Information Systems Journal

WILEY

RESEARCH ARTICLE OPEN ACCESS

How AI Helps to Compile Human Intelligence: An Empirical Study of Emerging Augmented Intelligence for Medical Image Scanning

Mechthild Pieper[1] | Rob Gleasure[2]

[1]Chair of General Management and Information Systems, University of Mannheim, Mannheim, Germany | [2]Department of Digitalization, Copenhagen Business School, Frederiksberg, Denmark

Correspondence: Rob Gleasure (rg.digi@cbs.dk)

Received: 1 September 2023 | **Revised:** 5 August 2024 | **Accepted:** 7 January 2025

Keywords: artificial intelligence | augmented intelligence | case study | emergence | future of work | healthcare

ABSTRACT

Artificial intelligence (AI) is advancing continuously. However, full delegation to an AI application is often not possible or desirable due to technical limitations, ethical concerns or legal issues. Augmented intelligence systems, where humans and AI work together jointly, have been proposed to improve decision making in complex, uncertain and failure-intolerant environments. Yet, this raises questions about how compatible human and AI knowledge are, and whether translating between the two increases decision making intelligence, or whether it effectively limits AI applications' capacity for computational agency and human agents' capacity to consider uniquely human knowledge. We explore this notion by looking at augmented intelligence in terms of systemic intelligence and mutual learning. Building on an emergence perspective, we perform a case study of an augmented intelligence system for image-based diagnostics in the radiology branch of a medical care centre. Our findings indicate a strong distinction between specialists' and non-specialists' intelligence augmentation with AI. This distinction fuels generative cycles which produce iteratively more sophisticated algorithms, human representations and practical routines. Drawing on this analysis, we propose three stages by which new forms of intelligence emerge from the addition of AI recommendation tools, specifically, intelligence by propagation, intelligence by specialisation and intelligence by articulation.

1 | Introduction

Recent artificial intelligence (AI) research has focused on the power of AI to support human agents' decision-making, rather than replace or automate it; a term referred to as 'augmented intelligence' (Jain et al. 2021). In augmented intelligence systems, an AI application provides human agents with a recommendation to help those human agents make decisions under uncertain and complex conditions (Bansal et al. 2019a). In principle, these systems allow human agents and the AI application to combine their relative strengths in a way that allows the AI application to harness its ability to process large amounts of data quickly, and human agents to consider issues that algorithms may miss, such as fairness (Teodorescu et al. 2021) and ethical or legal implications (Saunders et al. 2017; Fügener et al. 2021). These systems have been applied in fields such as medicine (Wang et al. 2016), crowdsourcing (Kamar and Horvitz 2012), and speech recognition (Gaur 2015). However, this process may be limited by two theoretical challenges.

First, augmented intelligence requires an interplay of mutual learning between the AI application and its users, with the result that algorithms come to replicate what human agents deem valuable (Van den Broek, Sergeeva, and Huysman Vrije 2021). This can be difficult because the human agents involved often possess varying backgrounds and knowledge (Kokkodis 2021), and the ability for AI to draw on 'appropriate' knowledge and produce appropriate intelligence requires some judgement about what constitutes appropriate knowledge (Samtani et al. 2023). Complicating matters further, it appears that once human agents overcome this challenge and converge on a shared understanding of what constitutes appropriate knowledge for an AI, they may become less capable of leveraging their unique relative strengths and diverse perspectives (Fügener et al. 2021). This may contribute to a larger loss of intelligence in augmented intelligence systems, as decision makers and groups become less likely to consider the types of 'uniquely human knowledge' that are not captured by their AI applications (in other words 'the knowledge a human has, but the AI does not') (Fügener et al. 2021, 1528). Thus, it seems that the act of integrating human intelligence with AI threatens to degrade that human intelligence.

Second, much of the espoused value of AI applications comes from their ability to solve problems in radically different ways from human agents (Verganti, Vendraminelli, and Iansiti 2020). This has been described as 'computational agency', and it distinguishes AI from other digital systems, 'Because the input–output relationships of such tools are unknowable, both ex ante and ex post, [meaning] designers never become fully knowledgeable with regards to the tool's behaviors' (Zhang, Yoo, et al. 2021, 1207). This implies that at least some of the value of AI applications arises because they can discover solutions in unpredictable, and often non-reproducible ways. However, for human agents to critique the outputs of an AI application and apply their judgement, 'users would have a need to understand the underlying algorithmic process and/or be offered explanations that accompany actions' (Jain et al. 2021, 678). This places restrictions on the types of algorithms and approaches that can be applied. Thus, it seems that the act of integrating human intelligence with AI also threatens to degrade the types of AI that can be used for intelligence augmentation.

Despite these challenges, augmented intelligence systems are already being implemented in organisations (Jain et al. 2021). These implementations assume that the interactions between humans and AI can enable new forms of intelligence, that is, intelligence which may not have manifested without the collaboration of humans with AI. Such new forms of intelligence may emerge in situations where the available human and artificial intelligence cannot independently address a problem, and so each is used to support different elements of problem-solving. There are multiple practical examples of augmented intelligence systems enabling new forms of intelligence. Jussupow et al. (2021) demonstrated that medical professionals can use AI recommendations to provide an alternative perspective which can trigger metacognition and self-reflection. Jia et al. (2024) provide another example, showing that knowledge workers can use AI tools as assistants to make additional intelligence available when engaging with customers, and so to stimulate new creative solutions. Such studies demonstrate that augmented intelligence can result in new forms of intelligence. However, they do not explain how these systems overcome the theoretical challenges outlined above, and how such systems can integrate human and artificial intelligence when each builds on diverging assumptions and capabilities.

This study explores the process by which such new forms of intelligence are discovered. Specifically, we ask: *how do new forms of intelligence emerge in augmented intelligence systems?* We argue that these systems can increase intelligence at a systemic level due to mutual learning, both between human agents and an AI application, and among human agents prompted by the enabling role of an AI application. We draw on existing literature on augmented intelligence, mutual learning, complexity and emergence. We perform a case study of an augmented intelligence system for image-based diagnostics in the radiology branch of a medical care centre. Our findings indicate a strong distinction in the ways that specialist and non-specialist interact with the AI. This distinction fuels generating mechanisms which produce iteratively more sophisticated algorithms, human representations, and practices. At the heart of this process, we propose three ways in which augmented intelligence emerges from the interplay of diverse human actors and the AI application, namely, intelligence by propagation, intelligence by specialisation and intelligence by articulation.

