Computer Supported Cooperative Work (CSCW) (2024) 33:993–1036
https://doi.org/10.1007/s10606-024-09491-0
RESEARCH

Invisible to Machines: Designing AI that Supports Vision Work in Radiology

Giulia Anichini[1], Chiara Natali[2]* & Federico Cabitza[2,3]
[1]*Cermes3, Centre de recherche médecine, sciences, santé, santé mentale, société, (CNRS UMR 8211, Université Paris Cité, Inserm U988, EHESS), Institut santé numérique en société (ISNS), Paris, France (E-mail: giulia.anichini@inserm.fr);* [2]*Università degli Studi di Milano-Bicocca, Milan, Italy (E-mail:chiara.natali@unimib.it);* [3]*IRCCS Ospedale Galeazzi Sant'Ambrogio Milano, Italy*

Accepted: 24 January 2024

Abstract. In this article we provide an analysis focusing on clinical use of two deep learning-based automatic detection tools in the field of radiology. The value of these technologies conceived to assist the physicians in the reading of imaging data (like X-rays) is generally assessed by the human-machine performance comparison, which does not take into account the complexity of the interpretation process of radiologists in its social, tacit and emotional dimensions. In this radiological vision work, data which informs the physician about the context surrounding a visible anomaly are essential to the definition of its pathological nature. Likewise, experiential data resulting from the contextual tacit knowledge that regulates professional conduct allows for the assessment of an anomaly according to the radiologist's, and patient's, experience. These data, which remain excluded from artificial intelligence processing, question the gap between the norms incorporated by the machine and those leveraged in the daily work of radiologists. The possibility that automated detection may modify the incorporation or the exercise of tacit knowledge raises questions about the impact of AI technologies on medical work. This article aims to highlight how the standards that emerge from the observation practices of radiologists challenge the automation of their vision work, but also under what conditions AI technologies are considered "objective" and trustworthy by professionals.

Keywords: Radiological work, Decision making, Decision support, Artificial intelligence

1 Introduction

This article addresses the practices of observation and interpretation of medical images – in particular chest X-rays and mammography – in the field of radiology, as well as the appropriation of new AI technologies for the task of anomaly detection, hence producing the "visible". The adoption of such medical technologies, based on machine learning algorithms and aimed at assisting clinicians in image reading tasks, is gaining momentum in Europe throughout public and private health establishments. In recent years, observations related to troublesome appropriation of medical AI seem to be overshadowed by a certain rhetoric surrounding the introduction of new diagnostic tools, especially so in fields such as radiology. Here, human limitations such as cognitive biases and fatigue, as well as the variability of practices, are cited as significant factors modulating the effectiveness of expert judgment (Sardanelli and Di Leo 2009). Within this narrative, AI is championed as the catalyst propelling the field towards practices that are more "accurate", "fair", "objective" and standardized (Lincoln et al. 2019): "radiology is now moving from a subjective perceptual skill to a more objective science" (Pesapane et al. 2018).

The techno-scientific promises (Joly 2010) of AI tools point to some desirable modifications in radiological practice such as a reduction in interpretation time and reporting, which are often perceived as mere time-consuming administrative tasks rather than as generative activities in which radiologists exercise and refine their judgment; likewise, AI tools promise a decrease in diagnostic error rate, which is considered still too high (Berlin 2014); an increase in time spent with patients; a solution to the lack of health professionals in rural areas. The role of technology would extend beyond performance improvement: it is entrusted with the resolution of problems of rationalization and redistribution of skills – linked, for example, to the lack of radiologists in certain territories – due in part to the explosion of imaging data (Anichini and Geffroy 2021).

With this article, we would like to contribute to the deconstruction of this narrative by discussing some work practices of radiologists, in order to demonstrate how the automation of the professional gaze is challenged by the adjustment between the standards conveyed by the machine and those structuring the work of radiologists. In a field such as radiology, which is time and time again predicted to be strongly impacted by AI to the point of automation of its basic tasks (detection) (Huisman et al. 2021), we aim to demonstrate that the "construction" of the visible cannot take place without routine "invisible" work (Star and Strauss 1999; Tellioğlu and Wagner 2001; Piras and Zanutto 2010) based on formal, social and "practical" norms.

Essentially, in what way does the work of radiologists inform us on the compatibilities and incompatibilities between professional norms and the technologies aiming to automate the reading of medical images?

Our aim is to uncover, in an even finer way, situations in which the norms mobilized by AI tools to distinguish "normal" from "pathological" in visual data are in contradiction, or converge with, the practices of radiologists. The nature of the data taken into account by radiologists to identify visible anomalies, as well as the writing practices used in reporting, shows that the procedures leveraged by humans and machines to evaluate pathological entities and bodies can diverge. But despite these incompatibilities, and depending on the context of use, these new decision support systems can also meet professional standards and elicit trust, as shown by sociological studies on other AI application areas (Kotras et al. 2021).

In any case we argue that these appropriations imply an arbitration process to establish the accuracy of automatic detection results and that this is a form of work, and even an invisible form of work (Star and Strauss 1999). The emergence of this kind of professional involvement in the adjustment of technologies to professional practice is a crucial topic for the CSCW field because, as Star and Strauss wrote, "(...) the systematic exclusion of certain forms of work mean a displacement of that work and a distortion of the representations of that work." (Star and Strauss 1999, pp. 19-20).

2 AI and vision work

Seeing something is not simply looking at it, as this activity involves a sort of recognition and understanding, as also the common expression 'I see', to mean 'I understand, I get it' suggests. In that respect, seeing is inextricably bound to skills that can be acquired, through study and practice, and more in general through social interactions. This dimension of seeing, which we could denote as sociological to recognize its complementarity to the cognitive dimension of perceptual interpretation, has been reported and discussed in several studies (Amann and Cetina 1988; Coopmans et al. 2014; Lynch 1985, 1988; Grasseni 2004). These analyses show how the "visible", rather than being an objective and immediately available piece of the reality, actually emerges in collaborative practices and from situated interactions among heterogeneous actors, and it is intertwined with the dimensions of the "pertinent" (to some aim) and the "relevant" (for somebody).

Anthropological accounts highlight how each discipline builds a particular relationship to images (Alac 2011; Henderson 1998; Prasad 2005; Traweek 1997), and how this relationship is embedded in a specific material culture, which includes artifacts and routines of use, which give visual evidence a structure (Amann and Cetina 1988) and prepare "the way for perception by pre-coding, geometrizing and normalizing the properties of what comes to be perceived." (Lynch 1985, p. 59). For instance, Aanestad et al. (2003) make the point that even a well-structured object such as a laparoscopic image is actually a collective result, since its quality is created and sustained by a socio-technical network consisting of both technical and non-technical actors (equipment, materials, human actions and skills - technical surgical skills, communication skills, local division of labor) performing this 'invisible work' (Star and Strauss 1999; Tellioğlu and Wagner 2001; Piras and Zanutto 2010).

In this sense, images never speak "for themselves". Instead, their significance is derived from the intricate activities they are embedded within, forming an interpretation that is deeply rooted in the socio-material practices and knowledge production processes in which these images are situated (Perrotta 2013). As clinicians navigate between patients, locations, technologies, and information sources, they constantly reconfigure the optimal alignment of people, resources, and knowledge (Bardram and Bossen 2005). Thus, medical data is not a mere static entity, but an epistemic object (Parmiggiani et al. 2022) which is inherently undefined and acquires meaning through practices, intertwined with its production context and the lived reality of work (Berg and Goorman 1999; Hartswood et al. 2023b).

Our study is in line with the aims of the ethnomethodological approach, which seeks to highlight, through a detailed analysis of practices, the way in which the emergence of the visibile is indebted to a series of situated material, gestural and literary operations. The interpretation of the visible depends on the mobilization of experiential and collective knowledge, on specific concrete actions, on the connection of heterogeneous data, alignment or domestication (Callon 1986) of technological artifacts. On top of this more "situated character" of vision is added a more community framing which responds to rules and conventions which are repeated and stabilized within professional boundaries.

This framework was defined by Goodwin by the concept of professional vision, a set of "socially organised ways of seeing and understanding events that are answerable to the distinctive interests of a particular social group". (Goodwin 1994, p. 606).

Many similar studies have been conducted in the last thirty years (Cohn 2007; Pentimalli 2020; Roepstorff 2007; Saunders 2008), which have all been aimed at unravelling the socio-material nature of practices of medical seeing, mostly inspired by the work by Goodwin (1994).

Radiological work, in particular, has been object of several studies that have aimed to show how medical gaze is embodied in collective expertise (Byrne and Stengel 2010; Hartswood et al. 1998; Rouncefield et al. 2003; Slack et al. 2016, 2010). Some of these studies have focused on the use of Computer-Aided Detection systems (CADs), designed for the automatic detection of lesions and to reduce omission errors. These tools report suspicious images to radiologists to make them more sensitive and less prone to false negative errors (which may have serious consequences in terms of missed or delayed detection of progressing conditions). Far from simply "helping" radiologists, these kind of software systems require them to exert efforts to become familiar with their strengths and weaknesses. Knowledge of the machine becomes therefore part and parcel of the radiologists' "professional vision" and further structures their skills (Slack et al. 2010).

To the notion of professional vision, which is supposed to constrain professionals through mental schemas and practices which lead them to embrace a way of seeing, we prefer the more general and less restrictive notion of "vision work". Conceiving vision as a form of work allows us, as Goodwin suggests, to define it as a collective activity, characterized by a set of skills, conventions, habits and know-how (Grasseni 2004). Secondly, it allows us to consider vision not as an act through which reality is revealed in a direct and universal manner, but as a situated and participatory activity (Latour 1986). On the other hand, the concept of vision work as opposed to professional vision leads us to think about non-stabilized practices that cannot yet be defined as part of the repertoire of professional skills. More precisely, when introducing technologies that intervene in the interpretation of the visible, normative work (Dodier and Barbot 2016) must be performed by professional for these machines to be effectively introduced into their technical, epistemic and social space. Through the vision work, which involves the inclusion and exclusion of technologies in the production of knowledge, we can then understand the emergence of norms that support the perceived "objectiveness" of the tools employed in reading the visible.

In addition, compared to professional vision, vision work enables us to conceive the act of seeing not only as a collectively structured practice, but also as an operation rooted in formal and informal knowledge mobilized at the individual level, hence not necessarily as the prerogative of more established professional conventions. Finally, vision work appears to us as a more general concept, since the practices mobilized in deciphering the visible may reflect the local realities and the communities to which they belong (a specific hospital or department, for example) and not necessarily the whole professional field. This concept therefore encompasses both individual and collective operations, and refers to both professional norms and community practices that influence the formation of a certain "way" of seeing within a local community of practice (Wenger 2009).

Vision work is used here to underline the active commitment of actors in the construction of the visible, a commitment which takes place not only through the exercise of shared conventions linked to professional cultures, but also through daily, sometimes invisible work that individuals carry out to cooperate and to respond to social imperatives. This concept therefore makes it possible to connect the action of seeing, conceived in its modalities (more particularly targeted by Goodwin's expression) with the activity of making something visible for certain purposes, such as convincing or reassuring a patient, stating and legitimizing a scientific fact, asserting one's professional identity in the eyes of the members of the community of practice. Focusing on this vision work, in both its modalities and its objectives, helps us to understand when and why the technologies involved in either supporting or automating visual detection fit in with, or conflict with, the norms of their users. Thus, it is also a theoretical concept that we propose for analytical purposes, requirement elicitation as well as technology design practices.

Recent years have seen a rise in studies focusing on radiologists' perception and expectations of AI tools (Cai et al. 2019; Pinto dos Santos et al. 2019), and on their impact on medical reasoning (Cabitza et al. 2017), but there is still a lack of empirical studies on effective integration of AI systems in radiological work, which makes it difficult to formulate recommendations truly rooted in the contexts of use and professional practices. Our goal is therefore to fill this gap by bringing out through empirical investigation the social and "practical norms" (Olivier de Sardan 2010) which characterise radiologists' work and to use them to propose some implications for the design of AI-based decision support systems in radiological tasks.

In the first part of this article, we will present our case study and how AI technologies and radiologists are defining the framework for apprehending the "visible". We will point out some potential obstacles to the articulation of automatic image labeling and clinical practice, discussing the difficulty of taking into account the context surrounding visual data, non-quantifiable information, or non-pathological anomalies. We will also introduce the tacit and informal knowledge involved in vision work, its significance in radiological expertise but also the uncertainties it raises, especially so in terms of the the successful adoption of AI tools. In the final part, the insights drawn from our case study will inform a series of implications for the design of AI tools that safeguard as well as enhance professional skill and responsibility.

