broken glass. Video. Available at: https://mediathek.hkw.de/en/video/hito-steyerl--the-language-of-broken-glass--englisch-?backscreen=search&overlayObjectSlug=giulia-bruno--armin-linke--luc-steels--language-agents&overlayObjectType=audio&fi=contentFilters%3ALecture+Performance (last accessed: 1 May 2024).

Strand M, Gammon D, Eng LS, and Ruland C (2017) Exploring working relationships in mental health care via an e-recovery portal: Qualitative study on the experiences of service users and health providers. *JMIR Mental Health* 4(4): e54.

Strauss AL (1975) *Professions, work and careers*. London: Routledge.

Strauss AL (1988) The articulation of project work: An organisational process. *Sociological Quarterly* 29: 163-178.

Suchman L (2007) *Human-machine reconfigurations: Plans and situated actions.* Cambridge University Press.

Suchman L (2023) The uncontroversial ‘thingness’ of AI. *Big Data & Society* 10(2).

Tai-Seale M, Dillon EC, Yang Y, et al. (2019). Physicians’ well-being linked to in-basket messages generated by algorithms in electronic health records. *Health Affairs* 38(7): 1073–1078.

Taylor A (2018) The Automation Charade. *Logic Magazine*.

Taylor L (2024) AI lacks ethics checks for human experimentation. *Issues in Science and Technology*: 80-81.

Tavory I and Timmermans S (2014) *Abductive analysis: Theorising qualitative research*. Chicago: University of Chicago Press.

Thylstrup NB (2019) Data out of place: Toxic traces and the politics of recycling. *Big Data & Society* 6(2).

Thylstrup NB, Hansen KB, Flyverbom M, and Amoore L (2022) Politics of data reuse in machine learning systems: Theorising reuse entanglements. *Big Data & Society* 9(2).

Timmermans and Berg M (2003) The practice of medical technology: The practice of medical technology. *Sociology of Health & Illness* 25(3): 97–114.

Timmermans S and Tavory I (2012) Theory construction in qualitative research: From grounded theory to abductive analysis. *Sociological Theory* 30(3): 167-18.

Topol Review (2019) Preparing the healthcare workforce to deliver the digital future. Final Report February 2019 — A call for evidence. Report. Health Education England. Available at: https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf (last accessed: 19 April 2024).

Tran BD, Chen Y, Liu S, and Zheng K (2020). How does medical scribes' work inform development of speech-based clinical documentation technologies? A systematic review. *Journal of the American Medical Informatics Association* 27(5): 808–817.

Trittin-Ulbrich H, Scherer A, Munro I, and Whelan G (2020). Exploring the dark and unexpected sides of digitalisation: Toward a critical agenda. *Organisation* 28(1): 8-25.

Trondsen MV, Tjora A, Broom A, and Scambler G (2018) The symbolic affordances of a video-mediated gaze in emergency psychiatry. *Social Science & Medicine* 197: 87–94.

Tsing AL (2015) *The mushroom at the end of the world: On the possibility of life in capitalist ruins*. Princeton and Oxford: Princeton University Press.

United States Environmental Protection Agency (EPA) (2016). "Nonpoint source: Hydromodification and habitat alteration." Report. Available at: https://www.epa.gov/nps/nonpoint-source-hydromodification-

and-habitat-alteration (last accessed: 19 April 2024).
van den Tweel JG and Taylor CR (2010) A brief history of pathology. *Virchows Archiv: An International Journal of Pathology* 457(1): 3–10.
van Dijck J (2011) *The transparent body: A cultural analysis of medical imaging*. Seattle: University of Washington Press.
Vicsek L (2021) Artificial intelligence and the future of work — lessons from the sociology of expectations. *International Journal of Sociology and Social Policy* 41(7/8): 842-861.
Vogel M, Kaisers W, Wassmuth R, and Mayatepek E (2015) Analysis of documentation speed using web-based medical speech recognition technology: Randomised controlled trial. *Journal of Medical Internet Research* 17(11): e247.
Vogelsmeier AA, Halbesleben JRB, and Scott-Cawiezell JR (2008) Technology implementation and workarounds in the nursing home. *Journal of the American Medical Informatics Association* 15(1): 114–119.
Volksgezondheid, Welzijn en Sport (VWS), Ministerie van (n.d.) Wet verplichte ggz. Report. Available at: https://www.dwangindezorg.nl/wvggz (last accessed: 19 April 2024).
Vorderstrasse AA, Hammer MJ, and Dungan JR (2014) Nursing implications of personalised and precision medicine. *Seminars in Oncology Nursing* 30(2): 130–136.
Weber M and Rohracher H (2012) Legitimising research, technology and innovation policies for transformative change. Combining insights from innovation systems and multi-level perspective in a comprehensive 'failures' framework. *Research Policy* 41: 1037-1047.
Westbrook JI, Li L, Georgiou A, Paoloni R, and Cullen J (2013) Impact of an electronic medication management system on hospital doctors' and nurses' work: A controlled pre–post, time and motion study. *Journal of the American Medical Informatics Association* 20(6): 1150–1158.
Wiedemann L (2021) Being on standby: On maintenance work in chronic disease management. *ephemera: theory & politics in organisation* 21(1): 31-58.
Winthereik BR, van der Ploeg I, and Berg M (2007). The electronic patient record as a meaningful audit tool: Accountability and autonomy in general practitioner work. *Science, Technology, & Human Values*

32(1): 6–25.
Woodward S (2016) Object interviews, material imaginings and 'unsettling' methods: Interdisciplinary approaches to understanding materials and material culture. *Qualitative Research* 16(4): 359–374.

Woolgar S and Lezaun J (2013) The wrong bin bag: A turn to ontology in science and technology studies? *Social Studies of Science*: 43(3): 321-340.
World Health Organisation (WHO) (2022) Human resources for health, Seventy-Fifth World Health Assembly, Agenda item 15 (WHA75.17) 28-05-2022. Available at: https://apps.who.int/gb/ebwha/pdf_files/WHA75/A75_R17-en.pdf (accessed 29 May 2023).
Wyatt S (2003) Non-users also matter: The construction of users and non-users of the Internet. In: Oudshoorn N & Pinch T (eds) *How users matter: The coconstruction of users and technologies*. Cambridge, MA: MIT Press, pp. 67–81.
Wyatt S (2004) Danger! Metaphors at work in economics, geophysiology, and the internet. *Science, Technology, & Human Values* 29(2): 242–261.
Wyatt S (2021) Metaphors in critical Internet and digital media studies. *New Media & Society* 23(2): 406–416.
Wyatt S, Graham T, and Terranova T (2002) They came, they surfed, they went back to the beach: Conceptualising use and non-use of the Internet. In: Woolgar S (ed) *Virtual society? Technology, cyberbole, reality*. Oxford, England: Oxford University Press, pp. 23–40.
Zadvinskis IM, Garvey Smith J, and Yen PY (2018) Nurses' Experience with health information technology: Longitudinal qualitative study. *JMIR Medical Informatics* 6(2): e38.
Zetka GR (2001) Occupational divisions of labour and their technology politics: The case of surgical scopes and gastrointestinal medicine. *Social Forces* 79(4): 1495-1520.
Zulli D (2018) Capitalising on the look: insights into the glance, attention economy, and Instagram. *Critical Studies in Media Communication* 35(2): 137-150.

Appendix

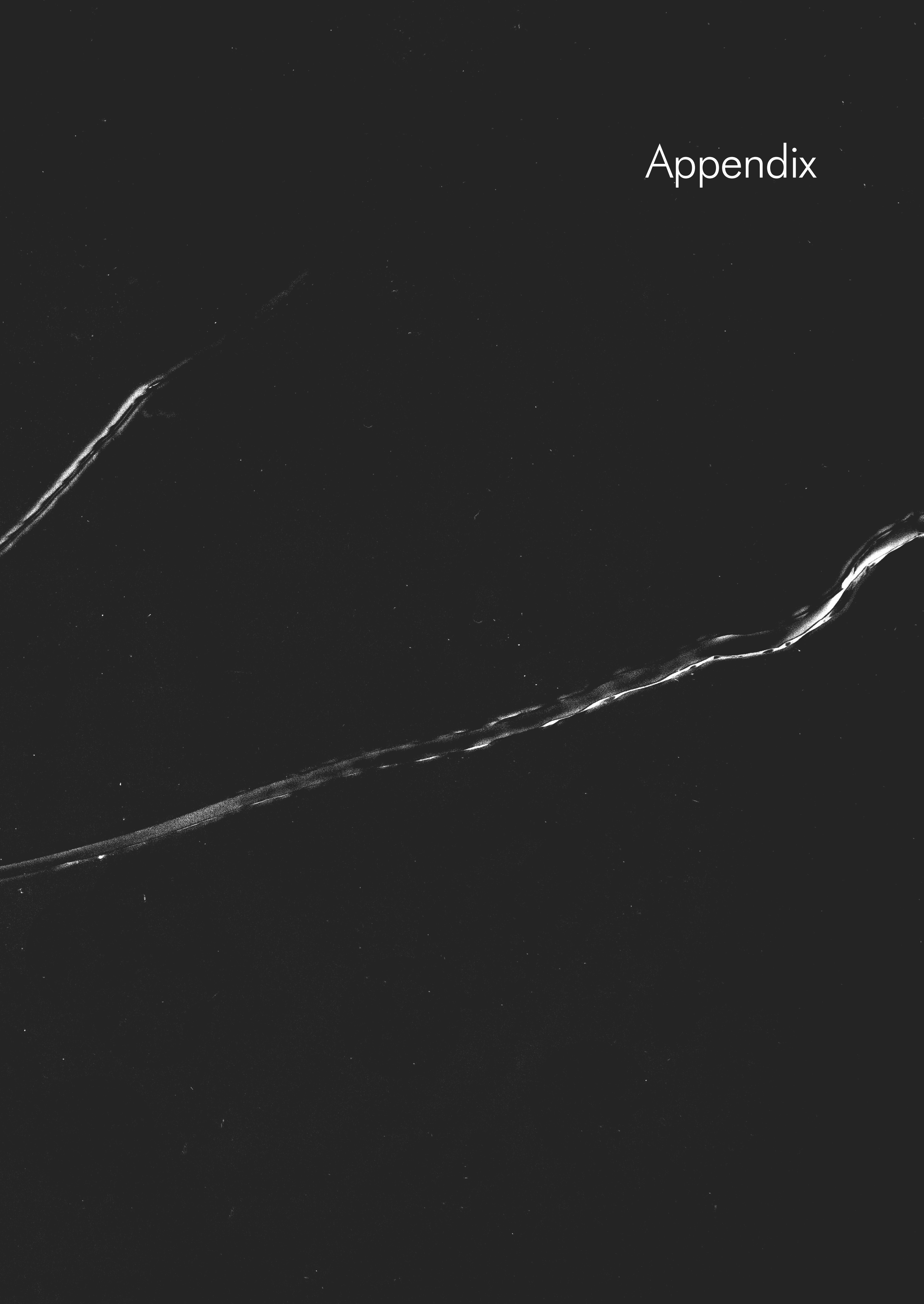


Table 1. Selected journals per discipline (* indicates interdisciplinary journals)

Selected STS journals	Selected sociology journals	Selected medicine journals
1. Science, Technology & Human Values;	1. Sociology: The Journal of the British Sociological Association;	1. JAMA — Journal of the American Medical Association;
2. Social Science & Medicine*;	2. Sociology of Health & Illness;	2. BMJ — British Medical Journal;
3. BioSocieties;	3. Journal of Health and Social Behavior;	3. Annals of Internal Medicine;
4. Social Studies of Science;	4. Social Theory & Health.	4. PLOS Medicine;
5. Technology in Society;	5. Information and Organisation;	5. BMC Medicine;
6. Science as Culture;	6. New Technology, Work & Employment;	6. Mayo Clinic Proceedings;
7. Big Data & Society;	7. Organisation Science;	7. Journal of Medical Internet Research
8. AI & Society;	8. Work and Occupations;	8. The Lancet Digital Health;
9. Philosophy & Technology;	9. Work, Employment & Society;	9. Journal of Healthcare Informatics Research;
10. Digital Health*.	10. Gender, Work & Organisation.	10. Journal of the American Medical Informatics Association.