2 | Augmented Intelligence and Mutual Learning

2.1 | The Interplay of Artificial and Human Intelligence

The concept of AI dates back to Alan Turing's 'Can Machines think?' in 1950 (Turing 1950). It was first defined during the preparation of the Dartmouth summer workshop in 1956 as 'making a machine behave in ways that would be called intelligent if a human were so behaving' (McCarthy et al. 1955, 11). More recently, the concept of AI has extended beyond simply replicating human intelligence, with the emphasis on actually solving problems that humans could not. For example, Berente et al. (2021, 1435) define AI as 'the frontier of computational advancements that references human intelligence in addressing ever more complex decision making problems'.

What it means to 'reference human intelligence' can change from one context to another. In some instances, algorithms are created to mimic human neural systems to help them learn from past experiences (Krogh 2008). In other instances, human intelligence is required to continuously reconfigure algorithms (Sturm et al. 2021). This is sometimes referred to as the 'human-in-the-loop' (Fügener et al. 2021; Morse et al. 2021; Teodorescu et al. 2021). A large body of research has argued that this human-in-the-loop is indispensable for successfully developing and implementing AI applications (Abdelzaher et al. 2017; Shah et al. 2010; Stankovic 2015; Talburt and Zhou 2015; Tomaszewski 2021; Wagner 2016). Humans often provide the labels or annotation of training and test data, they decide which data should be used for the training, and they may also decide which features of the data should be incorporated (Hastie, Tibshirani, and Friedman 2009; Shao, Yuan, and Wang 2020). Additionally, in many application areas of AI, human agents must entrench a 'ground truth', which represents the consensus of correct labels (Lebovitz, Levina, and Lifshitz-Assa 2021). This need for human input into AI-based learning suggests that AI cannot operate independently from human agents.

The reliance between AI and human input is not only one-way. There has been a surge in personal and professional scenarios in which human intelligence also relies on AI (Brynjolfsson and McAfee 2017). Examples include AI-coaches for training sales agents (Luo et al. 2021), AI-enabled voice assistants and smart household appliances (Puntoni et al. 2021), AI-enabled auditing and advisory functions (Munoko, Brown-Liburd, and Vasarhelyi 2020), as AI for medical purposes, such as diagnosis recommendations or chatbots for medical advice (Abràmoff et al. 2018; Gulshan et al. 2016; Haenssle et al. 2018; Huston 2017; Longoni, Bonezzi, and Morewedge 2019; O'Hear 2017). Embracing the collaboration between human agents and AI, in which augmentation is favoured over automation, is therefore viewed as among the most promising near-term directions for AI (Jain et al. 2021; Jussupow et al. 2021; Raisch and Krakowski 2021; Teodorescu et al. 2021). Hence, multiple IS scholars have called for further research to consider the implications of augmented intelligence (Lebovitz 2019; Van den Broek, Sergeeva, and Huysman Vrije 2021; Wilson and Daugherty 2018).

2.2 | The Paradox of Mutual Learning in Augmented Intelligence

The collaborative relationship between human agents and AI tools may vary, based on the specific groups of human agents involved. Many AI applications seem to integrate specialised human domain knowledge to help address complex problems and support difficult tasks (Dwivedi et al. 2021). The goal of this integration is that, together, 'humans and AI systems jointly evolve in their mutual understanding of each other's strengths and weaknesses' (Jain et al. 2021, 681). This interplay between human agents and AI ultimately leads to a paradox where humans unavoidably learn from the AI (simply by using it in their workflow) while they are at the same time expected to correct it. There are consequently concerns that this will result in the loss of unique human knowledge, as human agents enter a self-supporting loop of collective incremental learning, which reinforces a narrow understanding of a problem, and which inhibits a team's ability to draw on their diverse backgrounds and to apply alternative perspectives (Fügener et al. 2021; Jussupow et al. 2021).

To mitigate this threat, researchers and practitioners advocate an active, explicit learning process and increased engagement in addition to the implicit ad-hoc learning that takes place organically through the use of AI (Fügener et al. 2021; Jussupow et al. 2021; Lebovitz, Lifshitz-Assaf, and Levina 2022; Zhang, Mehta, et al. 2021; Zhou and Chen 2019). This not only applies to the users of the systems, but also the AI developers who can intentionally or unintentionally steer domain users' learning due to the way they design recommendation algorithms, provide information, or configure the AI application (Bansal et al. 2019a; Fügener et al. 2021). This is one reason why 'explainable AI' (Rai 2020) has become popular; to turn opaque AI applications into comprehensible ones which provide transparency and explanations for the AI algorithm's behaviour. These explanation are expected to support users in integrating the AI advice into their decision making, facilitate trust, and provide better means for accountability (Doran, Schulz, and Besold 2017; Holzinger 2018; Rai 2020).

These trends towards explainable AI and augmented intelligence appear complementary. Augmented intelligence suggests that systems are more effective if human agents can critically consider the outputs of AI applications before making decisions (Lebovitz, Lifshitz-Assaf, and Levina 2022; Jussupow et al. 2021), while explainable AI suggests human agents need some way of understanding how AI is arriving at its outputs to reliably evaluate them. Thus, both research streams argue that human reflection and learning is the key to achieve augmented intelligence and to advance the performance of systems of augmented intelligence.