3 Foundations in CSCW

The title of this paper is inspired by the works of Star and Strauss (1999) on the topic of *invisible* work and those by Lucy Suchman, particularly her book *Human-machine reconfigurations: Plans and situated actions* (Suchman 2007). In this latter work, Suchman critically analyzes the differences between human understanding and machine processing, noting that machines have access only to a very narrow subset of observable and codified actions and lack the rich contextual and inferential frameworks that humans use to make sense of the world. Vision work in radiology involves not just the technical reading of images but also a deep, nuanced understanding of patient histories, clinical data, and subtle cues that remain, in Suchman's words, "Not Available to the Machine", or *invisible*. In other words, we make the point that there remains an intrinsic gap between the depth of human perception and the surface-level interpretations of machines. This concept resonates deeply with the vision work in radiology, which goes beyond technical image interpretation to include a nuanced understanding of patient histories, clinical data, and subtle cues — as well as the rich, interactive process essential for accurate diagnosis — aspects that remain largely "invisible" to machine. Our observations contribute to the reflections of other scholars on the difficulties of algorithmic devices in capturing and defining the context that give meaning to the actions of a community (Dourish 2004; Seaver 2015), thus reinforcing the plea by Blois (1980): "The most important question appears not to be 'Where can we use computers?', but 'Where must we use human beings?"'. We will explore further the issue of human irreplaceability via the introduction of the concept of "Frictional AI" in Section 6.

Adding to this foundational understanding, we turn to Marc Berg's insights in "Accumulating and Coordinating: Occasions for Information Technologies in Medical Work" (Berg 1999), especially his understanding on "how distributed and interrelated entities can create new forms of activity, irreducible to and spanning over the actions of isolated elements". Berg (1999) argues for a relational understanding of technology's generative power, seeing it as embedded in and dependent upon work practices. This view aligns with our focus on radiology, where AI must be understood in relation to its embeddedness in complex, nuanced work practices. The paper is also indebted to the contribution by Luff et al. (1992) that introduced the concept of "Tasks-in-interaction", which emphasizes the need to understand tasks not just as individual cognitive activities but as situated within broader collaborative and communicative practices. Luff et al. (1992) argue that almost every task, regardless of the setting, is dependent upon an indigenous body of work and communicative practices. Furthermore, Berg (1999) emphasizes the transformative power of technological artifacts and the work activities mediated through them. For radiology, this means recognizing how AI tools and radiologists can co-create new forms of activity that are irreducible to the actions of isolated elements.

This scholarly background grounds our discussion on the role of AI in radiology by allowing us to acknowledge the limitations of AI in capturing the 'invisible' aspects of radiological vision work and to situate our research within the broader discourse of CSCW. By doing so, we aim to contribute to the CSCW literature by introducing the concept of "vision work", investigating reasons (Section 5) and design-oriented methods (Section 6) for successful appropriation of AI systems by radiologists.

4 Methods

We approached our objects of study – the AI systems conceived for the automated detection of visual anomalies – from the point of view of their reception and appropriation by their users. We draw out their specificity from the relationships they maintain with existing socio-technical norms, and the possible impact they could exert on such norms. Angèle Christin (2020) termed "algorithm refraction" the ability of computational devices not only to reflect the social and epistemic order of one (or more) social group(s) through their design and use, but also to reconfigure this very order. According to Christin, "applied to algorithms, studying refraction entails paying close attention to the changes that take place whenever algorithmic systems unfold in existing social contexts-when they are built, when they diffuse, and when they are used" (Christin 2020, p. 906).

We conducted a case study to first and foremost highlight how the vision work of radiologists may collide with, or rely on, AI tools. Our research questions were: what are the reasons why radiologists may reject automatic detection, and under what conditions does successful integration take place? What are the divergences/convergences between standards and knowledge that explain these different appropriations? To do so, we carried out an ethnographic survey over an 18-month period (between 2019 and 2020) and conducted interviews and observations of radiologists covering different hierarchical statuses (junior doctor, senior radiologists) (Table 1).

In conducting our ethnomethodological research, we adhered to a systematic approach to gather and analyze our data. Initially, we recorded and subsequently transcribed interviews with radiologists and engineers, carefully dissecting their discourse. This analysis was aimed at uncovering themes related to the socio-technical promises of AI, professional expectations, and the roles ascribed to AI tools within their work practices. Additionally, we meticulously recorded conversations observed during sessions where clinicians interpreted images (either without decision-support tools or when confronted with automatic reports), com-

Table 1. Overview of participants to our 18-month ethnographic survey (2019-2020).

Participant	Occupation	Type of decision support systems	Data collection	Institution
1	Radiologist (head of imaging department)	/	Interviews and work observation	Public Hospital 1
2	Radiologist	/	Interviews and work observation	Public Hospital 1
3	Radiologist	/	Interviews and work observation	Public Hospital 1
4	Radiologist	/	Interview	Public Hospital 1
5	Radiologist	/	Interview	Public Hospital 1
6	Radiologist	/	Interview	Public Hospital 1
7	Radiologist (head of imaging department)	/	Interview	Public Hospital 2
8	Radiologist	CAD	Interview	Public Hospital 2
9	Radiologist	/	Interview	Public Hospital 2
10	Resident doctor	AIT	Interviews and work observation	Public Hospital 2
11	Resident doctor	AIT	Interviews and work observation	Public Hospital 2
12	Resident doctor	AIT	Interviews and work observation	Public Hospital 2
13	Radiologist	AIM	Interview	Private Hospital
14	Radiologist (head of imaging department)	AIM	Interview	Cancer screening center 1
15	Radiologist	AIT	Interview	Cancer screening center 1
16	Radiologist	AIT	Interview	Cancer screening center 2
17	Radiologist	AIT	Interview	Cancer screening center 3
18	Start-up's engeneer		Interview and work observation	Start-up distributing AI solutions for medical imaging
19	Start-up's head of sale		Interview and work observation	Start-up distributing AI solutions for medical imaging

plementing these recordings with detailed field notes, thus identifying the data, knowledge and objects that radiologists leverage when assessing the normal or pathological nature of lesions. This analytical step was crucial in shedding light on the intricate, often tacit, reasoning that underpins vision work in radiology, offering profound insights into how AI tools are integrated and perceived in actual medical practice.

Our approach was deeply influenced by Olivier de Sardan (1995) and his concept of "politique du terrain" (policy of fieldwork), as we engaged in spontaneous dialogues with our subjects and iteratively refined our research questions based on the emergent data, ensuring a dynamic and responsive research process.

To understand the work of radiologists and their activities in the vision work of medical images, we first focused on professionals not using automatic detection tools (Public Hospital 1). This first part of the survey aimed at familiarizing with the work of radiologists, in particular during the sessions dedicated to the interpretation of medical images, in order to grasp the individual and collective strategies mobilized in the recognition of lesions and the categorization of visual data. The semi-directive interviews (N=6) focused on the vision work underlying the recognition of pathological anomalies, their previous and current experiences with computer technologies involved in the automation of medical work, their possible knowledge of AI tools for anomaly detection, their perception of AI in radiology.

We also observed, took ethnographic notes of, and recorded the vision work of three radiologists (about 10 hours of observation, distributed in 5 sessions). The conversations that took place during the observations, in which radiologists were encouraged to explicitly provide their own assessment of the images, was analyzed with two aims. On the one hand, we aimed at identifying the knowledge and strategies involved in the act of seeing; on the other hand, we aimed at taking account of the way in which this judgment was formulated in the medical report. The aim of the analysis is therefore the identification of how the "visible" is first understood by the radiologist and then communicated in written reports to other clinicians. This helped us in understanding the standards that were followed in the delimitation of pathological lesions during image reading and the written transmission of this reading.

In a second step and in a second hospital (Public Hospital 2), we followed the deployment of an AI tool (referred to as AIT) distributed by a French start-up and aimed at assisting radiologists in the interpretation of chest X-rays.

AIT is a deep learning system that was trained on an extensive collection of 2.5 million chest X-rays, sourced from 45 distinct global centers. This platform, based on convolutional neural networks (CNNs), can detect the following chest anomalies: 'blunted CP angle', 'cardiomegaly', 'cavity', 'consolidation', 'fibrosis', 'hilar enlargement', 'nodule', 'opacity', 'pleural effusion'. It was validated in a study by Putha et al. (2018) using two different datasets: the first encompassed 2000 chest X-rays, with the ground truth being the consensus of three radiologist; the second comprised of 100,000 X-rays, benchmarked by a radiologist's report.

This second hospital was chosen because it represented a first experience of clinical use of AIT at the national level, and it was therefore the privileged observation point for the first phases of implementation (technical installation, training sessions for radiologists, first tool uses) of AI solutions for decision support in radiology. We conducted interviews with radiologists (N=6) and start-up employees (N=2) and observed the first uses of the new AI software by three resident doctor (about 20 hours). The interviews focused on the hospital's commitment to AI projects, radiologists' expectations and fears and impressions about the technology's use. Analysis of the data collected was aimed at understanding the promises of AI technologies which emerge from the discourse, as well as the constraints that emerge from users, the nature of the uncertainties surrounding algorithmic tools, the convergence and divergence between automatic and "human" labeling of images.

We also followed, over several months, the negotiations between the clinicians and the engineers preceding the deployment of the technology, the training sessions and the user monitoring provided by engineers. At the same time, we were put in contact with users of another tool distributed by the same company, for comparative purpose and to foster collaboration with the start-up behind the development of the first tool. We therefore conducted interviews (N=5) with users of another detection tool for breast cancer detection in mammography (we will refer to it as AIM).

This system, compatible with various vendors, underwent rigorous training on over 9000 mammograms with cancer from four major vendors. It harnesses the power of deep learning CNN to accurately detect calcifications and soft-tissue lesions in the standard digital mammographic views: the craniocaudal and mediolateral oblique perspective. Each detected region is assigned a suspicion score, which is then synthesized into an examination-based proprietary score, providing an indication of the likelihood of cancer presence. (Rodríguez-Ruiz et al. 2019)

In the case of AIM, our interviewees were radiologists belonging to various facilities (cancer screening centers and private hospitals) that had integrated AIM in clinical practice about a year before, and interviews were conducted exclusively via videoconference due to the restrictions put in place during the covid-19 pandemic. Through a multi-sited study, we were able to initiate comparisons and bring out more clearly the diversity of issues that are specific to each context of use. Starting from a detailed description of medical practices, we investigate the possibility that the case study may incorporate a normative dimension (Latzko-Toth 2009) and endow our empirical approach with a more design-oriented purpose that we explore in the last section.

5 Findings from the fieldwork

In this section, we examine the vision work of radiologists, in particular through the introduction of AI technologies and the challenges posed by these systems.

We will focus on the normative framework within which the radiologists' gaze unfolds, exploring four axes: the processing of data and information that help in the detection and distinction of visible lesions, which often remain unquantifiable by technology (Section 5); the role of medical reports and writing practices in the radiologists' vision work, and the way in which these practices (which involve, for example, omitting or making more visible specific information in order to suggest therapeutic action, facilitate the interpretation of the report by the recipient, protect against legal action, etc.), which respond to social constraints, participate in the production of the visible (Section 5); the differentiation between pathological and benign anomalies, which remains problematic for AI tools due to an algorithmically-defined "super-normality" that may pathologize otherwise benign abnormalities (Section 5); and the connected over-structuring of vision work as automated by AI tools at the expense of the tacit and experiential knowledge involved in the anomaly recognition process (Section 5).

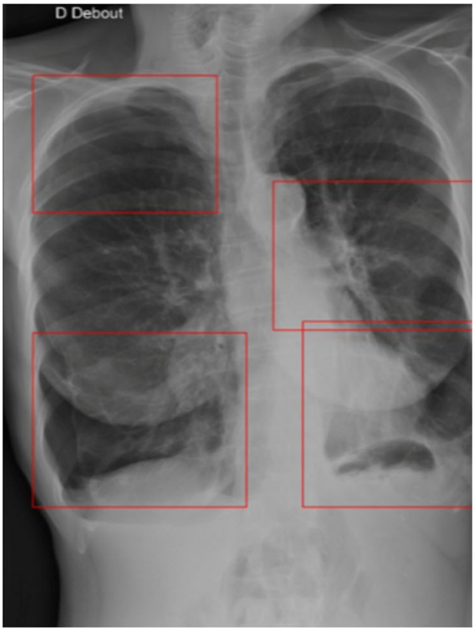
Informed by these four axes, we will illustrate an example of a more successful integration of technologies, which was possible thanks to the machine's ability to respond to the existing knowledge production regime. In particular, in cases where tacit knowledge (which is difficult to formalize) is mobilized to manage "small" uncertainties, the machine can reproduce a collegial decision-making process that is familiar to clinicians, hence providing a support that is perceived as more "objective".