Figures 1-3. Selection flow charts

Fig. 1. Selection flow chart for STS

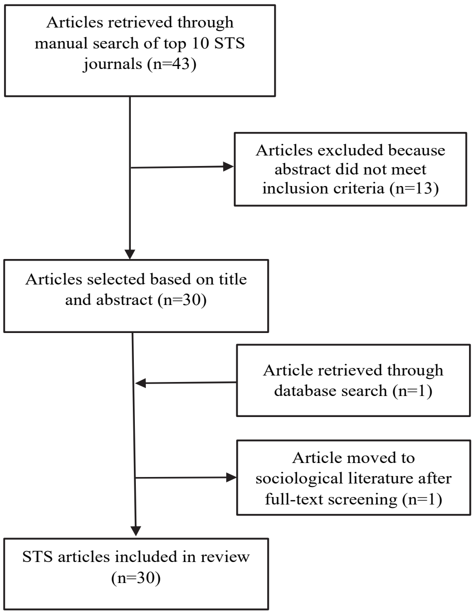


Fig. 2. Selection flow chart for sociology

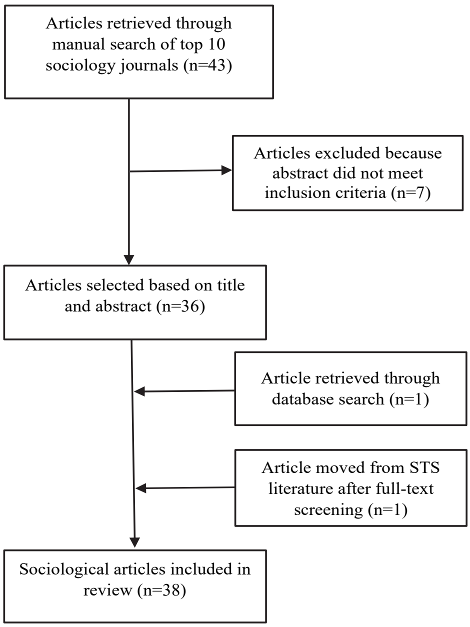


Fig. 3. Selection flow chart for medicine

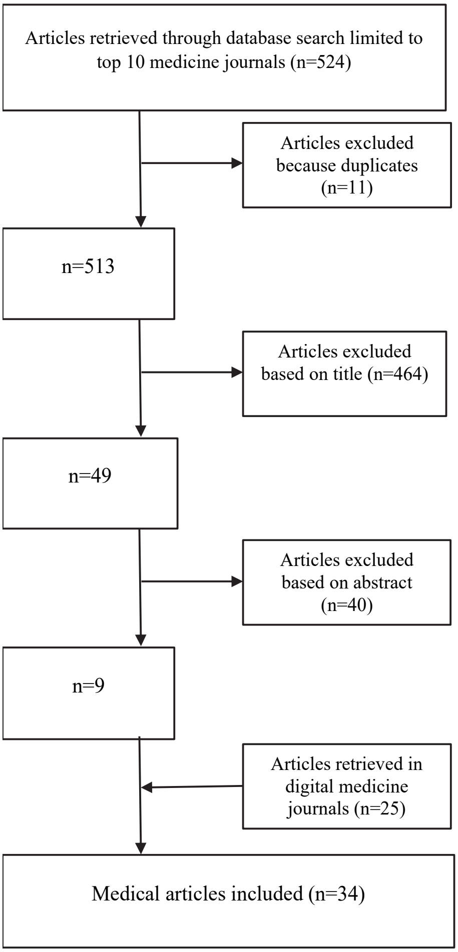


Table 3. Themes and metaphors per body of literature

	STS	Sociology	Medicine
Conceptualising the digitalisation of healthcare work	Network: the way a specific technology is embedded in professional practices is predicated on a process of negotiation at the intersection between several human and nonhuman actors. Materiality: the material characteristics of the technology steer the process of negotiation and the shape the network assumes.	Technology-in-practice: embedding of digital healthcare technologies in everyday professional work is an open empirical question. Steered innovation: digital healthcare technologies align in their functioning and requirements with managerial objectives imposed top-down.	Disattended promises: digital healthcare technologies have the potential to make healthcare work more efficient and meaningful, but often fail to deliver in practice. Good design: to deliver on their promises, technologies' design must fit with existing practices and meet professionals' needs.
Implications for individual professionals (nature of work /	Multiple and unpredictable: the introduction of new digital	Uneven: different professional groups experience	Rationalisation: digital healthcare technologies

practices)	healthcare technologies opens up possibilities for multiple reconfigurations. Aligned: ex-post, implications observed point towards an increased reliance on quantification and connected changes in the diagnostic process.	different implications, depending on the extent to which they are involved in innovation projects and their professional identities are embedded in technologies.	simplify workflows, make it easier to access information, can decrease error rates and alleviate documentation burden. Desktop medicine: digital healthcare technologies increase computer-based clerical work, working hours, and non-meaningful tasks.
Implications for patient-provider relationship	Invisible work: digital healthcare technologies enable more frequent communication, but they also require extra invisible work from	Invisible work: digital healthcare technologies create invisible work of explaining, reassuring, reminding, (re)establishing rapport with	Erosion of physician's authority: digital healthcare technologies try to involve patients more actively in their care trajectories; this can make

	professionals. **Task delegation**: technologies make possible to delegate to patients tasks formerly performed by professionals.	patients.	patients more inquisitive and hinder communication. **Invisible work**: physicians need to engage in extra sensory and emotion work.
Emotional implications	**Emotion work in the patient-provider relationship**: establishing rapport with patients takes extra emotion work.	**Distance and closeness in patient-provider relationship**: more emotion work is required to establish rapport with patients; easier to maintain emotional distance from patients.	**Burnout**: menial tasks take time away from the meaningful work of patient care, increasing exhaustion and likelihood of burnout.
Trade-offs of technological innovation	**Patient-provider information transfer**: qualitative and contextual information is not easily transferred and recorded in	**Interprofessional information transfer**: contextual, uncertain and subjective information is lost in	**Time and meaning**: menial tasks are crucial for digital healthcare technologies to function, but

	patient-provider interactions, and risks being lost.	interprofessional communication, or must be integrated informally. **Professional judgement**: technologies constrain the conditions under which professional judgements are formulated, resulting in less embodied, long-term and idiosyncratic knowledge of patients.	they take time and make healthcare work less meaningful.
Key metaphor	**Slime mould**: focus on open-ended exploration, information exchange and interconnections.	**Theatrical performances**: focus on the importance of unseen spaces and interactions in enabling performance.	**River engineering**: focus on directionality and unintended consequences of forceful changes.
Main insight from metaphor	Networks can be steered; their final configuration is	Digital healthcare technologies do not guarantee positive change	Goal-setting and visions are crucial in driving change

	not predictable, but interconnections are vital in ensuring an optimal one.	in healthcare work; invisible work is always necessary for coordination and information flow.	in healthcare organisations, but they can backfire when imposed forcefully.

Protocol for Critical Interpretive Synthesis

Guiding review question

How have the implications of digital technologies for healthcare professionals and organisations been conceptualised and described in the medicine, sociology, and STS literature, and what lessons can we learn by bringing together these insights?

Arguments for selecting these three domains

- Medicine: access to first-hand experiences and lively discussion of these topic (consequences of new technologies for work);
- Sociology: long-standing research and reflection on professions, thus perfectly positioned to study changes in professional roles;
- STS: analytical tools that allow to study the interplay between scientific, technological and social factors, with particular focus on technology and health (care).

Defining digital healthcare technologies

In order to specify the category of digital healthcare technologies in this paper, I build upon categories already defined within the health care domain. In 2019, the NHS-commissioned Topol Review devised a broad category of "digital healthcare technology," meant to span the technological innovations that, based on their current levels of institutional embedding and on their envisioning functioning in health care settings, are likely to have the biggest impact on the health care workforce. Although their focus is preeminently on the UK's NHS, in this paper I take up the Topol Review's (2019) category of digital health care technology, and focus on the following technologies:

- Genomics (both technologies related to reading the genome and genome editing);
- Digital medicine (i.e. telemedicine, smartphone apps, sensors and wearables for diagnostic monitoring, virtual and augmented reality);
- AI and robotics (i.e. speech recognition and natural language processing (NLP); automated image interpretation using AI; interventional and rehabilitative robotics; predictive analytics using AI).

My argument for adopting this definition goes beyond the Topol Review's claim that they are likely to have the most significant impact on health care professionals. Indeed, predictions and claims about the future are performative and contribute to the production of the future itself (van Lente 2012). Thus, although I do not aim to predict which technologies will be the most impactful for health care professionals, it is nonetheless justified to focus on the technologies that are presented as such in the professional and academic discourse. Adopting as a heuristic tool the Topol Review's (2019) broad yet highly impactful definition allows me to identify the most fertile areas of debate, within which the conceptualisation of technologies' implications is the most robust. My assumption here is that, given the wide resonance of reports such as the Topol Review, they are likely to set the terms of scholarly and professional conversations on these topics, thus fostering research and theorisation.

Selection of search terms

Search terms have been developed with the help of a library and information science specialist. They focus on bringing together the following dimensions:

- Technology — introduction, implementation, use, implication*, consequence*;
- Digital healthcare technologies[25] (thus particular emphasis on AI, Big Data, automation, robotics, health information technology, digital medicine);
- Healthcare professionals / workforce (and subcategories); workplace redesign; social innovation;
- Clinical / medical practice; health care (as context).

Search terms will be used to conduct manual search in top 10 STS, sociology and medical journals.

For medicine, manual searches will be complemented by a database search run in Embase and Web of Science. A Boolean search string has been developed with the help of a library and information science specialist to yield the most relevant medical articles on this issue.

[25] See definition above. Since genomic technologies still do not have widespread applications in clinical settings, I did not explicitly include them in the search terms. Nonetheless, articles referring to related technologies (e.g. precision medicine, genetic technologies) were included in the final selection when yielded by the database and manual searches.

Overarching criteria for journal selection

The manual search for STS and sociology journals will be conducted in the top 10 journals for each discipline. Conducting such a search in the STS and sociology fields is likely to yield the most relevant articles in terms of theory development and conceptualisations. Medicine differs in this respect, since top medical journals publish more research- and clinical-trials-based articles — which makes a database search the best option to retrieve relevant publications.

I will select the top 10 sociology and STS journals based on the following criteria, which combine impact factor with topical relevance:

- (Relatively) high 2018 Impact Factor, as listed on Web of Science's *Journal Citation Report*;
- Based on relevance, thematic journals can be included (e.g. medical and health sociology; critical data studies), as long as they are clearly sociological or STS in their approach;
- For medicine, more specialised journals can be included when the considered articles take a generalised (and thus applicable to other specialties) perspective.

Process of journal selection: Science and Technology Studies

Being an inherently interdisciplinary field, there is limited agreement on which journals qualify as STS ones. Therefore, no official rankings of STS journals are available, and the publications listed on relevant websites (STS wiki; Society for the Social Studies of Science) somewhat differ from one another. This complicates the process of selection. The top 10 STS journals were selected through the following steps:

- Gather titles of publications listed on 4S website and STS wiki;
- Review titles of publications listed under WoS Journal Citation Report's category *History and Philosophy of Science*;
- Combine titles available in those two lists by cross-referencing the ones not included in Scopus category by looking up their Impact Factor on WoS Journal Citation Report;
- Whenever a journal is unknown to the reviewer, and its domain of interest is not clear from its title, check *Aim and Scope* of the journal

on its website. Include journal if its description meets at least one of these three criteria:
  - Mentions Science (and Technology) Studies;
  - Refers to interdisciplinary social sciences approach to the study of science and technology;
  - Refers to philosophical and/or critical approach to the study of technology (or of a particular technology, e.g. AI) AND of its social consequences;
- If some journals are not listed in either the WoS Journal Citation Report's ranking or the STS websites, but are known to the reviewer to publish relevant STS-related research (e.g. *AI & Society*), look up their Impact Factor and add them to the list;
- Rank journals based on their Impact Factor;
- If Impact Factor is not available, include if journal publishes highly relevant research (e.g. *Big Data & Society*);
- Select the first 10 journals.