2.3 | The Potential of Human and Mutual Learning in Augmented Intelligence

Social learning is important for individuals seeking to develop specialised knowledge and skills (Bandura 1977). While formal instructions can help individuals accumulate standard behaviours and guiding principles, those individuals often rely on guidance and feedback from peers in their social surroundings if they wish to develop sophisticated practical skills and task-specific knowledge (Compeau and Higgins 1995; Warkentin, Johnston, and Shropshire 2011). This learning is often mutual in nature, even where some individuals are significantly more knowledgeable than others. Less knowledgeable and skilled individuals benefit from interaction with more knowledgeable and skilled individuals, as they are able to observe important behaviours that may be overlooked or difficult to explain (Kanfer and Ackerman 1989). Simultaneously, more knowledgeable and skilled individuals benefit from demonstrating and explaining behaviours to less knowledgeable and skilled individuals (Burgess, McGregor, and Mellis 2014). These benefits can be attributed to the pressure placed upon the more knowledgeable and skilled individual to externalise these behaviours, which often requires that they symbolise them and engage in higher psychological processes for reflection (Vygotsky 1980). The potential for mutual learning between individuals with specialised knowledge and others is one of the reasons that medical training often includes peer mentoring (Ten Cate and Durning 2007), and why systems development teams often pair programmers with high and low experience (Balijepally et al. 2009). Mutual learning is therefore an ongoing and adaptive social process which can allow a system to continuously adapt to changing circumstances, inside and outside the system.

A similar view of learning can be applied to the interaction among various user groups and AI applications. Given that AI represents the frontier of computational advancements, then by definition, AI applications are continuously evolving in search of more accurate responses to feedback from complex and uncertain application contexts (Raisch and Krakowski 2021). It is the human agents working with AI applications that are typically expected to coordinate feedback about instances of AI failure for reconfiguration and improvement of AI applications (Jain et al. 2021). This requires that the human agents develop insights into the AI's performance by building a mental model of its behavioural patterns to evaluate when the AI is correct or when it should be overruled (Bansal et al. 2019b). The result is a process of multidirectional adjustment and correction, which unfolds interdependently between the AI application and its human users. In this process, humans learn from the AI while simultaneously teaching it what they deem valuable for further improvements (Van den Broek, Sergeeva, and Huysman Vrije 2021). Thus, augmented intelligence appears to require a form of mutual learning among human agents and an AI application.

There are calls for augmented intelligence to embrace this mutual learning dynamic, especially when recommendations need to be evaluated in light of undesirable social or legal implications, such as algorithmic fairness, discrimination, accountability, responsibility and culpability (De Lima Salge and Berente 2017; Floridi et al. 2018). For example, there is evidence that AI can help to mitigate bias by focusing on measurable facts and replacing intuitive instincts of human decision making (Agrawal, Gans, and Goldfarb 2019; Van den Broek, Sergeeva, and Huysman Vrije 2021). On the other hand, there is also evidence that bias can become magnified in black-boxed AI algorithms when those systems are trusted to make decisions without human oversight (Faraj, Pachidi, and Sayegh 2018; Floridi et al. 2018). This suggests that, not only might augmented intelligence systems allow for more ethical decision making than standard AI systems; due to the capacity for mutual learning, they may broaden human awareness and provide more ethical decision making than systems which rely solely on human knowledge (Morse et al. 2021; Teodorescu et al. 2021). Building on the discussed literature, it appears the mutual learning dynamic of augmented intelligence systems deserves further theoretical attention.

3 | Augmented Intelligence as a Process of Emergence

Digital systems tend to evolve along fast-paced and uncertain trajectories (Munoko, Brown-Liburd, and Vasarhelyi 2020). This is because the underlying technologies do not just become more advanced, they both shape and are shaped-by their contexts of use (Orlikowski 2000). AI applications appear especially complex and unpredictable, given they include some of the most advanced computational approaches (Berente et al. 2021) and some of the most specialised human domain knowledge (Dwivedi et al. 2021). To help understand augmented intelligence systems, Benbya et al. (2020, 4) suggest that 'complexity theories such as emergence [...] offer an explanation of processes and outcomes, [in which] systems appear to constantly adapt and self-organise to create configurations that ensure compatibility with an ever-changing environment'. Building on this argument, we now explore the basic tenets of emergence, and how this perspective may be applied to augmented intelligence.

3.1 | Complexity and Emergence

Complexity is a characteristic of systems which are composed of semi-autonomous entities that interact with and influence each other in non-linear ways, while responding to external or internal tensions (Holland 1995; Waldrop 1992). Complex systems undergo constant change, as well as coevolution and adaptation of their parts, which may lead to unpredictable outcomes (Kauffman 1992; Lewin 1992). The concept of complexity builds on the idea that, despite all of this seeming instability, an order emerges through the interaction of parts of the system. The study of complexity aims to explain how such order emerges from self-organised, complex and non-linear combinations of interactions (Anderson 1999). As such, 'complexity theory, with its investigation into emergent phenomena, promises to provide both a methodology and a theoretical framework for studying something that is already playing a crucial function in our businesses and institutions' (Goldstein 1999, 69).

Depending on the system under investigation, different approaches can be taken to study and model complexity, including coevolutionary models (Allen and Varga 2006), complex adaptive systems (Nan 2011) and chaos theory (McBride 2005). This study applies an emergence perspective, which is defined in the context of augmented intelligence as 'a dynamic process of interactions among heterogenous agents [in this case humans and AI] that unfolds and evolves over time, resulting in various kinds of unexpected, novel, individual- and group-level configurations and/or broader social structures' (Benbya et al. 2020, 5). This means that emergence-based explanations of phenomena 'contain the claim that emergent phenomena are neither predictable from, deducible from, nor reducible to the parts alone' (Goldstein 1999, 57). The types of structures that emerge may take the form of either 'composition' or 'compilation' (Kozlowski and Klein 2000). Composition configurations describe the convergence of agents' perceptions and behaviour, that is, how they come to act as a single integrated system. Compilation configurations describe the divergence of perceptions and behaviours into distinct sub-systems, each contributing to larger system outcomes in different ways. Both types of configurations lead to novel outcomes, which are best understood as amalgamated unit-level phenomena that cannot be sufficiently understood by regarding only the sum of their constituent parts (Kozlowski and Klein 2000).

3.2 | Key Characteristics of Emergent Systems

Emergent systems can be described according to four key characteristics (Benbya et al. 2020). These four characteristics are summarised in Table 1, along with samples of supporting literature for each.