5.1 The undatafiable dimension

AIT, a deep learning-based tool, provides a classification of images both in terms of "normal" or "bnormal" and according to one of 12 categories (e.g., pleural effusion, pneumothorax, cardiomegaly, nodule). This tool was first conceived by an Indian team mainly for the detection of tuberculosis, which is endemic to the country. AIT is initially used by resident doctors confronted with reading emergency X-rays (see Figure 1). For many of the the junior radiologists we met, AIT provides an unsatisfactory labelling of anomalies. The machine can, for example, detect tuberculosis following the identification of certain cavities present on a chest X-ray. During an observation session, one of the young radiologist, confronted with this type of automatic labelling (see Figure 1), explained that these anomalies can in fact be correlated with tuberculosis, but that it mainly depends on the information accompanying the imaging data:

> "(...)Tuberculosis is not the only etiology that gives excavated lesions like that, according to the clinic and the samples they took, you have to correlate with all that, but (...) in fact if I have an X-ray like that and the emergency doctor tells me: he comes from the emergency room, he comes back from... I don't know... he lives in Morocco or in an endemic area where he may have contracted tuberculosis, he is coughing, there is reasonable doubt that we may be in front of tuberculosis." (Resident doctor N.10, University Hospital 2)

In fact, in this case, the patient's history and the emergency physician's notes relating to hospitalization for pneumonia led the radiologist to exclude tuberculosis. Another chest X-ray showed enormous lung volume, but the machine classified the image as "normal". Since this symptom can be associated with emphysema, the junior radiologist checked other contextual information. This information, inacces-



Observations sur radiographies de thorax et estimation

IMPRESSION

Examen anormal

OBSERVATIONS	PRÉSENCE	LOCALISATION
Anormal	☑	
Parenchyme pulmonaire		
Opacité	☑	
Cavité	☑	MG, IG, ID
Consolidation	☑	MG, IG, ID
Nodules		
Fibrose		
Autres opacités		
Médiastin		
Cardiomégalie		
Proéminence des régions hilaires		
Plèvre		
Épanchement pleural		
Emoussement de l'angle costophrénique		
Pneumothorax	☑	

Figure 1. Example of AIT's pneumothorax detection and automatic report: In the top X-ray, the rectangles outlined in red indicate the area where anomalies have been detected. Under the chest X-ray, the automatic report is displayed. The red marks indicate the general image classification (here the image is detected as abnormal) and the type of anomalies identified.

sible to the machine, will make it possible to decide on the pathological nature of the anomaly:

> "Is he being seen for pulmonary symptoms - and in that case we would have to investigate further, or is he being seen for something entirely different and he has no pulmonary problem at all and maybe he's just someone who may be very tall, you see (...) sometimes it's normal because the person is physiologically tall and skinny, sometimes it's abnormal because the person has a real thoracic distension linked to a pulmonary pathology, but then it is the clinical context that comes into play to explain that..." (Resident doctor N.10, University Hospital 2)

Having access to other contextual information that allowed him to grasp the specifics of the case, including the patient's physical characteristics and age, the radiologist also classified the abnormality as non-pathological. While the software's advice was consistent with the radiologist's interpretation, the elements included in the analysis resulting in the automatic classification were not the same. In the two cases described above, the definition of the pathological character of the anomaly referred to by the radiologist is not based on its appearance, at least not only, but emerges from the relation that the visible anomaly maintains with clinical information and experiential knowledge. Some of the categories of pathology used by the radiologist also seem to differ from those covered by the software. In particular, the radiologist can rely on visible clues that are not really pathological and do not correspond to anomalies listed by the algorithmic tool, thus escaping detection. During an ethnographic observation, we observed a young radiologist looking at an image where he saw small white dots. For AIT, the image is normal, but the radiologist, although there were no clear anomalies, considered it could be a chronic phenomenon (which is also reported by emergency physicians in their notes), saying:

> "I don't know, it's not a pathology, it's not normal, but it's a phenomenon that is a little more chronic." (Resident doctor N.10, University Hospital 2)

So the anomaly may not be traced back to a formal category and reflect a more general suspicion of a chronic pathology. Similarly, while a radiologist saw a bronchial syndrome in another chest X-ray, AIT classified the image as normal. The young radiologist, then, had the impression that while the software classifies certain anomalies correctly, it is blind to "large entities", as he explained, such as interstitial syndrome or bronchial syndrome which nevertheless guide the interpretation of many images. In the cases described here, the definition of the pathological character of the anomaly referred to by the radiologist is not based on its visible aspect, at least not only, but emerges from the relation that the visible anomaly maintains with several clinical information, knowledge and experiential knowledge. Moreover, regarding the recognition of visible anomalies, they refer to categories that go beyond those that guide the detection of the software, which allows them to embrace several forms of pathology. A particular challenge to these detection tools is the evaluation of some information which, as we will see, can also relate to the patient's own experience.

AIM is a tool that is based on a deep learning algorithm like AIT. It is conceived for the detection of masses and microcalcifications in mammograms (see Figure 2). The software produces an overall score of the abnormality of the examination (the maximum score being 10) and a score for each abnormal zone detected which indicates the probability (from 0 to 100) that it is a lesion. As for the interpretation of chest distension, in screening for breast cancer, age and as well as other information are important elements through which radiologists assess images. A radiologist explained how the patient's and his family's history are essential information and should be taken into account by AIM but, for the moment, they remain excluded from the machine's processing:

> "Finally, what should be integrated into AI is the patient's history, (...) we will also take into account her family history, if she has a history of breast cancer in the family, we will move faster towards the biopsy for example. If we have any doubts about a lesion, we will be a little more rigorous. We will also take into account her age, of course: on a young patient, we can tolerate certain anomalies; if they appear at a fairly old age, this is obviously something bad."
> (Radiologist N.15, Cancer Screening center 1)

The patient's age and medical history therefore guide the interpretation of an anomaly. The same lesion can be assessed differently, depending on the context of its emergence. Judging the seriousness of a lesion then entails going beyond its

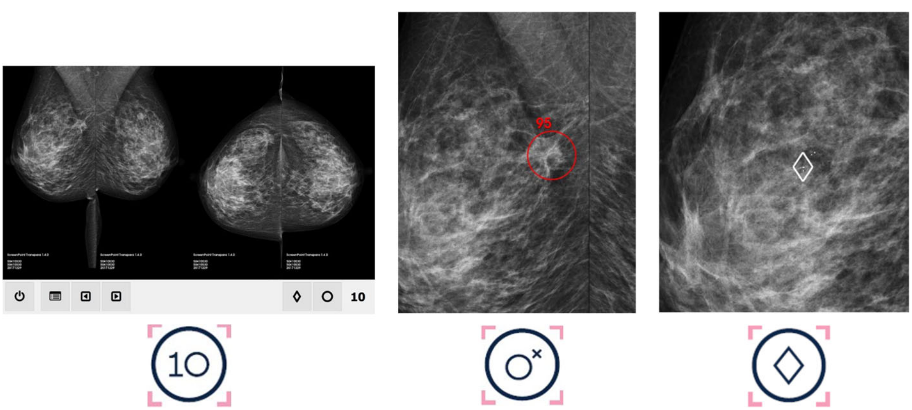


Figure 2. Examples of AIM's functionalities. From left to right, in the first visualization termed "exam score" the global score indicates the normality of the image on a scale of 0 to 10 (the higher the score, the greater the probability that it contains malignant lesions); In the second visualization, "region analysis", a regional score from 1 to 100 (in red) expresses the malignancy of a lesion; In the third visualization, "perception aid", the most suspicious microcalcifications are surrounded by a diamond shape.

visible appearance, because to determine its pathological character one must resort to a broader clinical picture. This information does not only concern the individual history: to assess both how at-risk the patient is and the probability that an anomaly is malignant, the senologist enquires about the patient's family. This helps her to guide the images interpretation, and in general to assess the risk surrounding a lesion.

To understand a lesion and give it meaning, the radiologist also refers to previous imaging exams, which work as benchmarks for the possible evolution of the anomaly over time. Comparing the state of a lesion at different times allows her to see a possible progression or deceleration of the disease, making it possible to better detect cancers that are particularly difficult to identify. One senologist gave us details about this process:

> "I always say to the women that they are cancer camouflage, normally they are... They look like the image. They really look like glands and how do you catch them? This is precisely because we compared them to previous images. You can see that there is an area of the breast where there is more gland, you are not supposed to have more gland in one place or another. Above all, it is the comparison that will help us."
> (Radiologist N.15, Cancer Screening center 1)

Viewing older images is sometimes essential for identifying a lesion. What is visible is also considered through the prism of clinical information gathered during the medical examination. The vision work of the mammography is thus combined with touch-based examination of the patient's affected area, which allows other signs of the disease to be grasped. In an interview, a breast specialist thus described this essential aspect of her work:

> "Sometimes there are even things that are infectious, inflammatory, there may be masses that correspond to abscesses so, we will ask the patient, we will look at her breast: is it red? Is it hot? Do you have a fever? But also the clinical examination…that is also very important."
> (Radiologist N.16, Cancer Screening center 2)

An exploration by palpation as well as asking the patient about her feelings are benchmarks which can guide the interpretation of images and which form part of the framework in which the interpretation of visible anomalies takes place. This is especially clear in this other excerpt:

> "I always start with the clinical examination, I always examine the patient, if I see an opacity on the mammogram, I'll look for it, I'll palpate it in that area, and I'll see that there's a mass, is it soft, is it hard? This will also help me in my final diagnosis, the consistency…(…) Sometimes the patients arrive and say: I can feel something. And so we'll necessarily go to …It also depends on the context…"
> (Radiologist N.16, Cancer Screening center 2)

The anomaly is therefore not (always) significant on its own: the radiologist needs to know its consistency by touch and the context of its appearance. If there is a discrepancy between what emerges from the clinical examination and the mammogram, the radiologist is often required to order an additional imaging test - like a magnetic resonance imaging (MRI) test.

In any case, the reference to the context remains a major pitfall in the use of AIM, since its detection is "blind" to a set of data that makes the radiologist lean towards a particular explanation of what is visible. The data considered in the vision work can relate to the individual and provide information about their medical history and feelings. Data can also regard the family sphere and go beyond the boundaries of the individual body. The assessment of cancer heritability, that includes family members, resulted in the use of the term "extended patient" (Bourret 2005). Pascale Bourret introduced this concept in the context of clinical work on gene mutations, to signify the role of data that goes beyond the patient's body and phenotype in the production of an individual prediction or diagnosis. The definition of the visible is therefore intertwined to an inferential process that uses information that comes from clinical examination, from individual history, but also from the family context.

5.2 Purposeful omission for useful representation

During the vision work, radiologists compare what they see with the clinical information available via the PACS (Picture Archiving and Communication System). As described in the previous subsection, knowledge of the patient's history is often essential for radiologists to "see" anomalies in medical images. For some pathologies, long minutes must be spent analyzing the data accompanying the images in order to start the interpretation of anomalies. This allows the radiologists to understand the clinical case in its "entirety" and to know how to orient the vision work according to the specificities of the patient. For example, they can broaden their gaze towards the possible consequences of a disease in areas other than those first targeted by the imaging investigation (such as bone metastases in the case of lung cancer) or focus on potential anomalies in "at risk" patients.

Despite this extended attention, the radiologist is also led to neglect certain anomalies when they consider them not worrisome, or to voluntarily restrict their research according to the pathology. Regarding this second point, when observing an exam prescribed for a case of bleeding in the intestine, a radiologist told us that, while in a tumoral context he would have also looked closely at the bones to search for secondary lesions, in this case he did not focus on them. But even when the radiologist sees certain anomalies, they can decide not to communicate them: when writing the medical report, they select the information to deliver but also what to omit to the clinicians because not relevant for the subsequent interventions.

During vision work the radiologist may, for example, see calcifications, cysts or micronodules, but will not necessarily mention them. When viewing a scan of a patient's lungs, a radiologist told us:

> "You don't even have to describe them, (e.g., a calcification), you don't have to describe them (…) no need to because it's irrelevant, it has no influence, it won't change anything, if the patient is 75 years old, it's the after-effects of something he did as a child or something else, we're not going to add that to the medical record."
> (Radiologist N.1, Public Hospital 1)

Other anomalies, considered otherwise harmless, can be described if they are associated with specific diseases. Another radiologist told us:

> "If we are in the initial assessment of a cancer and the patient's liver, for example, shows a metastasis and other lesions that are cysts, we will describe them (…) We will describe them even if they are of no interest because they are present with something that is (…) malignant. On the other hand (…) cysts on the kidneys, we won't necessarily mention them because there is no connection, they are completely irrelevant".
> (Radiologist N.1, Public Hospital 1)

Here, it is not the nature of the lesion itself that defines its dangerousness, but rather its location and the context of its emergence.

It is not enough to "see" an anomaly, as alluded in the maxim "seeing is for knowing" reported by Carlin et al. (2010). This implies that if viewing the images does not result in increased knowledge, or in this case, inform treatment, then there is no benefit to it. The pathological character of the anomaly must be recognized and defined starting from the medical information available: patient history, the questions asked by the clinician to the radiologists, but also the message that the radiologist wants to give to the other specialist.