List of selected STS journals:

1. Science, Technology & Human Values;
2. Social Science & Medicine;
3. BioSocieties;
4. Social Studies of Science;
5. Technology in Society;
6. Science as Culture;
7. Big Data & Society;
8. AI & Society;
9. Philosophy & Technology;
10. Digital Health.

Since *Social Science & Medicine* and *Digital Health* are intrinsically interdisciplinary journals, the scope of which exceeds unequivocally STS work, it was decided to examine their articles individually and to include them in the STS pool (or to move them to the sociology literature) on a case-by-case basis, according to the authors' framing of their specific contributions.

Process of journal selection: Sociology

Unlike STS, *Sociology* is a category in WoS Journal Citation Report. However, included in this category are also thematic journals of little or no interest. Publications in medical or health sociology are also left out of this category, and listed under *Social Sciences, Biomedical* instead (a category under which publications focusing on ethics and behavioural sciences are also listed). The top 10 journals for sociology were thus selected through the following steps:

- Gather titles first 10 publications listed under WoS Journal Citation Report category *Sociology*;
- Supplement list with relevant (i.e. sociological) titles listed under WoS Journal Citation Report category *Social Sciences, Biomedical*. This allows to include publications in the field of medical and health sociology;
- Compare the list with titles in other sociology rankings available online (5-Year Journal Impact Factor (TM) in ISI Journal Citation Report; h5-index) and add to the list possible missing journals, looking up their Impact Factor;
- Rule out all the non-sociological titles;
- When a journal is not included in the selected WoS Journal Citation Report categories, is not known to the reviewer and its domain of interest is not clear from its title, check *Aim and Scope* of the journal on its website. Include journal if its description meets at least one of these criteria:
    - Focuses on sociology and/or medical sociology;
    - Mainly publishes research or review articles (not methodology);
    - If the focus of the journal is thematic and exceeds medical sociology, the main domain of interest significantly overlaps with review question (e.g. thematic sociology journals on technology, IT-systems, etc.).
- If some journals are not listed in either the Scopus ranking or the STS websites, but are known to the reviewer to publish relevant sociological research, add them to the list;
- Rank journals based on their Impact Factor;
- Select the first 10 journals.

- If some journals yield no relevant results, substitute them with journals focusing on work and organisational sociology (as listed on https://www.scimagojr.com/journalrank.php?category=1407)

List of selected sociology journals:	Journals in sociology of work and organisations:	Final selection:
1. American Sociological Review [no results]; 2. Annual Review of Sociology [no results]; 3. American Journal of Sociology [no results]; 4. British Journal of Sociology [no results]; 5. Qualitative Research [no results]; 6. Sociology: The Journal of the British Sociological Association; 7. European Sociological Review [no results]; 8. Sociology of Health & Illness;	1. Organisation Studies [no results]; 2. Organisation Science; 3. Work and Occupations; 4. Work Employment and Society; 5. Gender Work and Organisation; 6. Information and Organisation; 7. New Technology Work and Employment.	1. Sociology: The Journal of the British Sociological Association; 2. Sociology of Health & Illness; 3. Journal of Health and Social Behavior; 4. Social Theory & Health. 5. Information and Organisation; 6. New Technology, Work & Employment; 7. Organisation Science; 8. Work and Occupations; 9. Work, Employment & Society; 10. Gender, Work & Organisation.

9. Journal of Health and Social Behavior; 10. Social Theory & Health.		

Process of journal selection: Medicine

Alongside the creation of a Boolean search string to conduct a database search, the top 10 journals for medicine were also selected. Combining these two search strategies allows, on the one hand, to select the most topical articles based on their use of key words and, on the other, to ensure that the review includes the articles likely to have the broadest resonance in the field, thus considerable shaping the discussion concerning the implications of digital healthcare technologies for professional roles in health care. The top 10 journals for medicine were selected base on WoS's Journal Citation Report category *General and Internal Medicine*.
List of selected medicine journals:

11. New England Journal of Medicine;
12. The Lancet;
13. JAMA — Journal of the American Medical Association;
14. BMJ — British Medical Journal;
15. Annals of Internal Medicine;
16. PLOS Medicine;
17. BMC Medicine;
18. Cochrane Database of Systematic Reviews;
19. Mayo Clinic Proceedings;
20. Canadian Medical Association Journal.

To make the article selection more thematically specific, and also in consideration of the fact that several of the aforementioned publications yielded no results, this list was supplemented with additional, thematic publications focusing specifically on digital health and medicine. Such publications were selected based both on recommendations of experts in the field of digital medicine, and through the list of digital medicine journals published by Springer (https://www.springer.com/gp/campaigns/digital-medicine). The following publications were thus identified:

1. Journal of Medical Internet Research (and its cognate publications JMIR mHealth and uHealth, JMIR Mental Health, JMIR Medical Informatics, JMIR Public Health and Surveillance, JMIR Human Factors);
2. Journal of Healthcare Informatics Research;
3. Journal of the American Medical Informatics Association.

The final selection thus consisted of the following journals:

21. JAMA — Journal of the American Medical Association;
22. BMJ — British Medical Journal;
23. Annals of Internal Medicine;
24. PLOS Medicine;
25. BMC Medicine;
26. Mayo Clinic Proceedings;
27. Journal of Medical Internet Research
28. The Lancet Digital Health;
29. Journal of Healthcare Informatics Research;
30. Journal of the American Medical Informatics Association.

Literature search

1. The identified search terms will be used to conduct manual searches within the identified top 10 medicine, STS and sociology journals. For medicine, a Boolean search string will also be run in *Embase* and *Web of Science* to retrieve the most relevant articles;
2. The articles published more than 20 years ago (i.e. before 2000) will be excluded;
3. The titles of the articles yielded through this search will be scanned;
4. An article's abstract will be read if its title meets the following criteria:
    a. Refers to medicine and/or health or some variant thereof

    AND

    b. Refers to digital healthcare technologies (either in general or to a specific one);
5. An article will be included in the pool of selected articles if its abstract meets the following criteria:

   a. Mainly or significantly focuses on health care professionals;
   b. Significantly focuses on digital healthcare technologies (either in general or on a specific one);
   c. Focuses on technologies used in clinical practice (broadly construed, including technologies — such as eHealth ones — the use of which goes beyond the physical space of 'the clinic');
   d. Is based on empirical research or on a review of relevant literature (i.e. no opinion pieces or commentaries);
   e. Establishes a link between technology use and change in work or professional practices;
   f. If focused on clinician-patient communication, clearly discusses implications for clinicians' role.
6. Exclusion criteria:
   a. No clear focus on technologies;
   b. Focus on technologies that are not data- and information-related;
   c. No clear focus on health care professionals;
   d. No clear focus on medical practice (e.g. focus on training or education);
   e. If research article or review on professional views of new technologies, does not focus on their expectations or experiences of changes in their work or role;
   f. If focused on clinician-patient communication, does not discuss implications for clinicans' role.
7. Whenever the first reviewer is unsure about whether an article qualifies for inclusion, title and abstract will also be checked by the second reviewer, and the article's inclusion will be jointly discussed.

First search results: STS

A manual search was conducted in the selected 10 STS journals. Articles were selected following the aforementioned criteria. This operation yielded a total of 29 articles. One of these articles (published on *Social Science & Medicine*) was framed as sociological, and thus moved to the sociological literature, bringing the number of STS articles included through manual search down to

28.
To check the exhaustiveness of this manual search, the results were compared with articles retrieved in STS-related journals that did not make it to the final selection (*Medicine, Health Care & Philosophy*; *New Genetics & Society*). This operation yielded 2 additional relevant articles, bringing the total of STS articles to 30.

First search results: Sociology

A manual search was conducted in the selected 10 journals. Articles were selected following the aforementioned criteria. This operation yielded a total of 36 articles.
To check the exhaustiveness of this search, the following search string was run in Scopus, limiting results by discipline (social science):
ALL (innovation* OR "machine learning" OR "artificial intelligence" OR "information technology" OR "algorithm*" OR robot* OR "decision-support system*") AND ALL (healthcare OR "health care" OR "health-care" OR hospital* OR clinic*) AND TITLE-ABS-KEY (profession* OR clinician* OR doctor* OR nurse* OR physician* OR practi* OR “social innovation” OR redesign) AND PUBYEAR > 2000 AND SUBJAREA (soci) AND (LIMIT-TO (EXACTSRCTITLE , "Social Science And Medicine") OR LIMIT-TO (EXACTSRCTITLE , "Sociology Of Health And Illness") OR LIMIT-TO (EXACTSRCTITLE , "Behaviour And Information Technology") OR LIMIT-TO (EXACTSRCTITLE , "Information Technology And People") OR LIMIT-TO (EXACTSRCTITLE , "Social Studies Of Science") OR LIMIT-TO (EXACTSRCTITLE , "Science Technology And Human Values") OR LIMIT-TO (EXACTSRCTITLE , "Health Care Analysis") OR LIMIT-TO (EXACTSRCTITLE , "New Genetics And Society") OR LIMIT-TO (EXACTSRCTITLE , "Scientometrics"))
However, no additional sociological articles were retrieved.
To further check that no relevant perspectives were left out, the results were compared with articles retrieved in social science journals that did not make it to the final selection (*Health Care Analysis*; *Health*; *New Media & Society*; *Information, Communication & Society*). This operation yielded 1 additional relevant article which, summed to the additional article moved from the pool of STS articles, brought the total number of sociological articles to 38.

First search results: Medicine

In order to ensure relevant articles from top-ranking publications also included, a query string was devised to limit the search to the top 10 *General and Internal Medicine* journals (as listed on WoS's Journal Citation Report). The following string was run through *Embase*:

('ai' OR 'artificial intelligence'/exp OR 'artificial intelligence' OR 'big data'/exp OR 'big data' OR 'machine learning'/exp OR 'machine learning' OR 'information technology'/exp OR 'information technology') AND (profession* OR 'physician'/exp OR physician OR 'clinician'/exp OR clinician OR role* OR nurse* OR doctor* OR 'medical practice'/exp OR 'medical practice' OR 'workforce'/exp OR workforce OR 'workplace'/exp OR workplace) AND ('healthcare'/exp OR healthcare OR 'health care'/exp OR 'health care' OR 'hospital'/exp OR hospital OR clinic*) AND (jama:jt OR nejm:jt OR lancet:jt OR bmj:jt OR 'annals of internal medicine':jt OR 'plos medicine':jt OR 'bmc medicine':jt OR 'cochrane database of systematic reviews':jt OR 'mayo clinic':jt)

This search yielded 521 articles.

- Duplicates (n=5);

n=516

- Articles published before 2000 (n=7);

n=510;

- Title screened:
    - articles not addressing implications of technologies for professionals in health care (n=464);

n=46;

- Abstract screened (n=40):
    - not about digital healthcare technology (n=1);
    - not research or review article (n=25);
    - not mainly focused on professionals (n=6);
    - not mainly focused on changing professional roles (n=7);
    - not mainly focused on technology (=1)

n=6.

Thus, after applying the aforementioned exclusion criteria, this search yielded a total of 5 included articles.

In order to ensure relevant articles were not left out of the selection, a Boolean search string was devised with the help of a library information specialist in order to yield the most relevant articles based on their use of mesh terms. The following search string was run through both *Embase* and *Web of Science* (adapted):
('nursing role'/mj/de OR 'role change'/mj/de OR (((profession* OR doctor* OR nurs* OR physicist* OR physiotherapist* OR paramedic* OR practitioner* OR Medical-Assistant* OR pharmacist* OR workforce OR worker* OR physician*) NEAR/6 (role* OR implication*)) OR ((role OR roles) NEAR/3 (chang* OR expectation*))):ti) AND ('health technology'/mj/de OR 'artificial intelligence'/mj/exp OR 'machine learning'/mj/de OR 'virtual reality'/mj/de OR 'innovation'/mj/de OR 'big data'/mj/de OR 'personalised medicine'/mj/de OR 'computer assisted therapy'/mj/de OR robotics/mj/exp OR 'robot assisted surgery'/mj/de OR 'biomedical engineering'/mj/de OR 'mobile application'/mj/de OR 'mobile phone'/mj/de OR (((new) NEAR/3 technolog*) OR ((medical OR biomedic* OR Health OR Healthcare OR Health-care OR nurs) NEAR/3 informatic*) OR innovation* OR (artificial* NEAR/3 intelligen*) OR ((machine OR deep) NEXT/1 learning) OR big-data OR Health-Information-Management OR virtual-realit* OR high-tech* OR (personali* NEAR/3 medicin*) OR ((computer*) NEAR/3 assisted NEAR/3 (treat* OR therap* OR surg*)) OR Biomedical-Engineer* OR robot* OR (mobile NEAR/3 (phone* OR telephone* OR application*))):ti)
This search yielded a total of 354 articles (Embase: 208; WoS: 146), which were reduced to 302 (Embase: 205; WoS 97) after checking for duplicates.