The first characteristics is the presence of *disequilibrium or beyond equilibrium situations*. 'Beyond equilibrium' refers to a situation in which a complex system had reached equilibrium, but that equilibrium has been disturbed due to changes in its environment. Complex systems thus 'move beyond their equilibria', as ever-changing contexts mean they are subject to continuous adaptive tensions, triggers and small events outside the norm (McKelvey 2001, 2004). These tensions, triggers and events force the system to adapt to maintain a sense of order among its various components (Benbya and McKelvey 2006a). For example, a hospital might acquire a new AI application which promises to help classify patients as high risk or low risk of contracting some disease, and the availability of this application might prompt that organisation to consider how it could change its structures and practices to integrate the new application. Augmented intelligence systems seem especially susceptible to disequilibrium situations, given the role of computational novelty (Berente et al. 2021). Human agents must therefore perpetually realign their way of working with their AI algorithms, anticipating and adjusting to new capabilities

TABLE 1 | Four characteristics of the process of emergence.

Characteristic	Explanation	Related literature
Disequilibrium or beyond equilibrium situations	Tensions, triggers and small events outside the norm allow for the amplification of random events. Tension in the form of external or internal pressures drive the system from one state to another.	(Goldstein 1999; McKelvey 2001, 2004; Benbya and McKelvey 2006a; Benbya et al. 2020)
Positive feedback and bursts of amplification	A process starting with (1) bottom-up dynamic interactions among lower-level entities which over time yield a phenomenon that manifests at higher, collective levels. (2) The emergent higher collective level influences the components' behaviours on the lower levels from which it simultaneously emerges. Positive feedback is the 'engine' of complex system adaptation.	(Choi, Dooley, and Rungtusanatham 2001; Benbya and McKelvey 2006b; Kozlowski et al. 2013; Kozlowski and Klein 2000; Dooley and Van de Ven 1999)
Self-organisation and coevolution	Creative, self-generated, adaptability seeking behaviour. Emergent phenomena are novel structures that confer the adaptability which allows for continuous change.	(Kauffman 1992; Holland 1995; Goldstein 1999; Chiles, Meyer, and Hench 2004)
Phase transitions and attractors	The system transitions from one phase into another due to its self-learning abilities. Potential outcome situation, called attractors, may manifest themselves with a burst of amplification so that new roles, structures or causal relationships emerge.	(Goldstein 1999; Pentland et al. 2012, 2020; Tanriverdi and Lim 2017)

and new challenges as they come to light. The resulting adaptive tension between humans and AI constitutes an important part of the dynamic interplay, as each enables and constrains the other (Jain et al. 2021).

The second characteristic is *positive feedback*. Positive feedback occurs when the changes in a system create changes in adjacent levels which amplify the associated conditions and interdependencies (see Benbya and McKelvey 2006b). Positive feedback therefore increases the pressure on specific components to interact in particular ways, and this pressure allows predictable and reusable patterns of interaction to form within a system, even if the precise nature of the emergent system is unpredictable (Dooley and Van de Ven 1999). For example, when the aforementioned hospital introduces the new AI application to help predict patients' risk of contracting some specific disease, many individuals may initially prefer to ignore or minimise use of the AI application, while others may interact with the AI application in unpredictable ways. However, once some individuals demonstrate positive results from using the AI predictions, then everyone is pressured to adopt similar practices and to find ways to leverage the AI application further. Thus, positive feedback is essential for the emergence of new higher level structures in complex systems, as it acts upon both systemic and task-specific outcomes, and so provides a means for top-down contextual forces to shape and influence lower level phenomena and interactions (Choi, Dooley, and Rungtusanatham 2001).

The third characteristic is *self-organisation*. Self-organisation describes 'the emergence of system-level order as an unintended consequence of the action and repeated interaction of lower level system components, without intervention by a central controller' (Chiles, Meyer, and Hench 2004, 502). Importantly, self-organisation allows components to organise themselves into higher level phenomena in ways that would not be possible via top-down direction alone (Kozlowski et al. 2013). It also allows systems to find novel configurations capable of addressing local complexity, without having to balance the particularities of other parts of the system. Returning to our hospital example, the need for self-organisation means that, as the doctors and AI application co-evolve, it is up to those doctors and AI applications to direct future adaptations, as the configurations that suit each human agent may not generalise to others in the system. For example, some doctors may work with elderly patients and others may work with children, and this difference could change how accurate the AI predictions are, or how to act upon those predictions. Thus, augmented intelligence systems demand that each human and digital component can create new ways to navigate the unique complexities they both suffer and create (Jain et al. 2021).

The fourth characteristic is *phase transitions*. Phase transitions occur when a complex system has changed from one stable equilibrium to another (Benbya and McKelvey 2006a). This concept borrows from studies of physical systems, where phase transition describes matter passing between 'states', for example, a liquid to a gas (Polese et al. 2021). In social contexts, phase transition occurs when the incremental process of component-level adaptation reaches a tipping point, triggering significant changes in the system's structure which create new structures and logics to replace older versions (Pentland et al. 2020; Tanriverdi and Lim 2017). At this point, the instability in the system reduces, and the complexity in the system may begin to slow change, rather than accelerate it. For augmented intelligence, such phase transitions may occur as systemic tensions cause human agents and/or the AI application to form new associations over repeated interactions until some stable configuration emerges. In our hospital example, when doctors begin using the new AI application they might continuously scrutinise

AI predictions based on their own domain knowledge. If the AI predictions and their own interpretations routinely converge, then those doctors may stop these scrutinising practices. At this point, the interdependencies among components begin to limit component-level adaptation so that a stable new equilibrium manifests in form of a composition or a compilation. Over time this stable equilibrium may then be challenged by new tension, triggers or events outside the norm, creating the next beyond equilibrium situation.

4 | Methodology

This study applied a qualitative research methodology to explore how augmented intelligence can emerge around AI applications. Specifically, we performed an interpretivist case-study (Klein and Myers 1999) that used the core concepts identified in the emergence literature as a sensitising lens. We chose to focus on a single case, rather than a comparative case study for two main reasons. First, while we built on existing studies of AI and augmented intelligence systems, we also sought new explanations that could address some of the challenges in the existing literature. A single case encourages the researchers to stay close to their empirical observations and to allow those observations to dictate the direction of theorising, so limiting the impact of pre-existing assumptions and biases (Flyvbjerg 2006; Ragin 1992). Second, our focus was on understanding complex patterns of emerging intelligence within the system of augmented intelligence. A single case encourages the researchers to engage with the full complexity of the system, while a comparative case often encourages researchers to look for more abstract patterns that could be easily compared with other contexts (Patton 2005).