On this last point, radiologists must make a selection, at the time of redacting the medical report, according to what they consider important to convey, with the aim of guiding the physicians towards a better understanding of the case, the production of a diagnosis, even towards the initiation of specific therapeutic action. This sometimes requires the radiologist to anticipate the possible effects their allocutions will have on the intended recipients and the reactions of the referring physicians. This is the case of a radiologist who voluntarily omitted the presence of micronodules in the image: as he did not consider them alarming, he found that description would be likely to cause unjustified concern for the physician who would receive the medical report, potentially leading them to initiate inappropriate medical actions (e.g. overdiagnosis). Furthermore, the medical report is an instrument which structures the relation between the radiologists and their colleagues, and it is often seen as a means to build their reputation.

> "If you are credible with your interlocutor …it (the medical report) shouldn't be a stereotypical thing: it is not this, it is not that …copy - paste always

> the same uh …voilá, they (physicians) are going to say: well he plugged it into the machine and then here is what comes out! On the other hand, if you say: there is a lesion of three centimeters at the head of the pancreas, with a contact on the vein of less than 180 degrees, which means that it is a patient who is borderline, that is no stenosis of the celiac trunk, the surgeons who receive it will know whether they can potentially operate the patient or not! They know that you are a specialist because you say the vessels are like that, and that it matters in the surgical set-up. And so they trust you, because they know what you're talking about. And they say to themselves: this guy, he knows what he's talking about"
> (Radiologist N.1, Public Hospital 1)

Even more strikingly, the radiologist's vision work also involves highlighting elements that go beyond the visible. If the omission of certain anomalies is considered necessary, other omissions are considered potentially harmful (for the patient or for the radiologist's reputation). Sometimes, and depending on the circumstances, the report may indicate the absence of visibility of certain parts of the body that the radiologist considers important in clarifying a clinical question, in order to suggest additional examinations or to protect himself against "an increasingly litigious patient population" likely to sue the doctor for malpractice.

It is clear that medical reports, by fulfilling an important social function – namely the collective construction of the professional identity of radiologists – contribute to the stabilization of the visible (and the invisible). It is in fact in writing that the radiologist establishes what must be seen by their interlocutors. The radiologist's expertise therefore does not consist in detecting visible anomalies, but in a selection of elements according to the clinical context, the request of the physicians and professional objectives (like reputational ones).

We are beginning to see that the activity of identifying anomalies is carried out in a context where the definition of the pathological depends, among other things, on disparate information (knowledge of the disease, the type of question that justifies the examination, clinical data available to the radiologist, the identity of the interlocutor) which are not confined to the image alone. The machine's failure to take this information into account is one of the first pitfalls we have observed in the use of AI tools in clinical practice.

5.3 Super-normality and alarm fatigue

Especially in the case of cancer, radiologists are required to identify anomalies on images often displaying organs that are impacted by pharmaceutical modifications (due to the treatments) or by surgery. In the images of patients which some radiologists refer to with the French adjective "techniqués", medical devices (drains, catheters) are visible and taken into account in the interpretation. Radiologists are thus constantly required to discriminate anomalies in contexts where the morphological appearance of bodies is altered by professional interventions and their visible traces. A senior radiologist explained, for example, how immunotherapies can cause venous thrombosis, and this should be verified in the image.

> "So, there are treatments that can cause vascular lesions, so at that point you have to zoom in and you're really going to look at the vessels, the veins, to see if there isn't something ... (...) In fact, when you close the windows like that, you can't see very well, and you have no way of knowing! But I know this because... I was taken in by these lesions which are therapeutic lesions in fact, linked to medication, incidental, which are not symptomatic and that... well if you haven't seen them before, then you know that if you don't widen the screen view, you won't see them!"
> (Radiologist N.3, Public Hospital 1)

To differentiate pathological anomalies from iatrogenic ones, the acquisition of knowledge is often necessary. This knowledge is built during the vision work. Once again, knowledge of the patient's clinical information (for example, whether a type of treatment is being administered) guides the radiologist's gaze. The recognition of certain lesions caused by medical interventions allows them to better isolate iatrogenic anomalies from the ones which attest to the progression of the disease.

This process of visual discrimination, at the heart of the vision work, represents a major limitation in the use of AIT and AIM. Images that are modified by therapeutic treatments or technical objects introduced into the patient's body are not correctly identified by the machine, which may classify them as pathological. This leads to the detection of false positives each time the patient's body does not meet the criteria of normality used by the software, which could in turn lead to alarm fatigue (Reyna et al. 2022) or automation bias (Skitka et al. 2000) over time. This problem, which concerns the inability of the machine to recognize modified bodies - by interventions, treatments or devices - constitutes a major pitfall in a context such as cancer screening where breast specialists are often required to monitor operated patients. The radiologist is in fact constantly confronted with alarms concerning the signs of a scar on the mammogram. The software often displays a high score (9 or 10) for images where the scar trace is mistaken for malignant abnormalities. A breast specialist explained:

> "(IAM) sometimes detects masses and assumes that they are pathological, but I often reinterpret them and say, 'No, it's scar tissue'. It doesn't know the patient's history, it can't make that assessment, saying it's a scar mass. It's very subtle, not very specific, so all the breasts operated on with scars, it'll tell us that there's an opacity, there's cancer, and it's right... Except that we actually know that this opacity is scar tissue and so we're able to correct the diagnosis."
> (Radiologist N.15, Cancer screening center 1)

The detection tools discussed here involve criteria of normality that do not adhere to those used by physicians in their daily work. In particular, the "pathologization" of iatrogenic lesions that is driven by the automatic classification shows the gap between a machine's definition of the "normal" body as it is found in the training data and the one directing the medical action of senologists. In fact, automatic classification embodies a "super-normality" (Beaulieu 2001), a sociological concept initially coined to refer to the choice of experimental subjects in neuroimaging research, which embodies a conception of normality in which non-pathological states (being pregnant, having undergone trauma or drug treatments) are excluded. Similarly, here, the software detects false positives whenever the body does not meet the machine's criteria of normality which are limited to non-operated or non-altered bodies.

The other side of the coin of alarm fatigue regards errors of omission. A minimal irregularity ("a little trifle"), which previously would have led to further investigation through an additional examination, is sometimes no longer taken seriously if the machine has not reported it.

The tool therefore not only works in focusing the doctor's eye on certain lesions, but also in diverting their attention to anomalies that they would otherwise have taken into consideration. In particular, a radiologist's intuitive medical concern based on reasons that are difficult to explain is more likely to be overlooked in favour of the machine's opinion, which is considered to be more 'objective' than this type of intuition. One radiologist in a cancer center explains:

> "Sometimes I can see opacities that it (IAM) does not detect, so it is reassuring for us to know that it is a glandular opacity, something that is not very dense, that is not suspicious..."
> (Radiologist N.16, Cancer screening center 2)

This confidence is also explained by the decision-making context, characterised by uncertainty. While it would still be possible for the radiologist to request additional examination, the automatic detection tool is considered a way to reassure them about doubts surrounding certain anomalies. Doubt becomes crucial in a context, such as that of breast cancer, where the radiologist's job is to "track" anomalies which are linked to "complex entities" (Bourret and Rabeharisoa 2008) where it is not easy to make a decision.

5.4 From tacit knowledge to over-structuring

The automation of vision work is challenged by heterogeneous reading practices that vary from person to person, each radiologist relying, among other things, on "tacit knowledge" (Collins 1974, 2001) which guides their diagnostic activity. In our observations of the activity of several radiologists, we identified various types of tacit knowledge.

For example, tacit knowledge can be used so as not to miss anomalies and to distribute the tasks by anticipating possible shortcomings in attention that could lead to negligence. In cases of monitoring and detection of hepatocellular carcinoma (HCC), one of the radiologists explained that he usually starts by viewing the areas that seem less problematic to him. He knows that liver exploration is going to require more time and more cognitive resources, and that he might not examine the rest of the body as closely if he does that first. In fact, once attention is directed to an organ that is likely to be more affected by the disease, it is difficult for him to focus with the same accuracy on the other parts of the body.

Various image-reading heuristics influence the trajectory of a radiologist's gaze when examining slices. For example, radiologists may scan the image in a specific direction that they will always follow, like a script that they repeat to guide their work. As not to overlook any part of the image, there are radiologists who divide organs into several sections when the surface area is too large to achieve sufficient efficiency in locating anomalies.

Other knowledge is involved to better target anomalies and testifies, for example, to a particular use of measurements in the reading process. The eye is in fact trained to recognise the pathology according to the quantitative values mentioned in the guidelines and learned by the radiologists. The pathological character of a lymph node depends, among other things, on its location and size, and this information is therefore indispensable to see an anomaly as such. However, even if an entity exceeds a value that confers it a pathological character, knowledge of the clinical case may lead to the intentional rejection of these measurements in the reading process. Certain treatments may cause a change in the size of the lesion, which leads the radiologist to consider these quantitative dimensions in the light of these circumstances and context. Intentional neglect of quantitative values is then informed by experience and implicitly guides the vision work.

Other habits and heuristics acquired through experience improve reading, but their empirical and personal nature makes them less acceptable. A radiologist, for example, developed a sensitivity to certain signs which he adopted as visual landmarks. Notably a retraction in the lower part of the liver helps him assess the condition of the organ at a glance. He refers this to a practice that has "no scientific value" as if to emphasise its low level of empirical evidence. Likewise, intuition is an element that is involved in the interpretation work and guides the radiologist's view:

> "There are a lot of recommendations and indications to follow. But in the end there is a patient and a doctor in front of him or her. And even if sometimes the patient doesn't tick all the alarm boxes, sometimes instinct and experience mean that they'll still give us cause for concern, and there are a whole bunch of things to keep the human being in check, but a lot of the time it's still experience..."
> (Radiologist N.17, Cancer screening center 3)

Concretely, this manifests itself, for example, by taking into account a set of elements contained in the medical record and possibly resulting from the patient's prior knowledge. Heterogeneous elements combined with knowledge of the disease and its expression sometimes lead the radiologist to go beyond what is visible (or not) on the picture. A radiologist explains:

> " There are plenty of things that come into play, it's even intuition actually, when discussing with the patients (...) there are plenty of things I think that influence us and sometimes I don't know, without being able to really explain it, I say to myself: I don't know, there is something that doesn't match, it's discordant, there is something that I don't like and I will go a little further. I'm going to do an MRI, I'm going to bring the person back because there is something that doesn't reassure me and sometimes it is difficult to say exactly what it is."
> (Radiologist N.16, Cancer screening center 2)

This intuition, which is judged by radiologists as "subjective" because of its elusive nature, can also be nourished by what is communicated by the patient, by their remarks and worries, both verbally and with their body language. These dimensions of the medical experience remain difficult to formalise but represent an important part of the work with patients:

> "It's a matter of subtle anomalies. Sometimes it's just an area that is a little more dense, an area on palpation that is a little less soft, it's really sometimes very, very discreet, very tiny things, but it's all of these things which incline me to be worried or not."
> (Radiologist N.16, Cancer screening center 2)

This knowledge - based on experiential data such as the doctor's feelings and the patient's experience - although considered variable and subjective, is essential to the professional vision (Goodwin 1994) of the radiologist and to the achievement of an interpretation capable to fulfill both medical and social objectives. The reporting of radiologists does not show what is visible, but what is significant for the objective they have set. This includes minimising harmless anomalies, supporting a therapeutic action, establishing a relationship with other professionals, reassuring a patient, and for this they need a set of skills and knowledge and that they refine on a daily basis in the course of their individual and collective work.

From our observations and interviews emerges the fact that the acquisition and exercise of this knowledge can be disrupted by the introduction of AI tools for anomaly detection.

Among residents, for example, AIT causes some apprehension about the possible influence of automated classification on their learning process, through which they acquire their judgment skills: this highlights how the introduction of AI tools has different implications for radiologists at various stages of their careers. While still in a learning phase, future specialists fear they will be destabilized even before they have consolidated their knowledge and experience. As long as the X-ray image is opened on the screen at the same time as the machine's report, the solution adopted by some radiologists is to look away and close the automatic detection window to avoid reading it before taking their own reading (cf. the *Hound*, or Human-first protocol in Cabitza et al. 2023c). The color associated with the classification of the images (green for "normal", red for "abnormal"), which is instantly visible on the screen, can in fact immediately inform the radiologist of the algorithm's classification, swaying their judgment from the start.