- no full access (n=106);

n=196;

- articles published in before 2000 (n=64);

n=132;

- articles published in non-medical journals (n= 64);

n=68;

- title screened:
    - articles not addressing implications of technologies for professionals in health care (n=52);

n=16;

- Abstract screened (n=13):
    - not research or review article (n=11);
    - not mainly focused on changing professional roles (n=2);

n=3.

Thus, after applying the aforementioned exclusion criteria, this search yielded a total of 3 included articles. These were combined with the previously selected publications, bringing the total of selected medical article to 9.

For the sake of symmetry with searches methods implemented for sociology and STS literatures, the database search was supplemented with a manual search conducted in more thematic medicine journals, with a specific focus on digital health (see above). The general inclusion and exclusion criteria were applied in this manual search, which yielded a total of 25 articles, bringing the total number of included medical articles to 34.

Included articles per discipline

STS

Abrishami, P. Boer, A. & Horstman, K. (2014). Understanding the adoption dynamics of medical innovations: Affordances of the da Vinci robot in the Netherlands. *Social Science & Medicine 117*, pp. 125–133.

Andreassen, H. K. Kjekshus, L. E. & Tjora, A. (2015). Survival of the project: A case study of ICT innovation in health care. *Social Science & Medicine 132*, 62–69.

Bjerring, J. C. & Busch, J. (2020). Artificial Intelligence and Patient-Centered Decision-Making. *Philosophy & Technology*.

Burri, R. V. (2008). Doing Distinctions: Boundary Work and Symbolic Capital in Radiology. *Social Studies of Science 38*(1), 35–62.

Burri, R. V. (2013). Visual Power in Action: Digital Images and the Shaping of Medical Practices. *Science as Culture 22*(3), 367–387.

Coeckelbergh, M. (2013). E-care as craftsmanship: Virtuous work, skilled engagement, and information technology in health care. *Medicine, Health Care and Philosophy 16*(4), 807–816.

Danesi, G. Pralong, M. Panese, F. Burnand, B. & Grossen, M. (2020). Techno-social reconfigurations in diabetes (self-) care. *Social Studies of Science 50*(2), 198–220.

Galetsi, P. Katsaliaki, K. & Kumar, S. (2019). Values, challenges and future directions of big data analytics in healthcare: A systematic review. *Social Science & Medicine 241*, 112533.

Garrety, K. McLoughlin, I. Wilson, R. Zelle, G. & Martin, M. (2014). National electronic health records and the digital disruption of moral orders. *Social Science & Medicine 101*, pp. 70–77.

Greenhalgh, T. & Stones, R. (2010). Theorising big IT programmes in healthcare: Strong structuration theory meets actor-network theory. *Social Science & Medicine 70*(9), pp. 1285–1294.

Joyce, K. (2005). Appealing Images: Magnetic Resonance Imaging and the Production of Authoritative Knowledge. *Social Studies of Science 35*(3), pp. 437–462.

Juciute, R. (2009). ICT implementation in the health-care sector: Effective stakeholders' engagement as the main precondition of change sustainability. *AI & Society 23*(1), pp. 131–137.

Lehoux, P. Sicotte, C. Denis, J.-L. Berg, M. & Lacroix, A. (2002). The theory of use behind telemedicine: How compatible with physicians' clinical routines? *Social Science & Medicine* 54, pp. 889-940.

Lehoux, P. Poland, B. Daudelin, G. Holmes, D. & Andrews, G. (2008). Displacement and Emplacement of Health Technology: Making Satellite and Mobile Dialysis Units Closer to Patients? *Science, Technology, & Human Values 33*(3), pp. 364–392.

Maslen, S. (2017). Layers of sense: The sensory work of diagnostic sensemaking in digital health. *Digital Health 3* 205520761770910.

May, C. Gask, L. Atkinson, T. Ellis, N. Mair, F. & Esmail, A. (2001). Resisting and promoting new technologies in clinical practice: The case of telepsychiatry. *Social Science & Medicine 52*(12), pp. 1889–1901.

Moerenhout, T. Fischer, G. S. & Devisch, I. (2020). The elephant in the room: A postphenomenological view on the electronic health record and its impact on the clinical encounter. *Medicine, Health Care and Philosophy 23*(2), pp. 227–236.

Mort, M. May, C. R. & Williams, T. (2003). Remote Doctors and Absent Patients: Acting at a Distance in Telemedicine? *Science, Technology, & Human Values 28*(2), pp. 274–295.

Nicolini, D. (2006). The work to make telemedicine work: A social and articulative view. *Social Science & Medicine 62*(11), pp. 2754–2767.

Nieboer, M. E. van Hoof, J. van Hout, A. M. Aarts, S. & Wouters, E. J. M. (2014). Professional values, technology and future health care: The view of health care professionals in the Netherlands. *Technology in Society 39*, pp. 10–17.

Oudshoorn, N. (2012). How places matter: Telecare technologies and the changing spatial dimensions of healthcare. *Social Studies of Science 42*(1), pp. 121–142.

Petrakaki, D. Barber, N. & Waring, J. (2012). The possibilities of technology in shaping healthcare professionals: (Re/De-)Professionalisation of pharmacists in England. *Social Science & Medicine 75*(2), pp. 429–437.

Pols, J. (2011). Wonderful Webcams: About Active Gazes and Invisible Technologies. *Science, Technology, & Human Values 36*(4), pp. 451–473.

Reich, A. (2012). Disciplined doctors: The electronic medical record and physicians' changing relationship to medical knowledge. *Social Science & Medicine 74*(7), pp. 1021–1028.

Rystedt, H. Ivarsson, J. Asplund, S. Johnsson, Å. A. & Båth, M. (2011). Rediscovering radiology: New technologies and remedial action at the worksite. *Social Studies of Science 41*(6), pp. 867–891.

Spatar, D. Kok, O. Basoglu, N. & Daim, T. (2019). Adoption factors of electronic health record systems. *Technology in Society*, *58*, 101144.

Trondsen, M. V. Tjora, A. Broom, A. & Scambler, G. (2018). The symbolic affordances of a video-mediated gaze in emergency psychiatry. *Social Science & Medicine 197*, pp. 87–94.

Turja, T. Rantanen, T. & Oksanen, A. (2019). Robot use self-efficacy in healthcare work (RUSH): Development and validation of a new measure. *AI & society 34*(1), pp. 137–143.

Vallor, S. (2011). Carebots and Caregivers: Sustaining the Ethical Ideal of Care in the Twenty-First Century. *Philosophy & Technology 24*(3) 251–268.

Winthereik, B. R. van der Ploeg, I. & Berg, M. (2007). The Electronic Patient Record as a Meaningful Audit Tool: Accountability and Autonomy in General Practitioner Work. *Science, Technology, & Human Values 32*(1), pp. 6–25.

Sociology

Andreassen, H. K. Dyb, K. May, C. R. Pope, C. J. & Warth, L. L. (2018). Digitised patient–provider interaction: How does it matter? A qualitative meta-synthesis. *Social Science & Medicine 215*, pp. 36–44.

Bailey, S. Pierides, D. Brisley, A. Weisshaar, C. & Blakeman, T. (2020). Dismembering organisation: The coordination of algorithmic work in healthcare. *Current Sociology 68*(4), pp. 546–571.

Bar-Lev, S. (2015). The politics of healthcare informatics: Knowledge management using an electronic medical record system. *Sociology of Health & Illness 37*(3), pp. 404–421.

Barrett, M. Oborn, E. Orlikowski, W. J. & Yates, J. (2012). Reconfiguring Boundary Relations: Robotic Innovations in Pharmacy Work. *Organisation Science 23*(5), pp. 1448–1466.

Beaudevin, C. Peerbaye, A. & Bourgain, C. (2019). 'It has to become true genetics': Tumour genetics and the division of diagnostic labour in the clinic. *Sociology of Health & Illness 41*(4), pp. 643–657.

Bloomfield, B. P. & McLean, C. (2003). Beyond the walls of the asylum: Information and organisation in the provision of community mental health services. *Information and Organisation 13*(1), pp. 53–84.

Campbell, M. L. & Rankin, J. M. (2017). Nurses and electronic health records in a Canadian hospital: Examining the social organisation and programmed use of digitised nursing knowledge. *Sociology of Health & Illness 39*(3), pp. 365–379.

Casper, M. J. & Morrison, D. R. (2010). Medical Sociology and Technology:

Critical Engagements. *Journal of Health and Social Behavior 51*(1_suppl), pp. S120–S132.

Constantinides, P. & Barrett, M. (2006). Negotiating ICT development and use: The case of a telemedicine system in the healthcare region of Crete. *Information and Organisation 16*(1), pp. 27–55.

Daskalopoulou, A. Palmer, M. Keeling, K. & Pritchard Jones, R. (2019). Discretionary technology bootlegging tensions in institutional healthcare work. *New Technology, Work and Employment 34*(1), pp. 73–89.

David, G. C. Garcia, A. C. Rawls, A. W. & Chand, D. (2009). Listening to what is said - transcribing what is heard: The impact of speech recognition technology (SRT) on the practice of medical transcription (MT). *Sociology of Health & Illness 31*(6), pp. 924–938.

Dupret, K. (2017). Working around technologies-invisible professionalism? *New Technology, Work and Employment 32*(2), pp. 174–187.

Engesmo, J. & Tjora, A. H. (2006). Documenting for whom? A symbolic interactionist analysis of technologically induced changes of nursing handovers. *New Technology, Work and Employment 21*(2), pp. 176–189.

Fältholm, Y. & Jansson, A. (2008). Telephone advisory services-nursing between organisational and occupational professionalism. *New Technology, Work and Employment 23*(1–2), pp. 17–29.

Findlay, P. Lindsay, C. McQuarrie, J. Bennie, M. Corcoran, E. D. & Van Der Meer, R. (2017). Employer Choice and Job Quality: Workplace Innovation, Work Redesign, and Employee Perceptions of Job Quality in a Complex Health-Care Setting. *Work and Occupations 44*(1), pp. 113–136.

Håland, E. (2012). Introducing the electronic patient record (EPR) in a hospital setting: Boundary work and shifting constructions of professional identities: EPR boundary work and constructions of identity. *Sociology of Health & Illness 34*(5), pp. 761–775.

Håland, E. & Melby, L. (2015). Negotiating technology-mediated interaction in health care. *Social Theory & Health 13*(1), pp. 78–98.

Halford, S. Obstfelder, A. & Lotherington, A.-T. (2010). Changing the record: The inter-professional, subjective and embodied effects of electronic patient records: Changing the record. *New Technology, Work and Employment 25*(3), pp. 210–222.

Heath, C. Luff, P. & Svensson, M. S. (2003). Technology and medical practice: Technology and medical Practice. *Sociology of Health & Illness 25*(3), pp. 75–96.

Højlund, H. & Villadsen, K. (2020). Technologies in caregiving: Professionals' strategies for engaging with new technology. *New Technology, Work and Employment*, 12161.

Lupton, D. & Maslen, S. (2017). Telemedicine and the senses: A review. *Sociology of Health & Illness 39*(8), pp. 1557–1571.

Maiers, C. (2017). Analytics in action: Users and predictive data in the neonatal intensive care unit. *Information, Communication & Society 20*(6), pp. 915–929.

Mueller, F. Valsecchi, R. Smith, C. Gabe, J. & Elston, M. A. (2008). 'We are nurses, we are supposed to care for people': Professional values among nurses in NHS Direct call centres. *New Technology, Work and Employment 23*(1–2), pp. 2–16.