To obtain rich data on the emergence of augmented intelligence, we needed to observe a context where it had been implemented previously and was actively in use before and during the time of study. Thus, as with much existing AI research (Davison and Martinsons 2015; Ghassemi, Oakden-Rayner, and Beam 2021; Lauritsen et al. 2020; Lebovitz, Lifshitz-Assaf, and Levina 2022), we chose to focus on a medical context. The interpretivist approach allowed us to synthetize the shared narratives of actors that develop, market, work with, and refine systems of augmented intelligence (Klein and Myers 1999). We also chose a case where we could talk to users as well as developers of the AI application, as this helped us to gain a holistic understanding of the case from multiple perspectives.

Overall, the approach drew heavily on Pentland (1999), who laid out how to build a process theory from narratives. This approach advocates that researchers focus on the 'stories' told by different participants, where 'stories' represent a version of events as told from a particular point of view. Such stories are not necessarily complete in isolation; rather, they capture a coherent and memorable embodiment of past experiences. Each story is assembled from different 'tellings' or 'texts' presented from a particular narrator. These texts can be performed in a variety of ways, including speech, written documents, recordings or other forms of documentation. Researchers can then synthesise these stories into a 'fabula', that is, a generic description of events and their relationships. This fabula synthesises the various stories into a deep structure that transcends the experience of any one individual. Finally, with the fabula in place, researchers can theorise the underlying structures that enable and constrain the process, or the 'generating mechanisms'. These generating mechanisms present the tensions that drive change in the system.

4.1 | Case Background

The selected case context surrounds an AI application, called AI-Rad Companion, which was developed by Siemens Healthineers for providing multi-modality imaging decision support in the field of radiology. This AI application had been developed in co-operation with a medical care centre in North-Rhein-Westfalia, Germany. The project also includes members of the radiology department of a university hospital in Bavaria, Germany.

We selected this context for two main reasons. First, while the AI application is continuously evolving, it is past the experimental stage, which could otherwise have made it challenging to separate characteristics of the system from issues with the initial implementation. The product's usefulness had been investigated and tested in multiple studies and is approved by the European Medicines Agency (EMA) with a CE (Conformité Européenne) mark. We focused on the 'AI-Rad Companion Chest X-ray', which as 'a member of the AI-Rad Companion family [...] automatically processes upright chest X-ray images [and] next to pneumothorax, pleural effusion and nodule detection [...] is able to indicate consolidation and atelectasis' (*AI-Rad Companion*, siemens-healthineers.com). Hence the algorithm can detect the five most common findings spotted in chest X-rays (Rudolph et al. 2021, 2022). When detected, patients often undergo additional diagnostic investigation to (dis)confirm the X-ray finding (e.g., CT scan for nodule detection). The AI application is constructed with multiple algorithms arranged within a neural network, and this network is trained on manually annotated training data. Thus, the AI application falls into the category of supervised deep learning. It can be tailored to new data and contextual nuances, and it is reconfigured twice a year to integrate new data and user feedback. The second reason for selecting this context is that the chosen medical care centre is the first private practice to implement and test the AI application on a large scale by integrating it into the daily routine of doctors. The centre is comprised of 17 sites, many of which are connected to affiliated hospitals to which they offer radiology services. This has created sufficient opportunity for complexity to develop and augmented intelligence to emerge. Further, all relevant stakeholder groups were familiar with the AI application and could be interrogated for our study; this included different AI users, as well as a development team seeking to further advance the technology.

We identified four categories of stakeholders to interview in our case context. First, we identified 'specialists', which included qualified radiologists with experience in advanced diagnosis. Second, we identified 'specialists-in-training', which included individuals who were training to become radiologists but were not yet fully qualified. Third, we identified 'non-specialists', which included internists, surgeons, and other clinicians who rely on services from radiology to assist their daily work and diagnoses. Non-specialists primarily use the AI application when radiology services are not easily available, such as during night or weekend shifts, or during normal working hours characterised by high workloads and long waiting times. Fourth, we identified a 'development team'. Since the AI application's modalities require the determination of a 'ground truth' for training and test data, members of the development team evaluated AI-processed images ex-post for research purposes and product improvements. The development team, thus, included trained radiologists who used their expertise to contribute to product enhancements, rather than using the AI application in their medical routines. It also included product managers and a data scientist, who were still collaborating closely with the medical care centre at the time of study. We interviewed members from all four categories of stakeholders in an effort to integrate multiple perspectives as we built our process theory, as 'there is a great deal of insight to be gained from a careful analysis of the same story from multiple, subjective points of view' (Pentland 1999, 714).

4.2 | Data Collection

We sought to identify multiple stories told by individuals with specific points of view, so, consistent with Chakraborty, Sarker, and Sarker (2010), we began our case exploration with two initial talks with the development team where we focused on scoping the analysis of the system. These talks helped to delineate some of the different groups, as well as verifying the fit of the case site, and improving our contextual understanding. We designed an interview protocol to encourage participants from each of these groups (doctors as users of the AI application and varying members of the development team) to describe their experiences with the system. Our exploration of the case included a strong deductive element, as we used existing emergence concepts to guide the construction of our interview protocol. Specifically, the interview protocol was informed by the concepts of disequilibria, positive feedback, self-organisation and phase transitions, which we expected to stimulate participants' communication of their stories in a way that could be related to the principles of emergence. Due to the variety of interviewees, the protocol was slightly adapted for each group and the semi-structured interview guide allowed for follow-up questions during the interviews.