> "(...) When I look at a chest X-ray I look first of all at that, and then that, and again that, but if I know that there is something abnormal when I open it (the image) I will say to myself: where is it abnormal? you see? (...)"
> (Resident doctor N.10, Public Hospital 2)

In fact, the mobilization of reading heuristics and experiential knowledge, now considered to be "subjective" in relation to the machine's process, can be disrupted and this also concerns senior radiologists. This is what a radiologist using AIM in her work explains:

> "AIM will find an increased intensity, on which the second viewer will probably draw attention to and this allows me to say I have to take an extra picture... and sometimes I want to take an extra picture for a small detail... a minor asymmetry... It's all very subjective... and when I see that AIM hasn't analysed it or hasn't noticed it, I say to myself that there's no need to do it... I can skip it without any problems..."
> (Radiologist N.13, Private Hospital)

AI is also seen here as a tool to reduce the anxiety that the breast specialist may feel when they have to communicate uncertainty to the patient. Invoking technology seems to reframe the emotional intensity of a diagnosis that may be ambiguous or doubtful, evoking instead a more objective and therefore more controllable register:

> "Earlier, when I said... the lady who came to check the classifications and I told her that the AI, there is an artificial intelligence tool which also uh... did not show any high probability, and that we can just monitor, that's reassuring! (...) There is a real human experience behind which is not simple. And frankly, if AI could help us decide on these cases, that would be great."
> (Radiologist N.13, Private Hospital)

Because of the type of (tacit) knowledge, the emotional dimensions of medical work and the uncertainty that characterizes the assessment of certain anomalies, some breast specialists seem to use automated detection as a means of reintroducing judgment into a more "objective" - and therefore reliable – framework.

This observation underscores the importance of understanding the intricate relationship between the radiologists' professional norms and the AI tools they employ. In our effort to ensure the successful appropriation of AI technologies to support vision work in radiography, we propose a series of design principles to uphold the existing practices of professionals as well as to deal with the inherent uncertainties in anomaly detection.

6 Implications for design

Building on our fieldwork findings, we recognize that the introduction of technologies in work settings influences existing practices: technology is not merely a tool, but something that actively shapes and is shaped by its interactions with users (Berg 1999; Winthereik and Vikkelsø 2005; Carroll et al. 1991; Shneiderman 2022). The mutual influence between technology and practice plays a pivotal role in what is commonly termed as 'appropriation' (Dix 2007).

Appropriation is a nuanced process (Debono et al. 2013; Simone et al. 2019) whose success hinges on several factors: these include the user-friendliness of the technology, its alignment with the genuine needs of users, and its ability to present clear advantages. The context of software usage can either facilitate or hinder appropriation, especially considering the uncertainties tied to the software, the role of tacit knowledge required for the task, and the user's need for reassurance (Turrini and Bourgain 2021).

Another equally important factor to trust-building and clinical adoption of technology is the alignment with pre-existing collaborative practices. For example, Alby et al. (2015) described three cooperative strategies deployed by clinicians to deal with complexity, limits of knowledge and cognitive difficulties: joint interpretation, intersubjective generation and validation of hypotheses, and postponing the diagnostic decision. This leads to the possibility of designing systems that are specifically built upon such unspoken yet recognized practices in an iterative way (Berg 1999) by promoting early user involvement and various forms of participatory design (Bratteteig and Wagner 2016), such as co-design (Steen 2013), continuous design (Joshi and Bratteteig 2015) or similar "user-led processes of adaption and adoption" of systems that express the professional vision of the users and afford their work practices (Hartswood et al. 2002). All these efforts can make appropriation more likely to happen as users play an active role, as they can express their needs and provide their suggestions when the design process is still ongoing, and at the same time they can anticipate the embedding of the technology into their situated practices (Cabitza and Simone 2015).

The CSCW community has consistently expressed a desire to "avoid underrating the skills and competencies that are required in even the most routine of tasks", as noted by Hartswood et al. (2003a), championing the importance of designing systems that complement and enhance human skills rather than replace them. This ethos is rooted in a recognition of the complex nature of collaborative work, and an understanding of technology as a tool for empowerment rather than displacement. Aiming to contribute to the discussion over "what to automate and what to leave to human skill and ingenuity" (Hartswood et al. 2003a), we leverage upon our observations and the preexisting literature to suggest some design solutions aimed at enabling hybrid decision-making agencies.

These solutions are informed by the concept of Frictional AI: an AI design principle that leverages programmed inefficiencies (Cabitza et al 2019a) to stimulate human cognitive activation and mitigate overreliance on automated systems (Cabitza et al. 2024). This approach is not about impeding efficiency but about enriching the diagnostic process with layers of critical thinking, professional intuition, and ethical consideration that are prompted by aptly-engineered phenomena of cognitive friction (Cooper 1999), instead of pushing for ever-faster and optimized interactions.

In this section, we will explore how the principles of Frictional AI can inform the design of AI systems in a way that aligns with the CSCW tenets. We will discuss how openness, multiplicity, and auxiliarity, as aspects of Frictional AI, can foster an environment where radiologists are not mere operators of a system but engaged, critical thinkers, and decision-makers. This involves designing AI systems that reflect the complexity of radiological vision work and adopting them while respecting the irreplaceable value of human expertise. The summary Table 2 succinctly presents our three design-oriented principles and explicitly provides linkage to the themes discussed in Section 5.

6.1 Openness

In the context of Artificial Intelligence, the term *openness* evokes imagery such as opening the black box (where openness hints to transparency and explainability) and open source (i.e., software that is freely available for use and redesign). However, we intend openness in the sense of being open to the (undatafiable) surrounding context in all its complexity and uncertainty, as well as being an open loop, i.e. allowing human input and feedback.

One of the main challenges in the effective use of AI in radiology is ensuring the appropriate integration into the radiologist's decision-making process of AI advice that is opaque (Lebovitz et al. 2022) and that is lacking in real-world, contextual interpretation possessed by radiologists, who ought to complement the AI output with their own professional vision (Hartswood et al. 2023b).

Not only should AI take into account more contextual data such as patient age, patient and family history (as suggested by a radiologist in Section 5) as well as iatrogenic anomalies (see Section 5): AI should be accorded the role of a case-mining instrument, an extension of the memory of professionals that, in the diagnostic task, may allow doctors to evaluate areas of concerns, or provide them with analogies to and differences from previous cases, to understand how to proceed in a group decision-making setting and make sense of these AI-generated outputs by using their own professional judgment. We reported in Section 5 how the software's advice, despite being consistent with the radiologist's interpretation, did not take

Table 2. Summary table presenting a short definition for each design principles, the related themes from Section 5, and some examples of the excerpts from interviews that informed the design process.

Design principle	Design requirements	Related themes	References to the text
Openness	AI systems should be transparent, sensitive to the surrounding context, considering its inner complexities, ambiguities and uncertainties, and allowing human input and feedback, and encouraging users to integrate its outputs with their professional judgment, while also seeking additional contextual information beyond just the digital output.	'The Undatafiable Dimension', 'From tacit knowledge to over-structuring' and 'Purposeful Omission for Useful Representation': AI systems should be transparent about what they can and cannot capture, nudging users to move beyond clear-cut categories and seek additional information beyond the computer screen or allow the user to omit irrelevant elements (according to the context). It also touches upon the theme of 'supernormality' through the proposal of introduction of iatrogenic anomalies in training data.	"if I see an opacity on the mammogram, I'll look for it, I'll palpate it in that area, and I'll see that there's a mass, is it soft, is it hard?…" (Section 5); "Finally, what should be integrated into AI is the patient's history …family history…We will also take into account her age, of course…" (Section 5); "There are a lot of recommendations and indications to follow. But in the end there is a patient and a doctor in front of him or her …here are a whole bunch of things to keep the human being in check, but a lot of the time it's still experience…" (Section 5); "…you don't have to describe them…no need to because it's irrelevant, it has no influence, it won't change anything…" (Section 5)

Table 2. continued

Multiplicity	AI outputs should be presented in multiple, varied and complementary forms, by allowing the interpretation and integration of these outputs in medical decision-making on the basis of the inherent complexity of the medical phenomenon, supporting more nuanced decision-making, and reducing reliance on singular algorithmic recommendations.	'Purposeful Omission for Useful Representation', 'The Undatafiable Dimension': emphasis on the value and complementary AI outputs, as well as taking into consideration contextual elements and applying comparison-based reasoning. Multiplicity is connected to 'Alarm Fatigue' in its proposed provision of multiple cognitive-activating outputs to act as a countermeasure to heuristic disruption and false alarms that frustrate professionals and hinder novices from effective learning.	"Above all, it is the comparison that will help us" (Section 5); "Don't give me decisions! give me probabilities! …I like more the IAM approach, where they give me a score, and it keeps the decision to a certain extent to the human being. And he comes to help me by saying, he sees that there is a 60% or 80% or 95% probability, and that alerts me! But the decision is mine" (Section 6.2); "If we are in the initial assessment of a cancer and the patient's liver, for example, shows a metastasis and other lesions that are cysts, we will describe them …We will describe them even if they are of no interest because they are present with something that is …malignant" (Section 5).

Table 2. continued

Auxiliarity	AI should serve as a supportive tool, which augments and complements human expertise without replacing it, and by performing ancillary tasks that facilitate diagnostic reasoning and empower professionals to make informed decisions, so as to prioritize the professional insights and intuitions of the user.	'From Tacit Knowledge to Overstructuring', 'super-normality and Alarm Fatigue': AI is positioned as a supportive tool that enhances rather than disrupts human expertise. It also touches upon the 'Purposeful omission for useful representation' theme in that auxiliarity protects and puts at the forefront professional discretion and heuristics in the preparation of the radiological report.	"When I see that AIM hasn't analysed it or hasn't noticed it, I say to myself that there's no need to do it …I can skip it without any problems…" (Section 5); "Sometimes I can see opacities that it (IAM) does not detect, so it is reassuring for us to know that it is …not suspicious" (Section 5); "(the medical report) shouldn't be a stereotypical thing …the surgeons who receive it will know whether they can potentially operate the patient or not! They know that you are a specialist" (Section 5); "(IAM) sometimes detects masses and assumes that they are pathological, but I often reinterpret them and say, 'No, it's scar tissue'. It doesn't know the patient's history, it can't make that assessment …Except that we actually know that this opacity is scar tissue and so we're able to correct the diagnosis" 5 "There are plenty of things that come into play, it's even intuition actually, when discussing with the patients" (Section 5); "The machine, the AI must stay in its place. It shouldn't tell us to do this, do that…" 6.2

into consideration the same elements to produce the classification. The aim would be to lower the feeling of disappointment over diverging purported "mindsets" between physician and AI system, enabling a constructive mental representation of the potentiality and shortcomings of AI advice.

A useful tool for this end is the representation of uncertainty, although "many visualization authors choose not to visualize uncertainty" (Hullman 2019). In particular, we pursue vagueness as a way to provide decision makers with a less cognitive and more immediate, concrete feeling of the uncertainty that affects a specific condition or prospect. It reflects the frictional principle of sustaining an "underdetermined environment conducive to human flourishing" (Frischmann and Selinger 2018), promoting thoughtfulness and cognitive enhancement. Vague visualizations, for example, have been shown to convey an appropriate perception of uncertainty without relying on numerical values or symbolic forms (Assale et al. 2020). This emphasis on nuanced interpretation and uncertainty also addresses ethical and sensitivity considerations: the vision work of radiologists also involves ethical judgment, consisting in "the need to accompany patients in the difficult decisional contexts brought on by our own technologies, the need to help patients understand that the uncertainty regarding the pertinence of findings cannot be simply resolved, and the professional requirement that we should not act as if we know" (Raymond and Trop 2007).

As for being an open loop, the capability of updating its reference data and correlative models would make an AI system capable of coping with an ever-changing environment and mitigate the risk of error due to concept drift (Zenisek et al. 2019). This entails a tighter relationship with users, which is not just unidirectional – the machine that gives humans advice – but rather it is bidirectional, in that the user provides feedback on the correctness and usefulness of the recommendations and the relevance of any explanations, with the machine that updates or recalibrates its estimates accordingly.

Yet, in order not to interfere with the vision work, systems such as AIM and AIT should ideally provide the least possible amount of diagnostically relevant information, as to afford preexisting reading heuristics without disrupting them and avoid information overload (Bawden and Robinson 2020) and alarm fatigue (Sendelbach and Funk 2013).

Managing information overload (van Leeuwen et al. 2022) is akin to identifying the right portion of information in the large amounts provided, while accounting for the missing one. Therefore, other than integrating more data, solutions to alarm fatigue can also be found by incorporating different design features.

Probabilistic outputs can reduce the likelihood of false-positive results. Contextual information about the patient's medical history and other relevant factors can also help radiologists interpret the AI system's output more accurately, and visualizations can highlight regions of potential concern.

In line with the principles of calm technology (Weiser and Brown 1996) and slow computing (Kitchin and Fraser 2020), we suggest nudging users to move beyond clear-cut categories and seek additional information beyond the computer screen (and behind the screen, by talking more with patients). Referring back to Section 5, we highlight the statement "if I see an opacity on the mammogram, I'll look for it, I'll palpate it in that area …Sometimes the patients arrive and say: I can feel something." Openness values and respects the tactile and emotional aspects in clinical practice, as well as the radiologist's professional experience.