Nickelsen, N. C. M. (2019). The infrastructure of telecare: Implications for nursing tasks and the nurse-doctor relationship. *Sociology of Health & Illness 41*(1), pp. 67–80.

Oudshoorn, N. (2008). Diagnosis at a distance: The invisible work of patients and healthcare professionals in cardiac telemonitoring technology. *Sociology of Health & Illness 30*(2), pp. 272–288.

Oudshoorn, N. (2009). Physical and digital proximity: Emerging ways of health care in face-to-face and telemonitoring of heart-failure patients. *Sociology of Health & Illness 31*(3), pp. 390–405.

Peiris, D. Usherwood, T. Weeramanthri, T. Cass, A. & Patel, A. (2011). New tools for an old trade: A socio-technical appraisal of how electronic decision support is used by primary care practitioners: Use of electronic decision support in primary care. *Sociology of Health & Illness 33*(7), pp. 1002–1018.

Petrakaki, D. Klecun, E. & Cornford, T. (2016). Changes in healthcare professional work afforded by technology: The introduction of a national electronic patient record in an English hospital. *Organisation 23*(2), pp. 206–226.

Petrakaki, D. & Kornelakis, A. (2016). 'We can only request what's in our protocol': Technology and work autonomy in healthcare. *New Technology, Work and Employment 31*(3), pp. 223–237.

Pols, J. (2010). The Heart of the Matter. About Good Nursing and Telecare. *Health Care Analysis 18*(4), pp. 374–388.

Pols, J. & Willems, D. (2011). Innovation and evaluation: Taming and unleashing telecare technology: Taming and unleashing telecare technology. *Sociology of Health & Illness 33*(3), pp. 484–498.

Prout, A. (1996). Actor-network theory, technology and medical sociology: An illustrative analysis of the metered dose inhaler. *Sociology of Health &*

Illness 18(2), pp. 198-219.
Reed, K. Kochetkova, I. & Molyneux-Hodgson, S. (2016). 'You're looking for different parts in a jigsaw': Foetal MRI (magnetic resonance imaging) as an emerging technology in professional practice. *Sociology of Health & Illness 38*(5), pp. 736–752.
Russell, B. (2012). Professional call centres, professional workers and the paradox of the algorithm: The case of telenursing. *Work, Employment and Society 26*(2), pp. 195–210.
Schwennesen, N. (2019). Algorithmic assemblages of care: Imaginaries, epistemologies and repair work. *Sociology of Health & Illness 41*(S1), pp. 176–192.
Timmermans, S. & Berg, M. (2003). The practice of medical technology. *Sociology of Health & Illness 25*(3), pp. 97–114.
Van den Broek, D. (2017). Perforated body work: The case of tele-nursing. *Work, Employment and Society 31*(6), pp. 904–920.
Wilson, R. Baines, S. Martin, M. Richter, P. McLoughlin, I. & Maniatopoulos, G. (2017). Who is telecaring whom? Exploring the total social organisation of care work in an Italian municipality. *New Technology, Work and Employment 32*(3), pp. 268–282.

Medicine

Alvandi, E. O. Van Doorn, G. & Symmons, M. (2019). Emotional Awareness and Decision-Making in the Context of Computer-Mediated Psychotherapy. *Journal of Healthcare Informatics Research 3*(3), pp. 345–370.
Blease, C. Kaptchuk, T. J. Bernstein, M. H. Mandl, K. D. Halamka, J. D. & DesRoches, C. M. (2019). Artificial Intelligence and the Future of Primary Care: Exploratory Qualitative Study of UK General Practitioners' Views. *Journal of Medical Internet Research 21*(3), p. e12802.
Bloem, B. R. & Munneke, M. (2014). Revolutionising management of chronic disease: The ParkinsonNet approach. *BMJ 348*, pp. g1838–g1838.
Bourla, A. Ferreri, F. Ogorzelec, L. Peretti, C.-S. Guinchard, C. & Mouchabac, S. (2018). Psychiatrists' Attitudes Toward Disruptive New Technologies: Mixed-Methods Study. *JMIR Mental Health 5*(4), p. e10240.
Buchan, T. A. Ross, H. J. McDonald, M. Billia, F. Delgado, D. Duero Posada, J. G. Luk, A. Guyatt, G. H. & Alba, A. C. (2020). Physician Judgement

vs Model-Predicted Prognosis in Patients with Heart Failure. *Canadian Journal of Cardiology 36*(1), pp. 84–91.

Callen, J. L. Westbrook, J. I. & Braithwaite, J. (2006). The Effect of Physicians' Long-term Use of CPOE on Their Test Management Work Practices. *Journal of the American Medical Informatics Association 13*(6), pp. 643–652.

Cresswell, K. Cunningham-Burley, S. & Sheikh, A. (2018). Health Care Robotics: Qualitative Exploration of Key Challenges and Future Directions. *Journal of Medical Internet Research 20*(7), p. e10410.

Fisher Wilson, J. (2009). Making electronic health records meaningful. *Annals of Internal Medicine 151*(4), pp. 293-296.

Glasgow, M. E. S. Colbert, A. Viator, J. & Cavanagh, S. (2018). The Nurse-Engineer: A New Role to Improve Nurse Technology Interface and Patient Care Device Innovations. *Journal of Nursing Scholarship 50*(6), pp. 601–611.

Grassl, N. Nees, J. Schramm, K. Spratte, J. Sohn, C. Schott, T. C. & Schott, S. (2018). A Web-Based Survey Assessing the Attitudes of Health Care Professionals in Germany Toward the Use of Telemedicine in Pregnancy Monitoring: Cross-Sectional Study. *JMIR MHealth and UHealth 6*(8), p. e10063.

Grünloh, C. Cajander, Å. & Myreteg, G. (2016). "The Record is Our Work Tool!"—Physicians' Framing of a Patient Portal in Sweden. *Journal of Medical Internet Research 18*(6), p. e167.

Hains, I. M. Georgiou, A. & Westbrook, J. I. (2012). The impact of PACS on clinician work practices in the intensive care unit: A systematic review of the literature. *Journal of the American Medical Informatics Association 19*(4), pp. 506–513.

Kurahashi, A. M. Weinstein, P. B. Jamieson, T. Stinson, J. N. Cafazzo, J. A. Lokuge, B. Morita, P. P. Cohen, E. Rapoport, A. Bezjak, A. & Husain, A. (2016). In the Loop: The Organisation of Team-Based Communication in a Patient-Centered Clinical Collabouration System. *JMIR Human Factors 3*(1), p. e12.

Lærum, H. Karlsen, T. H. & Faxvaag, A. (2003). Effects of Scanning and Eliminating Paper-based Medical Records on Hospital Physicians' Clinical Work Practice. *Journal of the American Medical Informatics Association 10*(6), pp. 588–595.

Li, R. C. Wang, J. K. Sharp, C. & Chen, J. H. (2019). When order sets do not align with clinician workflow: Assessing practice patterns in the electronic health record. *BMJ Quality & Safety 28*, pp. 987-996.

Mather, C. Cummings, E. & Gale, F. (2019). Nurses as Stakeholders in the

Adoption of Mobile Technology in Australian Health Care Environments: Interview Study. *JMIR Nursing 2*(1), p. e14279.

Maxson, E. Jain, S. Kendall, M. Mostashari, F. & Blumenthal, D. (2010). The Regional Extension Center Program: Helping Physicians Meaningfully Use Health Information Technology. *Annals of Internal Medicine 153*(10), p. 666.

Mayer, M. A. Rodríguez Blanco, O. & Torrejon, A. (2019). Use of Health Apps by Nurses for Professional Purposes: Web-Based Survey Study. *JMIR MHealth and UHealth 7*(11), p. e15195.

Mercer, K. Burns, C. Guirguis, L. Chin, J. Dogba, M. J. Dolovich, L. Guénette, L. Jenkins, L. Légaré, F. McKinnon, A. McMurray, J. Waked, K. & Grindrod, K. A. (2018). Physician and Pharmacist Medication Decision-Making in the Time of Electronic Health Records: Mixed-Methods Study. *JMIR Human Factors 5*(3), p. e24.

O'Malley, A. S. Draper, K. Gourevitch, R. Cross, D. A. & Scholle, S. H. (2015). Electronic health records and support for primary care teamwork. *Journal of the American Medical Informatics Association 22*(2), pp. 426–434.

Petersson, L. & Erlingsdóttir, G. (2018a). Open Notes in Swedish Psychiatric Care (Part 1): Survey Among Psychiatric Care Professionals. *JMIR Mental Health 5*(1), p. e11.

Petersson, L. & Erlingsdóttir, G. (2018b). Open Notes in Swedish Psychiatric Care (Part 2): Survey Among Psychiatric Care Professionals. *JMIR Mental Health 5*(2), p. e10521.

Prgomet, M. Georgiou, A. & Westbrook, J. I. (2009). The Impact of Mobile Handheld Technology on Hospital Physicians' Work Practices and Patient Care: A Systematic Review. *Journal of the American Medical Informatics Association 16*(6), pp. 792–801.

Schuster, R. Pokorny, R. Berger, T. Topooco, N. & Laireiter, A.-R. (2018). The Advantages and Disadvantages of Online and Blended Therapy: Survey Study Amongst Licensed Psychotherapists in Austria. *Journal of Medical Internet Research 20*(12), p. e11007.

Sieja, A. Markley, K. Pell, J. Gonzalez, C. Redig, B. Kneeland, P. & Lin, C.-T. (2019). Optimisation Sprints: Improving Clinician Satisfaction and Teamwork by Rapidly Reducing Electronic Health Record Burden. *Mayo Clinic Proceedings 94*(5), pp. 793–802.

Strand, M. Gammon, D. Eng, L. S. & Ruland, C. (2017). Exploring Working Relationships in Mental Health Care via an E-Recovery Portal: Qualitative Study on the Experiences of Service Users and Health Providers. *JMIR Mental Health 4*(4), p. e54.

Sykes, T. A. Venkatesh, V. & Rai, A. (2011). Explaining physicians' use of EMR systems and performance in the shakedown phase. *Journal of the American Medical Informatics Association 18*(2), pp. 125–130.

Tai-Seale, M. Dillon, E. C. Yang, Y. Nordgren, R. Steinberg, R. L. Nauenberg, T. Lee, T. C. Meehan, A. Li, J. Chan, A. S. & Frosch, D. L. (2019). Physicians' Well-Being Linked To In-Basket Messages Generated By Algorithms In Electronic Health Records. *Health Affairs 38* (7), pp. 1073–1078.

Tran, B. D. Chen, Y. Liu, S. & Zheng, K. (2020). How does medical scribes' work inform development of speech-based clinical documentation technologies? A systematic review. *Journal of the American Medical Informatics Association 27*(5), pp. 808–817.

Vogel, M. Kaisers, W. Wassmuth, R. & Mayatepek, E. (2015). Analysis of Documentation Speed Using Web-Based Medical Speech Recognition Technology: Randomised Controlled Trial. *Journal of Medical Internet Research 17*(11), p. e247.

Vogelsmeier, A. A. Halbesleben, J. R. B. & Scott-Cawiezell, J. R. (2008). Technology Implementation and Workarounds in the Nursing Home. *Journal of the American Medical Informatics Association 15*(1), pp. 114–119.

Vorderstrasse, A. A. Hammer, M. J. & Dungan, J. R. (2014). Nursing Implications of Personalised and Precision Medicine. *Seminars in Oncology Nursing 30*(2), pp. 130–136.

Westbrook, J. I. Li, L. Georgiou, A. Paoloni, R. & Cullen, J. (2013). Impact of an electronic medication management system on hospital doctors' and nurses' work: A controlled pre–post, time and motion study. *Journal of the American Medical Informatics Association 20*(6), pp. 1150–1158.

Zadvinskis, I. M. Garvey Smith, J. & Yen, P.-Y. (2018). Nurses' Experience With Health Information Technology: Longitudinal Qualitative Study. *JMIR Medical Informatics 6*(2), p. e38.