The first semi-structured interviews served to refine the interview questions, while contributing to a better general understanding. We adopted a 'theoretical sampling' approach in which we used the insights from ongoing analysis to direct further data gathering (Urquhart, Lehmann, and Myers 2010). This meant we often relied on a snow-balling approach to reach out to additional interviewees by asking each interviewed person which of their colleagues might be able to tell us more about specific contexts and practices. The main suitability criteria were that interviewees were familiar with the AI application, regularly engaged in intelligence augmentation with the AI-Rad Companion, and were willing to openly share and elaborate upon their experience with the AI application. Similarly, interviewees from the development team were engaged in the reconfiguration and improvement of the AI application, either by contributing to establishing the ground truth of new training data or by enhancing the technicalities of the AI application and its implementation. We were especially interested in exploring different 'stories' and sampling continued until we reached theoretical saturation, that is, additional interviews were neither identifying new stories, new stages, nor new theoretical nuances that required the stories and stages to be characterised differently. This required that data gathering and analysis was done iteratively throughout the process to integrate insights and allow the data to 'talk back' (Flyvbjerg 2006).

Theoretical saturation was reached after 16 interviews, at which point the open style of questions was no longer identifying new stories or stages that had not already been identified in previous interviews. Specifically, two distinct stories with six stages each had emerged: one for specialists/specialists-in-training and another for non-specialists. Each story could be told using six comparable stages. We interpreted the identification of these comparable stages as a sign that it was time to focus on theorising a generic description of events and relationships (fabula), and the underlying structures that enabled and constrained the process (generating mechanisms). We thus increased our efforts to relate our findings with existing literature on complexity, emergence, mutual learning and augmented intelligence. Simultaneously, we performed further eight interviews to challenge our interpretation and look for areas of ambiguity or misunderstanding. The interviews lasted approximately 40 min on average, with a longest duration of 62 min and a shortest duration of 21 min. One interview took place in person, was audio-recorded, and transcribed directly after. All remaining interviews were conducted via video call, fully recorded with interviewees' consent and transcribed within a maximum of 24 h after the interview. Table 2 and Figure 1 provide an overview of interviews.

4.3 | Data Analysis

For the analysis we adopted the narrative analysis approach proposed by Pentland (1999). This approach meant we theorised at the level of stories, the level of the fabula and the level of generating mechanisms. Building on Pentland (1999), we expected that the stories would be most explicit in the data, when compared with the more latent fabula and generating mechanisms. We accommodated this in our analysis in two ways. First, while we continuously referred to both data and literature when analysing each level, we tended to refer more often to the data when theorising the stories and more often to literature when theorising the fabula and generating mechanisms. Second, while we theorised each level continuously throughout the study, we focused our early analysis more on the stories than the fabula and generating mechanisms. This approach ensured we grounded our analysis firmly in the data and allowed us to stay responsive to unexpected observations. Figure 2 illustrates this approach.

Coding began with a thorough review of all transcripts while simultaneously listening to the recordings, highlighting important passages, noting down potentially relevant comments and thoughts, revisiting study notes, and browsing external media sources. This amounted to the type of 'line by line'

TABLE 2 | Overview of interviews.

Count	Category	Role	AI experience[a]	Duration	Affiliation
1	Development team (DT)	Medical CEO	Low	38 min	Medical Care Centre
2				45 min	
3		Product Manager 1	(High)	39 min	Provider
4		Product Manager 1	(High)	42 min	
5		Product Manager 2	(High)	43 min	
6		Data Scientist	(High)	62 min	
7		Data Scientist	(High)	54 min	
8		Medical Specialist	High	23 min	University Hospital
9	Specialists (S)	Project Lead	High	56 min	Medical Care Centre
10		Specialist 1		52 min	
11		Specialist 2	High	48 min	University Hospital
12		Specialist 3	High	21 min	Medical Care Centre
13		Specialist 4	Medium	24 min	
14		Specialist 5	Low	41 min	
15	Specialists in training (SIT)	Specialist-In-Training 1	Medium	40 min	Medical Care Centre
16		Specialist-In-Training 2	Medium	28 min	
17		Specialist-In-Training 3	High	41 min	
18		Specialist-In-Training 4	Medium	35 min	
19	Non specialists (NS)	Internist 1	High	36 min	Affiliated Hospitals
20		Surgeon 1	Medium	28 min	
21		Internist 2	Medium	22 min	
22		Internist 3	High	37 min	
23		Internist 4	High	41 min	
24		Internist 5	High	40 min	

[a]AI experience refers to interviewees level of experience with the AI Rad Companion depending how intensively or how long they have been accustomed to using the tool. Low: little regular use and/or very short use duration of up to 3 months; Medium: little regular use or short use duration of up to 6 months; High: consistent regular use and/or long use duration of up to a year.

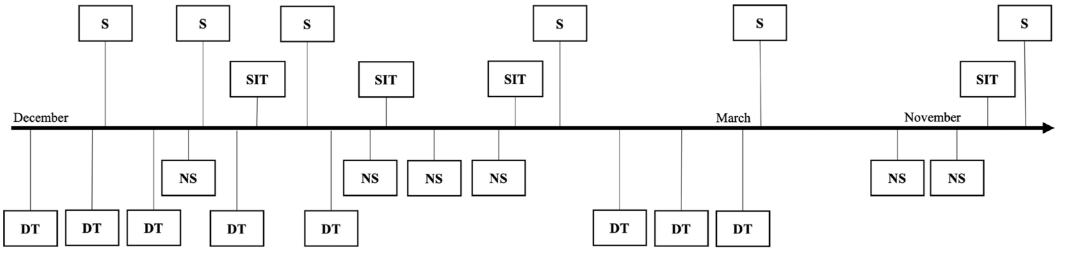

FIGURE 1 | Interview timeline.

exhaustive exploration described by Charmaz (2000). We observed two distinct 'stories' from the different 'tellings' of all narrators, one of which was *the specialists' story of emerging augmented intelligence* and the other was *the non-specialists' story of emerging augmented intelligence*. For this distinction, the specialists' story was primarily informed by interviews with specialists/specialists-in-training, while the non-specialists' story was informed by interviews with non-specialists, as well as specialists/specialists-in-training who worked together closely with non-specialists. We considered whether the development team also told a distinct, third story. However, such a third story was not evident in the data we gathered. Instead, members of the development team alternated between the specialists' story and the non-specialists' story, filling in gaps in each and contributing to a larger understanding of the interactions of all user groups with the AI