6.2 Multiplicity

As for the presentation of the output, our observations confirm the findings of Kohli-Laven et al. (2011): binary or dichotomous results are less accepted by physicians than those that express probabilities and leave more room for manoeuvre for medical decision making, such as the *docile* CAD machine prompts reported by Hartswood et al. (2023b), which highlight areas of interest without suggesting any specific action. In accordance with the concept of openness, is important to represent AI output not as the one of an oracle speaking of truth, but rather as an estimate, affected by uncertainty, like in the case of probabilities. This could include visualisations that represent the uncertainty associated with the output or confidence intervals that indicate the level of certainty associated with the result.

In our interviews we collected statements that are consistent with these findings. When a senologist compared IAM to a detection tool she previously tried, she noted:

> "I didn't like the idea (…) that (the previous software) would say 'call back' or 'don't call back the patient'. Don't give me decisions! give me probabilities! The machine, the AI must stay in its place. It shouldn't tell us to do this, do that…That's what gives me a probability. And I didn't want to have on the screen: 'call back the patient'. Maybe it's the same thing, because it has detected something, but I like more the IAM approach, where they give me a score, and it keeps the decision to a certain extent to the human being. And he comes to help me by saying, he sees that there is a 60% or 80% or 95% probability, and that alerts me! But the decision is mine."
> (Radiologist N.13, Private Hospital)

The senologist's frustration with the previous detection tool suggests that less dichotomous modes of expressing outcomes better fit the context of medical decision making and the distribution of responsibilities in patient care. Citing the Radiologist N.15, from the Cancer Screening center 1 (Section 5), "Above all, it is the comparison that will help us." This vouches for the usefulness of explanations in terms of sets of similar cases, whereby practitioners are presented with cases selected according to the highest correspondence following a similarity metric, and are provided with the final diagnosis for those similar cases (Cabitza et al. 2024). This function, akin to cognitive-forcing functions described by Buçinca et al. (2021), provides the health professional with the necessary support to reach a decision they would retain full accountability for, leveraging on their irreplaceable professional expertise and their unique grasp of the unquantifiable context surrounding the diagnosis. In fact, in line with Simone and Schmidt (1993), systems should not be "executable code but rather heuristic and vague devices to be interpreted and instantiated, maybe even by means of intelligent improvisation".

A requirement for multiplicity in AI output, rather than being aimed at confusing users, would reflect the inherent complexity and ambiguity of the phenomenon at hand, and thus mitigate phenomena such as automation bias, over-reliance, algorithmic aversion, or fallacious appeals to algorithmic authority. In a design-oriented perspective, a system could avoid proposing to users single pieces of advice or clear-cut categories, but would rather propose multiple and complementary indications, such as classes and the associated confidence scores, or conformal prediction intervals, or even possibly identical and diverging pieces of advice by different competing models such as models optimized for sensitivity, specificity, discriminative performance or utility (Lu et al. 2020).

Another example of multiple output is that of the *Evaluative AI* framework proposed by Miller (2023), "a machine-in-the-loop paradigm in which decision support tools provide evidence for and against decisions made by people, rather than provide recommendations to accept or reject." This function would be useful to users of the AIM and AIT technologies irrespective of their experience level, that is, both expert and novice health professionals; the cognitive activation elicited by multiple outputs could mean that expert radiologists would not be frustrated by heuristic disruption and false alarms, and novices would not be hindered from effective learning.

Multiplicity also spans the dimensions of configurability and customization as to accommodate the unique (and multiple) ways in which different radiologists conduct vision work. Communicating radiological insight via customized visualizations, adjusted sensitivity thresholds, or even certain AI suggestions being toggled on/off allows for maintaining the uniqueness and identity of each radiologist's reporting style.

6.3 Auxiliarity

As touched upon in the previous two design implications, the system should not (counterintuitively) make the radiologists' observation work "more efficient", nor the system should *necessarily* be too effective: as affirmed by two radiologists, noticing opacities that were not detected by the AI system led to a sense of reassurance on their benign nature, rather than arousing a constructive sense of doubt. According to Radiologist 15 in Section 5, "Sometimes I can see opacities that it (IAM) does not detect, so it is reassuring for us to know that it is a glandular opacity, something that is not very dense, that is not suspicious...". Likewise, for Radiologist 13 in Section 5, "when I see that AIM hasn't analysed it or hasn't noticed it, I say to myself that there's no need to do it…I can skip it without any problems…". Moreover, the system should not direct the specialists' professional gaze, breaking tacit reading heuristics and other work-oriented infrastructures, i.e. "highly complex and specialized practices whose properties are largely hidden for those who are not members of these communities (and which also the members are unconscious about)" (Hanseth and Lundberg 2001).

Human-AI collaboration protocols (HAI-CP) are a useful tool to address these concerns. They are representations of how humans and AI collaborate (or humans leverage the output of generative machines) to stipulate and evaluate how humans and AI can collaborate in cognitive tasks (Cabitza et al. 2023c). In designing HAI-CP, several features of the interaction between clinicians and their computational decision aids are stipulated (van Berkel et al. 2021), including the modality in which the AI output is presented to the user and the decision-making step at which the result is provided (e.g., before or after a first decision was made by the clinician, as in Bertrand et al. 2022). Other important considerations include the availability of eXplainable AI (XAI) solutions (e.g., feature rankings, pixel attribution maps, textual justifications), the calibration of the system (Vodrahalli et al. 2022), and its sensitivity or specificity (Cabitza et al. 2020). Echoing the configurability inherent to the previously explored concept of *multiplicity* (Section 6.2), collaboration protocols hint at the "radical conception of CSCW and CSCW systems" which advocates for these systems to offer an environment where users can create and adjust coordination mechanisms suitable for their specific context (Schmidt 2000, 1991; Schmidt and Simone 1996). These protocols are proposed as a means to identify the optimal conditions for AI to enhance human diagnostic skills in a particular work setting, while avoiding dysfunctional responses and cognitive biases that can undermine decision effectiveness.

As already mentioned in Section 5, introducing AI tools can cause apprehension among residents and future specialists, who fear that their yet-to-be-consolidated judgment skills may be influenced by automated classification, as in the case of the radiologist who looked away and closed the automatic detection window to avoid reading the machine's report before taking their own reading. A Human-First collaboration protocol as the one investigated in Cabitza et al. (2023c), also called *second-opinion* protocol, can help to ensure that radiologists rely on their own judgment and expertise, while still benefiting from the use of AI tools. However, these protocols are not immune to specific cognitive biases such as algorithmic aversion and conservatism bias, while AI-First protocols are associated with higher diagnostic accuracy than Human-First protocols (Cabitza et al. 2023c). Human-first protocols, supported by an analysis of reliance patterns (Cabitza et al. 2023a), have the advantage of enabling long-term technovigilance (Cabitza and Zeitoun 2019) of the effects of automation on human decision performance.

In line with the design suggestion put forward in Section 6.1, AI should be designed as a case-mining instrument, which focuses on leveraging its strengths in data analysis and decision-making support.

This approach aligns with the main tenet of Auxiliary (or Adjunct) AI, whereby AI is recognized for its helpful role as a decision support while being relegated to ancillary tasks, giving absolute precedence to the professional insight of the user – in the words of Hartswood et al. (2023b), the machine would be a 'dumb colleague'. In this sense, AI would not act as an oracular agent but, rather, as a catalyst (Miller and Masarie 1990), with a role of facilitating diagnostic reasoning and empowering physicians to make informed decisions, even going against the AI diagnosis (refer back to Radiologist N.15 in 5, "we actually know that this opacity is scar tissue and so we're able to correct the diagnosis") and protecting the value of professional intuition (see Radiologist N.16 in Section 5, "…there are plenty of things I think that influence us and sometimes I don't know, without being able to really explain it, I say to myself: I don't know, there is something that doesn't match, it's discordant, there is something that I don't like and I will go a little further").

Following the adjunction approach, AI systems are to be viewed as "supertools and active appliances, rather than teammates, partners, and collaborators" (Shneiderman 2021), with its recommendation being one of several considerations to be discussed in a group decision. Medical AI tools could allowing for shared annotations, real-time discussions, or second-opinion consultations directly within the platform, in order to promote and protect the collaborative nature of medical interpretation. Another crucial issue addressed by the design principle of auxiliarity is that of deskilling (Sambasivan and Veeraraghavan 2022; Chen et al. 2021), or the gradual loss of skills caused by over-reliance on automated advice.

This is why, following Sterponi et al. (2017), we urge caution in the uncritical adoption of digital technologies and we offer our perspective on the intentional design of HAI-CP in medical decision-making, aiming at incorporating the benefits of digital tools while preserving the semiotic resources of medical experiential and professional knowledge, which can be augmented, but not replaced, by AI (Van den Broek et al. 2021).

7 Conclusion

In this article, we made the point that the nuanced distinction between normal and pathological that emerges during vision work, that is the process of medical images interpretation in light of socio-professional norms and heuristics, is beyond the reach of current automatic AI detection tools. In fact, recognition of visible anomalies in radiology is based on quantifiable, codifiable and experiential elements, as well as on the radiologist's relationship with patients and other professionals, and on the application of formal as well as tacit knowledge.

The classification of images (normal/abnormal) - even by tools based on deep learning and despite their almost superhuman accuracy – is often opposed to the radiologists' process of recognising pathological anomalies, where diverse information helps them to make sense of what they see.

The radiologists we interviewed also insisted on their professional "role" and the value of their vision work, in regard to the ability to answer the prescribing physician's question while keeping a vigilant outlook for other signs they consider suspicious. Making the abnormality visible from a professional perspective is not only describing what one sees, but also building one's reputation with the help of the medical report by affirming one's competence to the patient and to other colleagues (Anichini and Geffroy 2021).

Vision work also entails the purposeful omission of visible anomalies whose description is considered unnecessary, as well as the reporting of elements that sometimes go beyond what is visible, but which allow the radiologists to define their intervention and assume their responsibility towards patients and other professionals.

Our analyses show the conditions in which the emergence of visible anomalies highlighted by automatic detection does not match the heuristics adopted by radiologists. The physician considers elements and features that cannot be quantified by the machine and that involve the use of senses employed both in the clinical examination of the patient's body, and in the understanding of the patient experience. In assessing cancerous lesions, considering information like family history raises the question on the boundaries of the patient, as well as presenting challenges to the scope of automatic detection, which is limited to the individual body and its digitized representation.

In addition, in the vision work of mammograms or chest X-rays, the criteria defining "normal" bodies diverged from those encoded in automatic detection. By misinterpreting the images of operated bodies or altered ones (by drug therapies or medical devices), the machine led to a pathologization of iatrogenic anomalies, reinforcing the criteria of super-normality.

We have also found that AI tools are likely to influence the physician's judgment, especially when the medical decision would involve tacit knowledge or physician's perceptions grounding on their feelings and intuition. Our case studies suggest that these informal dimensions, considered by defenders of AI technologies to be a source of inter-subjective variability and "noise" (Kahneman et al. 2021) negatively affecting diagnosis, are, on the contrary, activated by radiologists in their vision work to make observations more efficient and effective (Cabitza et al. 2019b). The implementation of routines and reading heuristics, the distribution of tasks in a specific order, the definition of informal visual landmarks are operations essential to the identification of a pathological anomaly in radiographic images.

Our work therefore suggests a reflection around the way in which AI tools, in certain contexts, can lead to a disruption in the acquisition - or even the socialization (Nonaka et al. 1996) - of this knowledge and to the weakening of confidence in medical judgment and hence self-confidence.

We have also observed situations where automated advice can be more favorably received, particularly to resolve minor uncertainties. In these cases, the system meets the radiologist's need for reassurance, and is conceived as a "colleague" on whom to rely for a second opinion.

Our article has shown how technologies, to be considered "effective", must align with the objectives arising from the contexts of use. These objectives (technical, social, epistemic) are complex and are not necessarily compatible with algorithmic quantification and its modalities.

Given this complexity, empirical research across diverse medical domains becomes essential to identify the distinct challenges posed by technology integration in each setting. We have initiated this process by defining the social and professional norms associated with vision work, which could then inform the adaptation of automatic detection tools to local constraints variable within the specialty itself (depending on the imaging modalities used, pathologies investigated, medical establishments where the professionals work).

In conclusion, it is imperative to acknowledge a critical tension at the heart of our proposed design principles of openness, multiplicity and auxiliarity. Our recommendations, grounded in the principles of Frictional AI, advocate for a design ethos that may appear, at first glance, to diverge from the prevailing market trends of increasing specificity and certainty in AI systems. The current push in the industry is towards AI solutions that promise greater efficiency, definitive outputs, and, ostensibly, a higher return on investment (ROI). Yet, the value of our principles emerges most distinctly when viewed through the lens of their counterintuitive nature, which we began investigating in (Cabitza et al. 2023b; Natali et al. 2023; Cabitza et al. 2024). The essence of Frictional AI — promoting thoughtful engagement, embracing ambiguity, and valuing human expertise — might seem antithetical to the relentless drive for faster, more accurate and efficient AI. However, it is within this apparent contradiction that our recommendations find their most profound justification and potential. By aligning AI design more closely with the intuitive and collaborative dimensions of vision work practices of professionals, we can achieve a deeper and more effective appropriation of technology for radiological vision work.