Summary

This book charts the trajectories of organisational and professional change that early-stage artificial intelligence (AI) has set in motion in clinical settings. Even though both lay and academic publics are increasingly interested in, and sometimes worried about, the way AI is going to shape our future, we currently have few empirical analyses of AI as it begins to enter our presents — including in clinical settings. This is, at least in part, because mapping out how clinical AI is generating change is not a straightforward enterprise. In the present moment, AI technologies are often not exactly implemented, or even present, in clinical practice. This (partial) absence can be either due to their early stage of development, to professional, ethical and legal concerns they raise, or simply to the chasm between AI's requirements and hospitals' current levels of digitalisation. Yet, even if it is not present strictly speaking, AI might still manage to accomplish concrete changes.

This book aims to trace contemporary manifestations of clinical AI, and their implications for how care is organised and provided. In doing so, it maps out the concrete ways in which clinical data are produced, the (new) objects that these data and their analysis introduce to clinical practice, and the ways in which professionals incorporate (or refuse to) these new technologies in their daily work.

Theoretically, inspired by artist's Hito Steyerl's suggestive portrayal of engineers breaking glass windows to train machine learning models, this book conceptualises as glass-breaking the shifts through which AI's presence in clinical settings currently manifests. The notion of glass-breaking contributes to current discussions around datafication in clinical settings by emphasising how both technological requirements and discourse around AI achieve concrete changes in organisations. In so doing, it takes seriously the connections that actors in the field articulate between changes they are currently undergoing and the AI futures these changes foreshadow. Moreover, it questions the inevitability of these futures by teasing out the painstaking work that it takes to bring them about.

Furthermore, still borrowing from Steyerl, this book proposes conceptualising the entwinement of AI presents and futures through the notion

of AI's shadow. It proposes that, even in its absence, AI already starts casting a shadow onto clinical settings: as AI is mobilised to address issues that might be overwhelmingly ethical or political, these very issues are turned into technical problems; yet other issues are being left unaddressed. This has implications for which futures are made possible. Tracing this shadow, by focusing on the changes — sometimes mundane, sometimes paradoxical — that take shape around and before clinical AI, is proposed as a methodology through which clinical AI futures can be made discernible in the present. If the futures we are faced with do not always appear desirable, making them an object of conversation and deliberation might help us work towards ones that are more just and sustainable.

Empirically, this book explores AI's current manifestations in clinical settings as diverse as pathology departments, intensive care units, and acute psychiatry clinics. It examines different clinical applications of machine learning (from automated image analysis, to algorithmic management, to behavioural prediction), with differential levels of distance from the clinic (from complete absence to piloting). It builds on three three-month-long ethnographies in a diversity of clinical departments, spanning observations of innovation-related meetings and daily work practices, as well as countless conversations with professionals, developers and managers. In their diversity, these stories and their analyses illuminate different aspects of clinical AI's presents and possible futures.

Chapter one sets the stage. It consists of a literature review that situates current developments around clinical AI in the broader context of the digitalisation of healthcare work. Through a Critical Interpretive Synthesis, this chapter engages with past instances of professional and organisational change related to digital healthcare technologies, picking apart the different ways such changes have been described and conceptualised across Science and Technology Studies, sociology, and medicine. Based on this review, it offers a conceptualisation of the digitalisation of healthcare work as a phenomenon spanning, at once, the open-endedness of situated changes in work practices, and the directionality of technological innovation trajectories. That is, the shifts observed when a specific healthcare organisation introduces new digital technologies are not determined by the technology itself, but a matter of the interplay between specific organisational structures, professionals' creativity and autonomy in (not) engaging with a technology, and of the affordances of the technology itself. However, across specific cases, the changes afforded by digitalisation point us towards specific futures characterised, for instance, by an increased administrative burden for professionals, an emphasis of quantitative data at the expense of more embodied and idiosyncratic modes of knowing, and a polarisation in the job market causing workers lower in the professional hierarchy to be stuck with less meaningful and more isolating work. These points re-emerge, in various ways, in the following empirical chapters.

Chapter two is a story of uncertainty. It takes us to a pathology department in which digital diagnostics are being introduced to pave the way to a future informed by AI-assisted diagnostics. This is a case of digitisation, where the glass slides that pathologists would normally use are being replaced with digital images. However, pathologists rarely perform their diagnoses on these images, which they consider of insufficient quality. This paradoxical situation — investments in technologies that provide limited added value to current care provision — is a prime example of glass-breaking: the shift to digital diagnostics is not meant to benefit pathologists' practices as much as to enable, in the future, AI-assisted diagnostics. This chapter shows how paving the way for an allegedly inevitable future can introduce tensions in the present. In this case, digital slides introduce three types of uncertainty in the diagnostic process: sensorial uncertainty, stemming from their insufficient sharpness, intra-active uncertainty, stemming from their disembodied nature, and fauxtomated uncertainty, stemming from the marginalisation of forms of situated knowing that go into the creation of glass slides. This chapter calls for taking issues of representation and knowledge seriously, and to not fall for narratives of automation that obliterate the distributed nature of knowing in organisations.

Chapter three is a story of embodied and datafied care provision. It explores two ICU departments in which machine learning is being mobilised to intervene on workforce shortages by making nurses' work more efficient. I argue that when it is mobilised to bring about efficiency in highly complex and knowledge-intensive settings, clinical AI, like the one powering the dashboard examined in this chapter, turns professionals into a source of attention. Subsequently, it attempts to allocate this attention efficiently harnessing real-time data. This combination of attention and efficiency is made possible by modernist notions of attention, casting it as a cognitive function that selects a single focus over surrounding noise, and that can be easily turned on and off and redirected. This view of attention, however, clashes with the practices of care provision that ICU nurses engage in daily. In highly datafied settings like ICUs, these practices strive to ensure, simultaneously, the physiological stability of patients' bodies and the production of reliable data about them. This chapter proposes attunement as an alternative concept that can describe and make justice to these practices better than attention. The notion of attunement emphasises the embodied, relational and affectively-charged nature of nurses' practices. As such, it surfaces questions around the desirability of an efficiency-driven approach to attention and, relatedly, workforce shortages.

Chapter four is a story about the ethically-laden decisions that medical professionals make every day. It follows the pilot of an algorithm tasked with predicting inpatient violence in two acute psychiatric clinics. Upon closer examination, the algorithmic predictions are revealed to be actually about pre-empting violent episodes — that is, about enabling professionals to intervene before the threat of violence actually manifests. However, even though risk scores are produced and circulated daily for each patient in the two clinics, local staff never mobilise them as legitimate sources of knowledge in their decision-making. Rather, they consider any prediction that deviates from their own judgement as simply wrong. Understanding this case of non-use requires dwelling on the practices and ethics of dealing with violence as articulated by local nurses. Unlike the algorithm, which takes words in reports as simply 'predictive' of violence, nurses constantly emphasise the importance of probing alternative explanations for the behavioural expressions they witness in acute patients. This chapter argues that the introduction of AI to predict risk and intervene in ethically-laden decision-making might enforce a more punitive logic in acute psychiatry. This logic might suggest, for instance, sedating patients more and earlier in an attempt to achieve a promise of de-risking an inherently volatile clinical environment. In short, this chapter suggests caution in applying machine learning to ethically-laden decisions.

Chapter five draws together the insights emerged across this study, articulating its contributions to different academic and societal discussions. First, it speaks to emerging scholarship on data work not only by restating its importance for current and future AI systems, but also by surfacing the pivotal issue of data quality. If various types of data have long been part and parcel of care provision, AI is starting to change the definition of good data. Data quality is thus increasingly assessed in terms of data's suitability for AI systems, rather than for professionals' care practices. In its relation to (clinical) AI, this book argues that data work thus becomes increasingly a matter of making care practices machine-readable. To contribute to care provision with data, not only does data work need to be supported at the organisational level, but the different requirements of data quality for different actors need to be articulated explicitly.

Second, this book offers the notion of attention redirection to debates around automation and labour. Like automation technologies in other industries, AI is being introduced in clinical practice chiefly to achieve increased efficiencies in the face of workforce shortages. However, it pursues these efficiencies in a decisively different ways from previous technologies. Whereas previous automation aimed at taking over tasks previously performed by workers, with AI we witness a qualitative shift: AI attempts to identify tasks that deserve urgent attention, leaving other tasks, potentially, unperformed. Intervening at the level of professional autonomy, this aspect of clinical AI suggests a measure of scepticism towards attempts at solving workforce shortages through technological means. Moreover, attention redirection warrants careful consideration when AI systems are being implemented in organisations: what value judgements are implicit in this selection, and which tasks are more likely to be left unattended?

Third, this book speaks to discussions around how AI is changing professional knowledge- and decision-making by highlighting the tension between the aura of certainty around algorithmic outputs and the centrality of doubting in care provision. AI systems, once finetuned, operate in a way that attempts to close down users' relation to data: the more they process data, the harder the stories and circumstances that produced those data become to discern. Conversely, an ethnographic examination of care practices reveals how professionals never take data at face value, constantly probing the technologies and the stories that have originated specific data points. This type of doubting is central to their decision-making ethics, and, if neglected, is likely to translate in a lack of engagement with new technologies.

In answering its overarching question, this book proposes that AI manifests in contemporary clinical practices through ontologies of partial absence. Even when they are not yet in clinical practice, AI technologies can hardly be said to be absent: they reconfigure practices, organisational structures, and notions of care. Vice versa, when they are introduced in clinical practice, AI technologies are never fully 'there:' because of their learning nature, which leaves them open for constant optimisation, they are always, necessarily, a prototype — a temporary version of themselves. Ontologies of partial absence make AI technologies a particularly slippery object, especially in legislative terms: they are warranted a special status as experiments, and yet, in their experimental nature, they achieve tangible changes in the world. Given the ultimate insufficiency of legislative approaches to steer AI ethics approaches, and given the concreteness of the changes this technology is managing to achieve in clinical settings, this book argues that professionals themselves, as well as patients and managers, should play a central role in imagining better futures not through, but alongside, clinical AI.

Samenvatting

Dit boek brengt de organisatorische en professionele veranderingstrajecten in kaart die early-stage kunstmatige intelligentie (KI) in klinische settings in werking heeft gezet. Hoewel zowel in publieke als academische discussies steeds meer interesse — en soms ook zorgen — ontstaan over de manier waarop KI onze toekomst gaat vormgeven, zijn er op dit moment maar weinig empirische analyses van KI op het moment dat het zijn intrede begint te doen in specifieke domeinen, zoals in klinische settings. Dit komt, in ieder geval gedeeltelijk, doordat het niet eenvoudig is om in kaart te brengen hoe klinische KI veranderingen teweegbrengt. Op dit moment zijn KI-technologieën vaak nog niet geïmplementeerd, of zelfs nog niet volledig aanwezig, in de klinische praktijk. Dat heeft te maken met de vroege ontwikkelingsfase waarin ze zich bevinden, de professionele, ethische en juridische bezwaren die ze oproepen, of simpelweg vanwege de kloof tussen de eisen die KI stelt en het huidige digitaliseringsniveau van ziekenhuizen. Maar zelfs als KI-technologieën strikt genomen niet aanwezig zijn kan KI toch concrete veranderingen teweegbrengen.

Dit boek tracht de hedendaagse manifestaties van klinische KI en hun implicaties voor de manier waarop zorg wordt georganiseerd en geleverd te achterhalen. Daarbij brengt het de concrete manieren in kaart waarop klinische data worden geproduceerd, de (nieuwe) objecten die deze data en hun analyse introduceren in de klinische praktijk, en de manieren waarop professionals deze nieuwe technologieën onderdeel maken van hun dagelijkse werk (of hoe ze dat weigeren te doen).

Mijn analyse is theoretisch geïnspireerd door Hito Steyerls suggestieve beeld van ingenieurs die glazen ruiten breken om modellen voor machine learning te trainen. Ik conceptualiseer de verschuivingen waarmee de aanwezigheid van KI in klinische settings zich momenteel manifesteert als vormen van glass-breaking. De notie van glass-breaking draagt bij aan de huidige discussies over dataficatie in klinische settings door te benadrukken hoe zowel technologische eisen als het discours rond KI concrete veranderingen in organisaties teweegbrengen. Op die manier helpt de notie van glass-breaking me om de verbanden die actoren in het veld leggen (tussen veranderingen die ze momenteel ondergaan en de KI-toekomsten die deze veranderingen aankondigen) serieus te nemen. Ook stelt het de onvermijdelijkheid van deze toekomsten ter discussie door het moeizame werk te schetsen dat nodig is om ze tot stand te brengen.