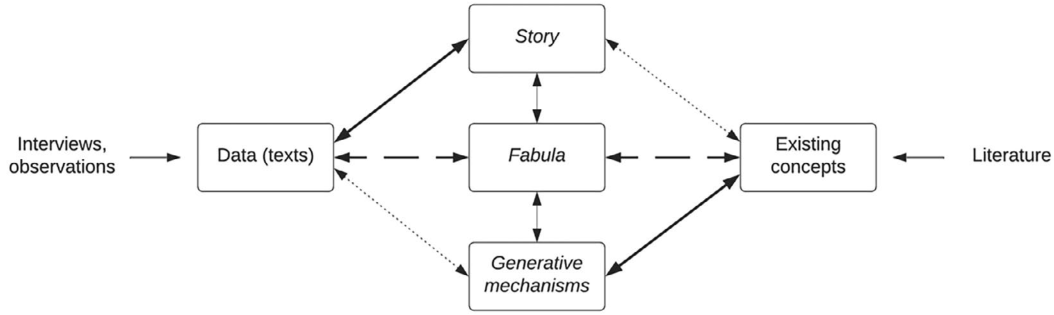

FIGURE 2 | The integration of data and literature during theorising.

application, including how the AI application was adapted in response to users' changing expectations over time.

To make sense of these stories, we considered preliminary open codes (see Weick 2007) which were continuously related to the four stages of emergence identified from literature, that is, 'disequilibrium or beyond equilibrium situations', 'positive feedback and bursts of amplification', 'self-organisation and coevolution', and 'phase transitions and attractors'. As the analysis progressed, we discussed specific observations that both supported and challenged these four stages of emergence identified from literature, with a particular focus on data that did not fit the coding scheme. We found the need to decompose 'disequilibrium or beyond equilibrium situations' into *externally generated equilibrium* and *new emerging tensions*. This need to differentiate was because externally generated equilibrium referred to generative forces outside the system, while new emerging tensions described changes among human actors and AI applications within the system which stimulated adaptive pressure. We also found the need to separate 'self-organisation and coevolution' into *human self-organisation* and *AI/developer self-organisation*. Human self-organisation referred to the ability of specialists' and non-specialists' to creatively adapt new structures for continuous change. AI/developer self-organisation referred to the ability of AI applications, under the control and guidance of the AI developers, to creatively adapt new algorithms, data, structures, modalities, and interaction norms to enable continuous change. Each stage is described in more detail later in the study.

As our understanding of the distinct stories began to crystalise, this allowed us to dig deeper into the *fabula* that linked different stories as part of one larger narrative (Pentland 1999), and to relate observations to the generating mechanisms for change described by Van de Ven and Poole (1995). This process can be compared with the scaling up from 'substantive theory', which provides an abstract account of observed phenomena, into 'formal theory', which builds on the substantive account by comparing it with other theories to search for deeper conceptual mechanisms and relationships (Urquhart, Lehmann, and Myers 2010). This scaling up process resulted in a fabula that included three common underlying stages of mutual learning at the heart of the emergence process. The first stage was *learning by propagation*. We identified this stage by observing similarities between the first two stages in the parallel stories and existing accounts of mutual learning in organisations, which showed how knowledge is shared through demonstration. The second stage was *learning by specialisation*. We identified this stage by contrasting the third and fourth stages in the parallel stories with existing literature which showed how self-organising systems allow actors to be become more focused in their roles and responsibilities. The third and final stage in the fabula was *learning by articulation*. We identified this stage by contrasting the fifth and sixth stages in the parallel stories with existing literature which showed how actors can improve their own understanding by making their knowledge explicit to others.

For the generating mechanisms, throughout the study the research team continuously engaged in an iterative and collaborative theorization process. We realised that the generating mechanisms appeared to reflect the type of 'dialectic' generating mechanism that is described by Van de Ven and Poole (1995), that is, 'conflicts emerge between entities espousing opposing thesis and antithesis that collide to produce a synthesis, which in time becomes the thesis for the next cycle of a dialectical progression' (Van de Ven and Poole 1995, 520–521). In our study context, this dialectic emerged from the systemic need to adapt to competing forces and values within and among groups of specialists and non-specialists. Specifically, we found two generating mechanisms related to competing desires: the competing desire for standardised practices and unique human knowledge and the competing desire for computational agency and explainability.

As the analysis matured, this process of zooming in and out of the different levels of the theory, and zooming in and out of data and relevant literature became increasingly interwoven and iterative. We eventually settled on a stable interpretation which appeared to be both empirically and logically consistent (Eisenhardt 1989). We challenged this selected interpretation to look for issues related to theoretical validity, interpretive validity, descriptive validity, evaluative validity and transferability (see Appendix C for more detail of the steps taken). We further continued to challenge this interpretation after this point, using the data and literature to propose alternative stages in both the stories (e.g., training or testing) and the fabula (e.g., intelligence by reflection), to challenge the distinction between user groups (e.g., whether the development team told a distinct story), and to link the process with alternative generating mechanisms for change (e.g., power dynamics or the embedded biases in the assumed 'ground truth' for diagnoses). None of these other theorizations were able to explain the findings as well as the selected interpretation, so we concluded the analysis. Figure 3 illustrates the coding procedure and provides selected open codes from the analysis.