Author Contributions

Giulia Anichini devised and conducted the ethnographic study, took the pictures, transcribed the interviews and wrote the first draft of this manuscript. Chiara Natali wrote the "Implications for Design" section, as well as deeply revised the initial draft and edited all the figures. Federico Cabitza edited the whole paper, wrote the introduction, revised the conclusion and provided the main insights to discuss in the Implications for Design section.

Funding

Open access funding provided by Universitá degli Studi di Milano - Bicocca within the CRUI-CARE Agreement. Agence Régionale de Santé (ARS) Pays de la Loire. F. Cabitza acknowledges funding support provided by the Italian project PRIN PNRR 2022 InXAID - Interaction with eXplainable Artificial Intelligence in (medical) Decision making. CUP: H53D23008090001 funded by the European Union - Next Generation EU. C. Natali gratefully acknowledges the PhD grant awarded by the Fondazione Fratelli Confalonieri, which has been instrumental in facilitating her research pursuits.

Availability of data and materials

Interview transcriptions available upon request to the first author: giulia.anichini@inserm.fr.

Declarations

Competing interests

The authors declare no competing interests.

Ethical approval

This study, focusing solely on professionals instead of patients, did not necessitate institutional ethical approval.

References

Aanestad, M., B. Edwin, and R. Mårvik. 2003. Medical image quality as a socio-technical phenomenon. *Methods of Information in Medicine* 42 (04): 302–306.

Alac, Morana. 2011. *Handling digital brains: A laboratory study of multimodal semiotic interaction in the age of computers.* MIT Press.

Alby, Francesca, Cristina Zucchermaglio, and Mattia Baruzzo. 2015. Diagnostic decision making in oncology: Creating shared knowledge and managing complexity. *Mind, Culture, and Activity* 22 (1): 4–22.

Amann, Klaus, and K. Knorr Cetina. 1988. The fixation of (visual) evidence. *Human Studies* 11 (2/3):133–169.

Anichini, Giulia, and Bénédicte Geffroy. 2021. L'intelligence artificielle á l'épreuve des savoirs tacites. analyse des pratiques d'utilisation d'un outil d'aide à la détection en radiologie. *Sciences Sociales et Santé* 39 (2): 43–69.

Assale, Michela, Silvia Bordogna, and Cabitza Federico. 2020. Vague visualizations to reduce quantification bias in shared medical decision making. In: *VISIGRAPP 2020 - Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 3: 209–216.

Bardram, Jakob E., and Claus Bossen. 2005. A web of coordinative artifacts: collaborative work at a hospital ward. In: *GROUP'05*, 168–176. ACM.

Bawden, David, and Lyn Robinson. 2020. *Information overload: An overview*. Oxford University Press.

Beaulieu, Anne. 2001. Voxels in the brain: Neuroscience, informatics and changing notions of objectivity. *Social Studies of Science* 31 (5): 635–680.

Berg, Marc. 1999. Accumulating and coordinating: Occasions for information technologies in medical work. *Computer Supported Cooperative Work* 8 (4): 373–401.

Berg, Marc, and Els Goorman. 1999. The contextual nature of medical information. *International Journal of Medical Informatics* 56 (1–3): 51–60.

Berlin, Leonard. 2014. Radiologic errors, past, present and future. *Diagnosis* 1 (1): 79–84.

Bertrand, Astrid, Rafik Belloum, James R. Eagan, and Winston Maxwell. 2022. How cognitive biases affect xai-assisted decision-making: A systematic review. In: *Proceedings of the 2022 AAAI/ACM conference on AI, ethics, and society,* 78–91.

Blois, Marsden S. 1980. Clinical judgment and computers. *New England Journal of Medicine* 303 (4): 192–197.

Bourret, Pascale. 2005. Brca patients and clinical collectives: New configurations of action in cancer genetics practices. *Social Studies of Science* 35 (1): 41–68.

Bourret, Pascale, and Vololona Rabeharisoa. 2008. Décision et jugement médicaux en situation de forte incertitude: l'exemple de deux pratiques cliniques à l'épreuve de la génétique. *Sciences Sociales et Santé* 26 (1): 33–66.

Bratteteig, Tone, and Ina Wagner. 2016. Unpacking the notion of participation in participatory design. *Computer Supported Cooperative Work (CSCW)* 25: 425–475.

Buçinca, Zana, Maja Barbara Malaya, and Krzysztof Z. Gajos. 2021. To trust or to think: Cognitive forcing functions can reduce overreliance on ai in ai-assisted decision-making. *Proceedings of the ACM on Human-Computer Interaction* 5 (CSCW1): 1–21.

Byrne, Paula, and Katrina Stengel. 2010. *Heartfelt and heartseen: making a diagnosis through coronary angiogram*. Ethnographies of Diagnostic Work: Dimensions of Transformative Practice 133–148.

Cabitza, Federico, and Carla Simone. 2015. Building socially embedded technologies: Implications about design. In: *Designing socially embedded technologies in the real-world*, Wulf V., K. Schmidt, D. Randall (eds.). Springer, London. https://doi.org/10.1007/978-1-4471-6720-4_11

Cabitza, Federico, and Jean-David Zeitoun. 2019. The proof of the pudding: In praise of a culture of real-world validation for medical artificial intelligence. *Annals of Translational Medicine* 7 (8): 161.

Cabitza, Federico, Raffaele Rasoini, and Gian Franco Gensini. 2017. Unintended consequences of machine learning in medicine. *Jama* 318 (6): 517–518.

Cabitza, Federico, Andrea Campagner, Davide Ciucci, and Andrea Seveso. 2019a. Programmed inefficiencies in DSS-supported human decision making. In: *Modeling Decisions for Artificial Intelligence: 16th International Conference, MDAI 2019, Milan, Italy,* 4–6, September 2019, Proceedings 16, 201–212. Springer.

Cabitza, Federico, Angela Locoro, Camilla Alderighi, Raffaele Rasoini, Domenico Compagnone, and Pedro Berjano. 2019b. The elephant in the record: On the multiplicity of data recording work. *Health Informatics Journal* 25 (3): 475–490.

Cabitza F., A. Campagner, F. Del Zotti, A. Ravizza, and F. Sternini. 2020. All you need is higher accuracy? On the quest for minimum acceptable accuracy for medical artificial intelligence. In: *Proceedings of the e-Health Procedings of the 12th International Conference on e-Health, Online*, 21–23.

Cabitza, Federico, Andrea Campagner, Riccardo Angius, Chiara Natali, and Carlo Reverberi. 2023a. Ai shall have no dominion: on how to measure technology dominance in ai-supported human decision-making. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*.

Cabitza, Federico, Andrea Campagner, Lorenzo Famiglini, Chiara Natali, Valerio Caccavella, and Enrico Gallazzi. 2023b. Let me think! Investigating the effect of explanations feeding doubts about the ai advice. In: *International Cross-Domain Conference for Machine Learning and Knowledge Extraction*, 155–169. Springer.

Cabitza, Federico, Andrea Campagner, Luca Ronzio, Matteo Cameli, Giulia Elena Mandoli, Maria Concetta Pastore, Luca Maria Sconfienza, Duarte Folgado, Marília Barandas, and Hugo Gamboa. 2023c. Rams, hounds and white boxes: Investigating human-ai collaboration protocols in medical diagnosis. *Artificial Intelligence in Medicine* 138: 102506.

Cabitza, Federico, Chiara Natali, Lorenzo Famiglini, Andrea Campagner, Valerio Caccavella, and Enrico Gallazzi. (2024). Never tell me the odds: Investigating pro-hoc explanations in medical decision making. *Artificial Intelligence in Medicine* 150 (3): 102819. https://doi.org/10.1016/j.artmed.2024.102819

Cai, Carrie J., Samantha Winter, David Steiner, Lauren Wilcox, and Michael Terry. 2019. "Hello AI": uncovering the onboarding needs of medical practitioners for human-ai collaborative decision-making. *Proceedings of the ACM on Human-computer Interaction* 3 (CSCW): 1–24.

Callon, Michel. 1986. Éléments pour une sociologie de la traduction: La domestication des coquilles saint-jacques et des marins-pêcheurs dans la baie de saint-brieuc. *L'Année Sociologique (1940/1948-)* 36: 169–208.

Carlin, Leslie, Helen Smith, Flis Henwood, Steve Flowers, Andrea Jones, Rebecca Prentice, and Ken Miles. 2010. Double vision: an exploration of radiologists' and general practitioners' views on using picture archiving and communication systems (pacs). *Health Informatics Journal* 16 (2): 75–86.

Carroll, John M., Wendy A. Kellogg, and Mary Beth Rosson. 1991. Designing Interaction: Psychology at the Human-computer Interface. *The task-artifact cycle*, 74–102.

Chen, Yaru, Charitini Stavropoulou, Radhika Narasinkan, Adrian Baker, and Harry Scarbrough. 2021. Professionals' responses to the introduction of ai innovations in radiology and their implications for future adoption: A qualitative study. *BMC Health Services Research* 21 (1): 1–9.

Christin, Angèle. 2020. The ethnographer and the algorithm: Beyond the black box. *Theory and Society* 49 (5–6): 897–918.

Cohn, Simon. 2007. Seeing and drawing: The role of play in medical imaging. *Skilled Visions: Between Apprenticeship and Standards* 6: 91–105.

Collins, Harry M. 1974. The tea set: Tacit knowledge and scientific networks. *Science Studies* 4 (2): 165–185.

Collins, Harry M. 2001. Tacit knowledge, trust and the Q of sapphire. *Social Studies of Science* 31 (1): 71–85.

Cooper, Alan. 1999. The Inmates are Running the Asylum. In: Arend, U.,E. Eberleh, and K. Pitschke (eds) Software-Ergonomie'99. Berichte des German Chapter of the ACM, 53. Vieweg+Teubner Verlag. https://doi.org/10.1007/978-3-322-99786-9_1

Coopmans, Catelijne, Janet Vertesi, Michael E. Lynch, and Steve Woolgar. 2014. *Representation in scientific practice revisited*. MIT Press.

Debono, Deborah S., David Greenfield, Joanne F. Travaglia, Janet C. Long, Deborah Black, Julie Johnson, and Jeffrey Braithwaite. 2013. Nurses' workarounds in acute healthcare settings: A scoping review. *BMC Health Services Research* 13 (1): 1–16.

Dix, Alan. 2007. Designing for appropriation. In: *Proceedings of HCI 2007 The 21st British HCI Group Annual Conference University of Lancaster, UK*, 1–4. https://doi.org/10.14236/ewic/HCI2007.53

Dodier, Nicolas, and Janine Barbot. 2016. La force des dispositifs. In: *Annales. Histoire, sciences sociales*, 71: 421–448. Cambridge University Press.

Dourish, Paul. 2004. What we talk about when we talk about context. *Personal and Ubiquitous Computing* 8: 19–30.

Frischmann, Brett, and Evan Selinger. 2018. *Re-engineering humanity*. Cambridge University Press.

Goodwin, Charles. 1994. Professional vision. *American Anthropologist* 96 (3): 606–633.

Grasseni, Cristina. 2004. Skilled vision. An apprenticeship in breeding aesthetics. *Social Anthropology* 12 (1): 41–55.

Hanseth, Ole, and Nina Lundberg. 2001. Designing work oriented infrastructures. *Computer Supported Cooperative Work* 10 (3–4): 347–372.

Hartswood, Mark, Rob Procter, and Linda J. Williams. 1998. Prompting in practice: How can we ensure radiologists make best use of computer-aided detection systems in screening mammography? *Digital Mammography: Nijmegen,* 13: 363–370. Dordrecht: Springer.

Hartswood, Mark, Rob Procter, Mark Rouncefield, and Roger Slack. 2002. Performance management in breast screening: A case study of professional vision. *Cognition, Technology & Work* 4: 91–100.

Hartswood, Mark, Rob Procter, Mark Rouncefield, and Roger Slack. 2003a. Making a case in medical work: Implications for the electronic medical record. *Computer Supported Cooperative Work (CSCW)* 12: 241–266.

Hartswood, Mark, Rob Procter, Mark Rouncefield, Roger Slack, James Soutter, and Alex Voss. 2003b. 'repairing' the machine: A case study of the evaluation of computer-aided detection tools in breast screening. In: *ECSCW 2003: Proceedings of the Eighth European Conference on Computer Supported Cooperative Work 14–18 2003,* 375–394. Helsinki, Finland: Springer.

Henderson, Kathryn. 1998. *On line and on paper: Visual representations, visual culture, and computer graphics in design engineering*. MIT press.