Bovendien stelt dit boek, nog steeds in navolging van Steyerl, voor om de verstrengeling van het heden en de toekomst van KI te conceptualiseren door middel van de notie van de schaduw van KI. Het stelt voor dat KI, zelfs in zijn afwezigheid, al een schaduw werpt op klinische settings: terwijl AI wordt ingezet om problemen aan te pakken die misschien overwegend ethisch of politiek zijn, worden juist deze problemen omgezet in technische problemen. Tegelijkertijd blijven andere problemen onaangeroerd. Dit heeft gevolgen voor welke toekomsten mogelijk worden gemaakt. Ik beargumenteer dat het traceren van deze schaduw, door te focussen op de veranderingen - soms alledaags, soms paradoxaal - die vorm krijgen rond en voor klinische KI, een bruikbare methodologie is om de impliciete toekomst van klinische KI in het heden zichtbaar te maken. Als de toekomsten waarmee we geconfronteerd worden niet altijd wenselijk lijken, kan shadow-tracing ons helpen om te werken aan toekomsten die rechtvaardiger en duurzamer zijn door ze onderwerp van gesprek en overleg te maken.

Empirisch onderzoekt dit boek de huidige manifestaties van KI in diverse klinische omgevingen: pathologie, intensive care en acute psychiatrie. Ik onderzoek verschillende klinische toepassingen van machine learning (geautomatiseerde beeldanalyse, algorithmische vormen van management en gedragsvoorspelling), met verschillende niveaus van afstand tot de kliniek (van volledige afwezigheid tot proefprojecten). Het is gebaseerd op drie etnografieën van drie maanden op verschillende klinische afdelingen, met observaties van innovatiegerelateerde vergaderingen en dagelijkse werkpraktijken, en talloze gesprekken met professionals, ontwikkelaars en managers. In hun diversiteit belichten deze verhalen en hun analyses verschillende aspecten van de huidige en mogelijke toekomst van klinische KI.

In hoofdstuk één positioneer ik de huidige ontwikkelingen rond klinische AI in een breder kaderdoor middel van een literatuuronderzoek naar de digitalisering van professioneel werk in de gezondheidszorg. Door middel van een Critical Interpretive Synthesis gaat dit hoofdstuk in op eerdere analyses van professionele en organisatorische veranderingen die verband houden met digitale technologieën in de gezondheidszorg. Daarbij onderzoek ik de verschillende manieren waarop dergelijke veranderingen worden beschreven en geconceptualiseerd in Science and Technology Studies (STS), sociologie en geneeskunde. Op basis van deze synthese ontwikkel ik in dit hoofdstuk een conceptualisering van de digitalisering van werk in de gezondheidszorg als een fenomeen dat zowel de openheid van gesitueerde veranderingen in werkpraktijken als de gerichtheid van technologische innovatietrajecten omvat. Dat wil zeggen dat de verschuivingen die worden waargenomen wanneer een specifieke zorgorganisatie nieuwe digitale technologieën introduceert, niet louter worden bepaald door de technologie zelf, maar dat ze een kwestie zijn van de wisselwerking tussen specifieke organisatiestructuren, de creativiteit en autonomie van professionals in het (niet) omgaan met een technologie, en de affordances van de technologie zelf. Tegelijkertijd zien we dat de veranderingen als gevolg van de digitalisering ons in de richting wijzen van specifieke toekomsten, die bijvoorbeeld gekenmerkt worden door een toegenomen administratieve last voor professionals, een nadruk op kwantitatieve gegevens ten koste van meer belichaamde en idiosyncratische manieren van weten, en een polarisatie op de arbeidsmarkt waardoor werknemers lager in de professionele hiërarchie vastzitten aan minder zinvol en meer geïsoleerd werk. Deze punten komen op verschillende manieren terug in de volgende empirische hoofdstukken.

Hoofdstuk twee is een verhaal van onzekerheid. Het neemt ons mee naar een afdeling pathologie waar digitale diagnostiek wordt geïntroduceerd om de weg vrij te maken voor een toekomst met KI-ondersteunde diagnostiek. Dit is een geval van digitalisering, waarbij de glazen coupes die pathologen normaal gesproken gebruiken, worden vervangen door digitale beelden. Pathologen baseren hun diagnoses echter zelden op deze beelden, omdat ze die van onvoldoende kwaliteit vinden. Deze paradoxale situatie - investeringen in technologieën die een beperkte toegevoegde waarde leveren aan de huidige zorgverlening - is een schoolvoorbeeld van glass-breaking: de verschuiving naar digitale diagnostiek is niet zozeer bedoeld om de huidige praktijk van pathologen ten goede te komen, maar om in de toekomst KI-ondersteunde diagnostiek mogelijk te maken. Dit hoofdstuk laat zien hoe het effenen van de weg voor een zogenaamd onvermijdelijke toekomst spanningen in het heden kan introduceren. In dit geval introduceren digitale coupes drie soorten onzekerheid in het diagnostische proces: zintuiglijke onzekerheid, die voortkomt uit de gebrekkige scherpte van de digitale beelden; intra-actieve onzekerheid, die voortkomt uit hun immateriële aard, en “fauxtomated" onzekerheid, die voortkomt uit de marginalisatie van gesitueerde kennispraktijken rondom het maken van objectglaasjes. Ik concludeer in dit hoofdstuk dat we kwesties van representatie en kennis serieus dienen te nemen en niet te snel moeten vertrouwen op automatiseringsverhalen die de gedistribueerde aard van kennis in organisaties uitvlakken.

Hoofdstuk drie is een verhaal over belichaamde en gedataficeerde zorgverlening. In dit hoofdstuk onderzoek ik twee IC-afdelingen waar KI wordt ingezet om in te spelen op personeelstekorten door het werk van verpleegkundigen efficiënter te maken. Wanneer het wordt ingezet om efficiëntie te bereiken in zeer complexe en kennisintensieve omgevingen, beschouwt klinische KI, zoals de KI die het dashboard aandrijft dat in dit hoofdstuk wordt onderzocht, professionals primair als ‘bronnen van aandacht.’ Vervolgens probeert het deze aandacht efficiënt toe te wijzen door gebruik te maken van real-time gegevens. Deze zienswijze volgt uit modernistische opvattingen over aandacht, waarbij aandacht wordt gezien als een cognitieve functie - het in staat zijn om één focus te kiezen boven omgevingsruis - die gemakkelijk aan- en uitgezet en omgeleid kan worden. Deze zienswijze botst echter met de zorgpraktijken waar IC-verpleegkundigen zich dagelijks mee bezighouden. In een omgeving met veel gegevens, zoals de IC, streven deze praktijken ernaar om tegelijkertijd de fysiologische stabiliteit van de lichamen van patiënten en de productie van betrouwbare gegevens over hen te garanderen. In dit hoofdstuk wordt attunement voorgesteld als een alternatief concept dat deze praktijken beter kan beschrijven en er meer recht aan kan doen dan aandacht. Het begrip benadrukt de belichaamde, relationele en affectief geladen aard van verpleegkundige praktijken. Als zodanig roept het vragen op over de wenselijkheid van een op efficiëntie gerichte benadering van aandacht en, daarmee samenhangend, personeelstekorten.

Hoofdstuk vier is een verhaal over de ethisch beladen beslissingen die medische professionals elke dag nemen. Ik volg de pilot van een algoritme dat geweld in twee acute psychiatrische klinieken probeert te voorspellen. Bij nadere bestudering blijkt dat de algoritmische voorspellingen eigenlijk gaan over het voorkomen van gewelddadige episodes - dat wil zeggen, over het professionals in staat stellen om in te grijpen voordat de dreiging van geweld zich daadwerkelijk manifesteert. Maar ook al worden er dagelijks risicoscores gemaakt en verspreid voor elke patiënt in de twee klinieken, het lokale personeel gebruikt deze nooit als legitieme bronnen van kennis in hun besluitvorming. Om dit geval van niet-gebruik te begrijpen, moeten we stilstaan bij de praktijken en ethiek van het omgaan met geweld zoals die door de plaatselijke verpleegkundigen worden verwoord. In tegenstelling tot het algoritme, dat woorden in rapporten simpelweg 'voorspellend' vindt voor geweld, benadrukken verpleegkundigen voortdurend het belang van het onderzoeken van alternatieve verklaringen voor de gedragsuitingen van acute patiënten. Dit hoofdstuk beargumenteert dat de introductie van KI om risico's te voorspellen en in te grijpen in ethisch beladen besluitvorming, een meer bestraffende logica in de acute psychiatrie zou kunnen afdwingen. Deze logica zou bijvoorbeeld kunnen suggereren om patiënten meer en eerder te sederen om de belofte van het verminderen van risico's in een inherent explosieve klinische omgeving waar te maken.

Hoofdstuk vijf bundelt de inzichten die in dit onderzoek naar voren zijn gekomen en verwoordt de bijdragen aan verschillende academische en maatschappelijke discussies. Ten eerste relateren mijn analyses aan de opkomende studies over data work.Mijn analyses bevestigen het belang van 'datawerk' voor huidige en toekomstige KI-systemen, maar voegen daar ook de cruciale kwestie van datakwaliteit aan toe. Hoewel verschillende soorten data al lange tijd deel uitmaken van de zorgverlening, begint KI de definitie van goede data te veranderen. Datakwaliteit wordt steeds vaker beoordeeld in termen van geschiktheid van data voor KI-systemen, in plaats van voor de zorgpraktijk van professionals. In relatie tot (klinische) KI stelt dit boek dat datawerk zo steeds meer een kwestie wordt van het machine-leesbaar maken van (zorg)praktijken. Om bij te dragen aan zorgverlening met data moet datawerk niet alleen op organisatieniveau ondersteund worden, maar moeten de verschillende eisen van datakwaliteit voor verschillende actoren expliciet gearticuleerd worden. Kortom, dit hoofdstuk suggereert voorzichtigheid bij het toepassen van KI op ethische beslissingen.

Ten tweede biedt dit boek nieuwe inzichten met betrekking tot debatten over automatisering en arbeid in relatie tot KI. Net als automatiseringstechnologieën in andere sectoren wordt KI in de klinische praktijk voornamelijk geïntroduceerd als middel om de efficiëntie te verhogen in het licht van een tekort aan arbeidskrachten. Deze efficiëntie wordt echter op een heel andere manier nagestreefd dan met eerdere technologieën. Waar eerdere automatisering gericht was op het overnemen van taken die voorheen door werknemers werden uitgevoerd, zien we bij KI een kwalitatieve verschuiving die het best getypeerd kan worden als het ombuigen van aandacht: KI probeert taken te identificeren die dringend aandacht verdienen, waardoor andere taken mogelijk niet worden uitgevoerd. Dit aspect van klinische KI, dat ingrijpt op het niveau van professionele autonomie, suggereert een zekere mate van scepsis ten opzichte van pogingen om personeelstekorten op te lossen met technologische middelen. Bovendien rechtvaardigt het ombuigen van aandacht een zorgvuldige afweging wanneer KI-systemen worden geïmplementeerd in organisaties: welke waardeoordelen zijn impliciet in deze selectie en welke taken zullen eerder onbeheerd blijven?

Ten derde sluit dit boek aan bij discussies over hoe KI kennis- en besluitvormingspraktijken van professionals verandert door de spanning te benadrukken tussen de aura van zekerheid rond algoritmische output en de centrale rol van twijfel in de zorgverlening. KI-systemen werken, eenmaal afgestemd, op een manier die de relatie van gebruikers tot gegevens probeert af te sluiten: hoe meer ze gegevens verwerken, hoe moeilijker het wordt om de verhalen en omstandigheden te onderscheiden die deze gegevens hebben voortgebracht. Omgekeerd laat mijn etnografisch onderzoek van zorgpraktijken zien hoe professionals data nooit als vanzelfsprekend aannemen en voortdurend de technologieën en de verhalen die aan de basis liggen van specifieke datapunten bevragen. Dit soort twijfel staat centraal in hun besluitvormingsethiek en zal, als het wordt verwaarloosd, waarschijnlijk leiden tot een gebrek aan betrokkenheid bij nieuwe technologieën.

Als antwoord op de overkoepelende vraag stelt dit boek voor dat de manier waarop AI-toepassingen zich in de hedendaagse klinische praktijk manifesteren het best geduid kan worden als ontologieën van gedeeltelijke afwezigheid. Zelfs als ze nog niet in de klinische praktijk worden gebruikt, kan van KI-technologieën nauwelijks worden gezegd dat ze afwezig zijn: ze leiden immers nu al tot hervormingen van praktijken, organisatiestructuren en opvattingen over zorg. Omgekeerd zijn KI-technologieën, wanneer ze worden geïntroduceerd in de klinische praktijk, ook nooit volledig aanwezig: vanwege hun lerende aard, die altijd ruimte laat voor voortdurende optimalisatie, zijn ze altijd, noodzakelijkerwijs, een prototype - een tijdelijke versie van zichzelf. Ontologieën van gedeeltelijke afwezigheid maken KI-technologieën tot een bijzonder vloeibaar object, vooral in termen van regulering : ze krijgen een speciale status als experimenten en toch bereiken ze door hun experimentele aard tastbare veranderingen in de wereld. Gezien de uiteindelijke ontoereikendheid van wetgevende benaderingen om KI-ethiek te sturen, en gezien de concreetheid van de veranderingen die deze technologie in klinische omgevingen weet te bewerkstelligen, stelt dit boek dat professionals zelf, evenals patiënten en managers, een centrale rol moeten spelen in het bedenken van betere toekomsten, niet door, maar naast klinische KI.

Acknowledgements

Many people have entered, or stayed in, my life while I was working on my PhD. Celebrating them — their work, support, and love — as I celebrate the end of this PhD is not just a custom but a genuine joy. Here is thus a necessary incomplete, voluntarily to-the-point (it is a long book already…), but thoroughly heartfelt, list of people to whom I wish to express my gratitude as I round off this book. It will get personal and bilingual — and it could not be otherwise.

To my supervisors, for their unwavering support and guidance throughout. To Antoinette de Bont, for her clarity, her emotional and intellectual investment, the endless primers in academic politics. To Rik Wehrens, for his intellectual generosity, his paper-trimming skills and, his infinite patience and understanding. To Romke van der Veen, for always holding me to the highest standards.

To my colleagues at the Erasmus School of Health Policy & Management, and particularly in the section Health Care Governance, for providing a stable professional community in which to develop my work. In particular, to Roland Bal, Marcello Aspria, Martijn Felder, and Jaqueline van Oijen, for engaging so generously with my papers in these four years. And to my fellow PhDs — current and past — for sharing, listening, discussing, co-conspiring, giving one another strength and refusing to give up: Amalia Hasnida, Embus Fanda, Hugo Peeters, Iris van de Voort, Jolien Sanders, Jonathan Berg, Kyra Mulders, Leonoor Gräler, Margot Kersing, Renee Michels, Sabrina Huizenga, Syb Kuipers and Sydney Howe, to name but a few, for being a real community, not just a professional one. To Susan Hoefnagel, for always making things happen — and with extreme kindness. To Anne Marie Weggelaar, for the generosity, the no-nonsense attitude, and the care.

To all my research participants, whom I cannot mention by name, for the patience, the openness, and for putting up with my Dutch.

To the other collegial communities that I participated in during these four years. To the always-morphing WTMC community, for keeping alive my enthusiasm for STS during a pandemic, and gifting me with some of my best academic friends (shoutout here to the wonderful IWLHBC members: Carla

Greubel, Lotje Siffels, Mike Grijssels, Rose Bieszczad and Wouter van Rossem – for the discussions, reading recommendations, and the snacks). In particular, to Bernike Pasveer, Anne Beaulieu, Andreas Weber, and Alexandra Supper, for their enthusiasm and for providing us with the most stimulating graduate training a young STS scholar could ask for.

To the Medical Delta programme From Prototype to Payment, for partially funding my PhD and for making me part of a close-knit group of social scientists engaging with medical technologies. In this context, to Diana Delnooj, Erik Schutt, Erik van Raaij, Frank Eijkenaar, Hamraz Mokri, Maureen Rutten — van Molken, Payam Abrishami and Sanne Allers for their interest in and valuable feedback on my work.

To the wonderful community that I had the luck joining during my visit at the Management, Society and Communication department at Copenhagen Business School. Specifically, to Nanna Bonde Thylstrup, for her enthusiasm in making my research visit happen, for her intellectual generosity, and for her ongoing encouragement in research pursuits. To the TechSoc cluster, for the most extensive and generative feedback any of my papers had ever received, and in particular to Mikkel Flyverbom for the sharp and engaged comments. To my new colleagues at TU Dresden, and in particular to Orit Halpern and Celia Brightwell, for their warm welcome, their patience while I was finishing my PhD, and their inspiring collaboration.

To all the research participants, whom I cannot name for reasons of anonymity, for making my work possible, for providing generous insights, and for taking care of me whenever I felt squeamish or slightly shocked on the ward or in the lab.

To my now thoroughly dismembered Rotterdam family. To Simone, María and Hannah, for their love, care and laughter, and for continuing to choose one another even with oceans between us. To Rose, for her mind-blowing generosity, strength and unwavering support. To Lola, for being the best hoop buddy I could have asked for. To my Maastricht wives, Mara and Katrin, for the free therapy sessions, the understanding, and for helping me build the person I am today.

Alla mia famiglia vera, o biologica, diciamo: a mamma e babbo (ovviamente!), per la libertà, l'amore, il supporto incondizionato. E per tutti i pacchi da giù. E a tutti gli altri che, che io ci sia o non ci sia, ci sono sempre.

Come disse il poeta, siete la forza mia.

An in closing, both because it is customary and because *dulcis in fundo*, to Frederik, for knowing how to make me laugh even under the most impossible circumstances, for gifting me with a kind of love that I had never known or known to be possible, and for inspiring me every day, hopefully for many more days to come.

PhD Portfolio

Conference presentations

28/06/2023	Datafied nurses? On real-time data, attention and attunement in the ICU. STS Italia 2023, Panel 7: “Where’s the ‘intelligence’ in AI? Mattering, Placing and Deindividuating AI.” Bologna, Italy.
08/07/2022	Digitising and knowing: A new-materialist reading of change and uncertainty in digital pathology. Politics of Technoscientific Futures — Biennial conference of the European Association for the Study of Science and Technology (EASST), Madrid, Spain.
23/05/2022	Eye for an AI: More-than-seeing and uncertainty in digital pathology. Mid-term symposium of RN24 (European Sociological Association), Helsinki, Finland.
21/04/2022	Medical Delta: From prototype to payment. With Mokri H, Michels R and Allers S. Innovation 4 Health Conference, Amsterdam, the Netherlands.
04/11/2021	More than meets the AI: Automation, expectations, and not knowing how to see in digital pathology. Anthropology of Technology Conference, Aarhus, Denmark.
06/10/2021	Conceptualising the digitalization of healthcare work: A Metaphor-based Critical Interpretive Synthesis. Annual meeting of the Society for Social Studies of Science (4S), Toronto (online).
05/07/2021	More than images: On materiality, seeing and knowing in digital pathology. Lancaster Intellectual Party, Lancaster (online).

17/06/2021	Conceptualising the work-related implications of digital healthcare technologies: A metaphor-based Critical Interpretive Synthesis. Dis/Entangling technoscience: Vulnerability, responsibility and justice. STS Italia (online).

Graduate training

WTMC workshop	The view from somewhere: Geographies of knowledge and STS	2022
Akademie der Künste Autumn school	AI Anarchies: Experiments in study, collective learning and unlearning	2022
WTMC workshop	Trust and truth	2022
RISBO	Essentials of delivery of education	2022
EGSH Dean's masterclass	Feminism and research	2021
WTMC Summer school	Epistemic corruption	2021
WTMC workshop	Datafying nonhumans	2021
WTMC Winter school	A new political sociology of science	2021
RISBO	Group dynamics	2020
NIG workshop	Qualitative data analysis	2020
NIG workshop	Writing ethnographic fieldnotes	2020

RISBO	Basic didactics	2020
WTMC workshop	Care	2020

Teaching

2023-2024	Guest lecturer, Qualitative Innovation Analytics. Utrecht University
	Advanced Research Methods (MSc Health Economics, Policy and Law; MSc Healthcare Management) — Erasmus School of Health Policy & Management
2022-2023	Quality & Safety (MSc Healthcare Management) — Erasmus School of Health Policy & Management
	Technologie en Innovatie (BSc Gezondheidswetenschappen) — Erasmus School of Health Policy & Management
2021-2022	Data-Driven Dreams (elective MSc Healthcare Management) — Erasmus School of Health Policy & Management
	Technologie en Innovatie (BSc Gezondheidswetenschappen) — Erasmus School of Health Policy & Management
	Advanced Research Methods (MSc Health Economics, Policy and Law; MSc Healthcare Management) — Erasmus School of Health Policy & Management
2020-2021	Quality & Safety (MSc Healthcare Management) — Erasmus School of Health Policy & Management
	Technologie en Innovatie (BSc Gezondheidswetenschappen) — Erasmus School of Health Policy & Management

2019-2020 Comparative Health Policy (MSc Health Economics, Policy and Law; MSc Healthcare Management) — Erasmus School of Health Policy & Management

Peer-reviewed publications

Carboni C (under review) Automating (against?) scarcity: On machine learning and the future of clinical work. In Marent B (ed) *De Gruyter Handbook of Digital Health & Society* (invited).

Egher C, Carboni C, Wehrens R (under review). The role of affective labor in expertise: bringing emotions back into expert practices. *Medicine Anthropology Theory*.

Allers S, Carboni C, Eikenaar F, Wehrens R (under review). Understanding the complexities of eHealth innovation scale-up: a cross-disciplinary analysis and a qualitative case study. *JMIR*.

Carboni C, Wehrens R, van der Veen R, and de Bont A (under review). From attention to attunement: Unpacking data-driven efficiency and more-than-human care provision in the work of ICU nurses. *Science, Technology & Human Values*.

Carboni C, Wehrens R, van der Veen R, and de Bont A (2024). Doubt or punish: On algorithmic pre-emption in acute psychiatry. *AI & Society*.

Carboni C, Wehrens R van der Veen R and de Bont A (2023). Eye for an AI: More-than-seeing, fauxtomation, and the enactment of uncertain data in digital pathology. *Social Studies of Science* 53(5): 712-737.

Carboni C, Wehrens R van der Veen R and de Bont A (2022). Conceptualising the digitalization of healthcare work: A metaphor-based Critical Interpretive Synthesis. *Social Science & Medicine* 292: 114572.

Awards

Erasmus Graduate School of the Social Sciences and Humanities PhD Excellence Awards: Best Article 2021.

Other activities

2023	Visiting fellow at Copenhagen Business School, Denmark.
2022-2023	Faculty council representative, Erasmus School of Health Policy & Management.

About the author

Chiara Carboni (1994) studied Medical Anthropology at University College London and Cultures of Arts, Science and Technology at Maastricht University. During her studies, she conducted a research internship at NYU Tandon School of Engineering. Her Master's dissertations focused, respectively, on ontologies of risk in areas affected by contested environmental illnesses, and on logics of care emerging from more-than-human health frameworks such as the One Health model.

Chiara pursued her PhD at the Erasmus School of Health Policy & Management of Erasmus University Rotterdam. During her PhD, she collaborated with several healthcare organisations in the Netherlands, contributing her ethnographic observations to their projects of organisational and technological innovation. Her PhD was co-funded by the Medical Delta programme *From Prototype to Payment*, in the context of which she was involved in extensive interdisciplinary discussions on medical technology. In the last year of her PhD, Chiara was a visiting fellow at the Copenhagen Business School, hosted by Nanna Bonde Thylstrup, with whom she collaborates to this day.

Currently, Chiara is a postdoctoral fellow at the Faculty of Linguistics, Literature and Cultural Studies at the Technische Universität Dresden. She is part of Orit Halpern's Chair of Digital Cultures, and she contributes to the EU-funded project CYMEDSEC. She lives in Copenhagen and Berlin.