Huisman, Merel, Erik Ranschaert, William Parker, Domenico Mastrodicasa, Martin Koci, Daniel Pinto de Santos, Francesca Coppola, Sergey Morozov, Marc Zins, Cedric Bohyn, Ural Koç, Jie Wu, Satyam Veean, Dominik Fleischmann, Tim Leiner, and Martin J. Willemink. An international survey on AI in radiology in 1,041 radiologists and radiology residents part 1: Fear of replacement, knowledge, and attitude. *European Radiology* 31: 7058–7066.

Hullman, Jessica. 2019. Why authors don't visualize uncertainty. *IEEE Transactions on Visualization and Computer Graphics* 26 (1): 130–139.

Joly, Pierre Benoît. 2010. On the economics of techno-scientific promises. *Débordements Mélanges offerts à michel callon* 203–222.

Joshi, Suhas Govind, and Tone Bratteteig. 2015. Assembling fragments into continuous design: On participatory design with old people. In: *Scandinavian Conference on Information Systems*, 13–29. Springer.

Kahneman, Daniel, Olivier Sibony, and Cass R. Sunstein. 2021. *Noise: A flaw in human judgment*. Hachette UK.

Kitchin, Rob, and Alistair Fraser. 2020. *Slow computing: Why we need balanced digital lives*. Bristol University Press.

Kohli-Laven, Nina, Pascale Bourret, Peter Keating, and Alberto Cambrosio. 2011. Cancer clinical trials in the era of genomic signatures: Biomedical innovation, clinical utility, and regulatory-scientific hybrids. *Social Studies of Science* 41 (4): 487–513.

Kotras, Baptiste, Pauline de Pechpeyrou, and Bernard Quinio. 2021. Indices, algorithmes et chasseurs-cueilleurs: Hybridation et friction des épistémologies dans un dispositif d'archéologie computationnelle. *Réseaux* 4: 61–93.

Latour, Bruno. 1986. Visualization and cognition. *Knowledge and Society* 6 (6): 1–40.

Latzko-Toth, Guillaume. 2009. L'étude de cas en sociologie des sciences et des techniques. CIRST. https://depot.erudit.org/id/004067dd

Lebovitz, Sarah, Hila Lifshitz-Assaf, and Natalia Levina. 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.

Lincoln, Christie M, Ritodhi Chatterjee, and Marc H. Willis. 2019. Augmented radiology: Looking over the horizon, 1 (1): e180039 .

Lu, Danni, Chenyang Tao, Junya Chen, Fan Li, Feng Guo, and Lawrence Carin. 2020. Reconsidering generative objectives for counterfactual reasoning. *Advances in Neural Information Processing Systems* 33: 21539–21553.

Luff, Paul, Christian Heath, and David Greatbatch. 1992. Tasks-in-interaction: Paper and screen based documentation in collaborative activity. In: *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*, 163–170.

Lynch, Michael. 1988. The externalized retina: Selection and mathematization in the visual documentation of objects in the life sciences. *Human Studies*, 201–234.

Lynch, Michael. 1985. Discipline and the material form of images: An analysis of scientific visibility. *Social Studies of Science* 15 (1): 37–66.

Miller, Tim. 2023. Explainable ai is dead, long live explainable ai! Hypothesis-driven decision support. arXiv:2302.12389

Miller, R. A., and F. E. Masarie Jr. 1990. The demise of the greek oracle model for medical diagnostic systems. *Methods of Information in Medicine* 29 (01): 1–2.

Natali, Chiara, Lorenzo Famiglini, Andrea Campagner, Giovanni Andrea, La Maida Enrico, Gallazzi and Federico Cabitza. 2023. Color shadows 2: Assessing the impact of xai on diagnostic decision-making. In: *World Conference on Explainable Artificial Intelligence*, 618–629. Springer.

Nonaka, lkujiro, Takeuchi Hirotaka, and Umemoto Katsuhiro. 1996. A theory of organizational knowledge creation. *International Journal of Technology Management* 11 (7–8): 833–845.

Olivier de Sardan, Jean-Pierre. 1995. La politique du terrain. sur la production des données en anthropologie. *Enquête Archives de la revue Enquête* 1: 71–109.

Olivier de Sardan, Jean-Pierre. 2010. Anthropologie médicale et socio-anthropologie des actions publiques. dispositifs de recherche, commanditaires, réformes.... *Anthropologie & Santé Revue Internationale Drancophone d'anthropologie de la santé* 1.

Parmiggiani, Elena, Thomas Østerlie, and Petter Grytten Almklov. 2022. In the backrooms of data science. *Journal of the Association for Information Systems* 23 (1): 139–164.

Pentimalli, Barbara. 2020. L'œil diagnostique. voir et apprendre à voir dans un laboratoire d'hémodynamique. *Revue d'anthropologie des Connaissances* 14 (14-3).

Perrotta, Manuela. 2013. The study of technoscientific imaging in sts. *TECNOSCIENZA: Italian Journal of Science & Technology Studies* 3(2): 163–176.

Pesapane, Filippo, Marina Codari, and Francesco Sardanelli. 2018. Artificial intelligence in medical imaging: Threat or opportunity? Radiologists again at the forefront of innovation in medicine. *European Radiology Experimental* 2: 1–10.

Pinto dos Santos, Daniel, Daniel Giese, Sebastian Brodehl, SeungHun Chon, W. Staab, R. Kleinert, D. Maintz. and B. Baeßler. 2019. Medical students' attitude towards artificial intelligence: A multicentre survey. *European Radiology* 29: 1640–1646.

Piras, Enrico Maria, and Alberto Zanutto. 2010. Prescriptions, x-rays and grocery lists. designing a personal health record to support (the invisible work of) health information management in the household. *Computer Supported Cooperative Work (CSCW)* 19: 585–613.

Prasad, Amit. 2005. Making images/making bodies: Visibilizing and disciplining through magnetic resonance imaging (mri). *Science, Technology, & Human Values* 30 (2): 291–316.

Putha, Preetham, Manoj Tadepalli, Bhargava Reddy, Tarun Raj, Justy Antony Chiramal, Shalini Govil, Namita Sinha, Manjunath KS, Sundeep Reddivari, Ammar Jagirdar, Pooja Rao, and Prashant Warier. 2018. Can artificial intelligence reliably report chest x-rays?: Radiologist validation of an algorithm trained on 2.3 million x-rays. arXiv:1807.07455

Raymond, Jean, and Isabelle Trop. 2007. The practice of ethics in the era of evidence-based radiology. *Radiology* 244 (3): 643–649.

Reyna, Matthew A., Elaine O. Nsoesie, and Gari D. Clifford. 2022. Rethinking algorithm performance metrics for artificial intelligence in diagnostic medicine. *JAMA* 328 (4): 329–330.

Rodríguez-Ruiz, Alejandro, Elizabeth Krupinski, Jan Jurre Mordang, Kathy Schilling, Sylvia H. Heywang-Köbrunner, Ioannis Sechopoulos, and Ritse M. Mann. 2019. Detection of breast cancer with mammography: Effect of an artificial intelligence support system. *Radiology* 290 (2): 305–314.

Roepstorff, Andreas. 2007. *Navigating the brainscape: When knowing becomes seeing, 191*. Skilled Visions: Between Apprenticeship and Standards.

Rouncefield M., M. Hartswood, R. Procter, R. Slack, and A. Voss. 2003. "Reparing" the machine: A case study of evaluating computer aided detection tools in breast screening. In: *European Conference on Computer Supported Cooperative Work*.

Sambasivan, Nithya, and Rajesh Veeraraghavan. 2022. The deskilling of domain expertise in ai development. In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 1–14.

Sardanelli, Francesco, and Giovanni Di Leo. 2009. Biostatistics for radiologists: Planning, performing, and writing a radiologic study. Springer Science & Business Media.

Saunders, Barry F. 2008. CT suite: The work of diagnosis in the age of noninvasive cutting. Duke University Press.

Schmidt, Kjeld. 1991. Riding a tiger, or computer supported cooperative work. In: *Proceedings of the Second European Conference on Computer-Supported Cooperative Work ECSCW'91*, 1–16. Springer.

Schmidt, Kjeld. 2000. The critical role of workplace studies in CSCW. Cambridge University Press Cambridge.

Schmidt, Kjeld, and Carla Simone. 1996. Coordination mechanisms: Towards a conceptual foundation of cscw systems design. *Computer Supported Cooperative Work (CSCW)* 5: 155–200.

Seaver, Nick. 2015. The nice thing about context is that everyone has it. *Media, Culture & Society* 37 (7): 1101–1109.

Sendelbach, Sue, and Marjorie Funk. 2013. Alarm fatigue: a patient safety concern. *AACN Advanced Critical Care* 24 (4): 378–386.

Shneiderman, Ben. 2021. Human-centered ai: A new synthesis. In: *Human-Computer Interaction–INTERACT 2021: 18th IFIP TC 13 International Conference, Bari, Italy, August 30–September 3, 2021, Proceedings, Part I 18*, Springer, 3–8.

Shneiderman, Ben. 2022. *Human-centered AI*. Oxford University Press.

Simone, Carla, and Kjeld Schmidt. 1993. *Computational Mechanisms of Interaction for CSCW.* Lancaster University.

Simone, Carla, Angela Locoro, and Federico Cabitza. 2019. Drift of a corporate social media: The design and outcomes of a longitudinal study. In: *Organizing for the Digital World*, 189–201. Springer.

Skitka, Linda J., Kathleen Mosier, and Mark D. Burdick. 2000. Accountability and automation bias. *International Journal of Human-Computer Studies* 52 (4): 701–717.

Slack, Roger, Mark Hartswood, Rob Procter, and Mark Rouncefield. 2016. Cultures of reading: On professional vision and the lived work of mammography. In: *Orders of ordinary action*, 175–193. Routledge.

Slack, Roger S., Rob Procter, Mark Hartswood, Alexander Voss, and Mark Rouncefield. 2010. *Suspicious minds?, 227–244*. Ethnographies of Diagnostic Work: Dimensions of Transformative Practice.

Star, Susan Leigh, and Anselm Strauss. 1999. Layers of silence, arenas of voice: The ecology of visible and invisible work. *Computer Supported Cooperative Work (CSCW)* 8: 9–30.

Steen, Marc. 2013. Co-design as a process of joint inquiry and imagination. *Design Issues* 29 (2): 16–28.

Sterponi, Laura, Cristina Zucchermaglio, Francesca Alby, and Marilena Fatigante. 2017. Endangered literacies? affordances of paper-based literacy in medical practice and its persistence in the transition to digital technology. *Written Communication* 34 (4): 359–386.

Suchman, Lucille Alice. 2007.Human-machine reconfigurations: Plans and situated actions. Cambridge University Press.

Tellioğlu, Hilda, and Wagner Ina. 2001. Work practices surrounding pacs: the politics of space in hospitals. *Computer Supported Cooperative Work (CSCW)* 10: 163–188.

Traweek, Sharon. 1997. *Iconic devices: Toward an ethnography of physics images, 103–115*. Cyborgs and citadels: Anthropological Interventions in Emerging Sciences and Technologies.

Turrini, Mauro, and Catherine Bourgain. 2021. Appraising screening, making risk in/visible. the medical debate over non-rare thrombophilia (nrt) testing before prescribing the pill. *Sociology of Health & Illness* 43(7): 1627–1642.

van Leeuwen, Bart, Richard Gasaway, Gerke Spaling, and B.V. Netage. 2022. Adopting ai to support situational awareness in emergency response: a reflection by professionals. In: *Proceedings of the First International Conference on Hybrid Human-Artificial Intelligence. Frontiers in Artificial Intelligence and Applications*, IOS Press.

van Berkel, Niels, Mikael B. Skov, and Jesper Kjeldskov. 2021. Human-ai interaction: intermittent, continuous, and proactive. *Interactions* 28 (6): 67–71.

Van den Broek, Elmira, Anastasia Sergeeva, and Marleen Huysman. 2021. When the machine meets the expert: An ethnography of developing ai for hiring. *MIS Quarterly* 45(3).

Vodrahalli, Kailas, Tobias Gerstenberg, and James Zou. 2022. Uncalibrated models can improve human-ai collaboration. arXiv:2202.05983

Weiser, Mark, and John Seely Brown. 1996. Designing calm technology. *PowerGrid Journal* 1 (1): 75–85.

Wenger, Etienne. 2009. Communities of practice: The key to knowledge strategy. In: *Knowledge and communities*, 3–20. Routledge.

Winthereik, BritRoss, and Signe Vikkelsø. 2005. Ict and integrated care: some dilemmas of standardising inter-organisational communication. *CSCW* 14 (1): 43–67.

Zenisek, Jan, Florian Holzinger, and Michael Affenzeller. 2019. Machine learning based concept drift detection for predictive maintenance. *Computers & Industrial Engineering* 137 (106): 031.

Publisher's Note