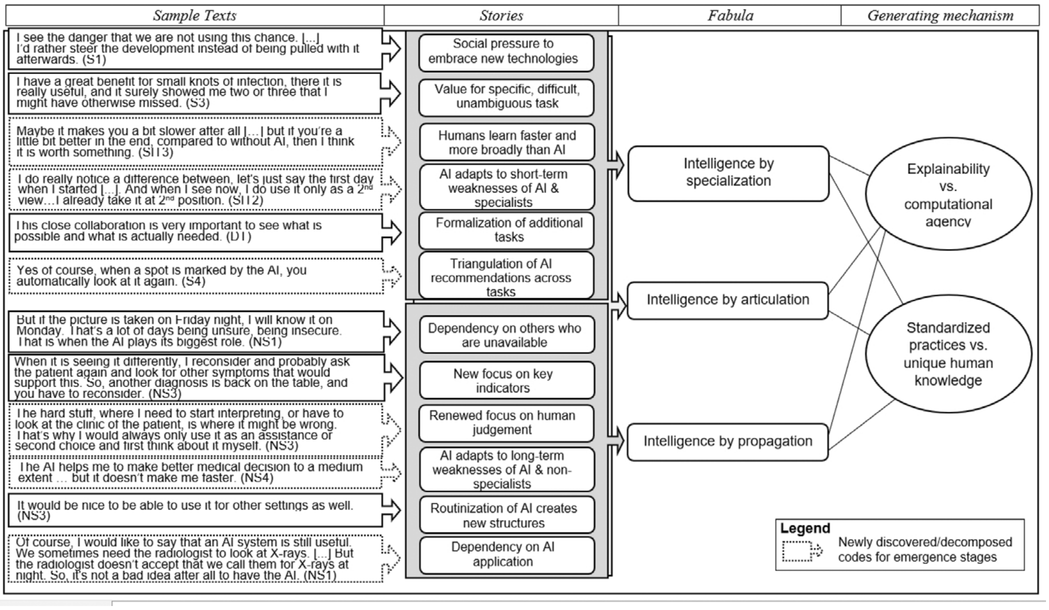

FIGURE 3 | Illustration of coding.

5 | Findings

The distinction between specialists' and non-specialists' interaction with the AI was evident in the way human agents described themselves and others, and in the different stories they told. These differences mainly arose from the diverging application contexts in which the two groups operate, as well as each group's diverging perceptions of how the AI could create value for them. Despite these differences, there were also many areas of convergence. First, participants agreed in principle that the recommendations from the AI application were less trustworthy than those from specialists. This elevation of human judgement informed many of the other dynamics, as the AI application was configured to fill in the gaps left by human agents, rather than the other way around. Second, each of the stories shared a similar set of six stages that characterised the process of emerging augmented intelligence, specifically (i) externally generated disequilibrium (ii) positive feedback, (iii) human self-organisation (iv) AI/developer self-organisation (v) phase transition and (vi) new emerging tensions. Despite the consistency of these key stages in the stories told by specialists and non-specialists, each described the stages differently, based on the forces they experienced in their roles.

5.1 | Externally Generated Disequilibrium

Each story began with a 'beyond equilibrium' stage that resembled those described in general emergence literature. Notably, however, the disequilibrium in this context was generated by external, social influences. For specialists, they believed that they were expected to be curious and forward-looking, and so, they should embrace and leverage the potential of new technologies. For non-specialists, they often found themselves reliant for others, particularly specialists, and this meant they were limited in the speed and quality of healthcare they could provide patients. These non-specialists initially saw the AI application as a way to limit their dependency on others who were often unavailable when needed.

5.1.1 | Specialists Experience Social Pressure to Embrace New Technologies

The motivation to develop an augmented intelligence system was driven by specialists' discussion of new technologies with hospital administrators and peers in other institutions, and the desire to stay ahead of potentially disruptive trends linked to those technologies. Specialist 1, the specialist who initiated the collaboration, explained that he 'would rather steer the development instead of being pulled with it afterwards'. Several other specialists echoed this general sense of curiosity, admitting they were looking at the advice provided by the AI application to learn about what it can do, rather than because it solves an existing problem for them. Project Manager 1 explained that this type of curiosity is essential, as it means they are collectively 'open for new technologies, for new possibilities and clinical pathways'. External actors also contributed to this curiosity around new technologies, many of whom assumed the use of these new technologies was an indicator of doctors' commitment to improving practices and outcomes. Specialist-In-Training 3 summarised this sentiment,

> 'Even if for the doctors it is only a little more certainty, I could imagine that patients feel much safer, because maybe you attribute flawlessness to the AI rather than to humans. Then the patients feel safer, because their diagnosis was double-checked by an AI system. These psychological factors can be helpful [...] and very good advertisement'. Ultimately this can lead to the fact 'that patients say they rather go to this medical practice for screening, because it offers an additional, independent computer program that checks the images. This then leads to an economic benefit as well'.

These forces combined to push specialists to embrace new AI technologies, even if it was not yet clear how they could increase intelligence and add practical value.

5.1.2 | Non-Specialists Become Aware of a Dependency on Others Who Are Unavailable

The motivation to introduce augmented intelligence was not only driven by the availability of new technologies, but also by growing political and clinical demands for more frequent scanning and diagnosis. Product Manager 2 in the development team explained that 'the workload of radiologists gets higher and higher every year, because there is more and more imaging, and imaging is getting more and more important' and that 'in many clinical situations it is reality that X-ray images are not always reviewed in time by radiologists'. The practicalities of this trend meant that non-specialists are increasingly asked to step in and scan images, even though it may not be their domain of expertise. Specialist 2 explained that,

> 'In real life it is not only the board-certified radiologists with 20 years of experience who looks at the images and decides on the therapy from the image, there are also the very young resident radiologists or even, the surgeons, internal medicine doctors, and residents in the emergency department who see or have to detect pneumonia. For those people who don't have the experience, the algorithms that are close to or as good as the board-certified radiologist, have a great benefit'.

This created an opening for the AI tool to be used by non-specialists, as external circumstances meant it could add value without being able to replicate the full value of a consultation with a specialist.

5.2 | Positive Feedback

Consistent with previous research on emergence, positive feedback played an important role in the emergence of new forms of intelligence. While both specialists' and non-specialists were working with the same AI application, they often responded to different forms of positive feedback around the system. Specialists saw the role of the AI application as relatively narrow, so they responded positively when those capabilities could be leveraged for some highly specific task, such as detecting pulmonary nodules. Non-specialists saw the role of the AI application as potentially broader, so they responded positively when the AI application could provide some guidance on a wider variety of tasks, even if the quality of guidance was not as high as they would receive from specialists.

5.2.1 | For Specialists, the AI Provides Value for a Specific, Difficult, but Unambiguous Task

Many specialists found the system most useful for a set of difficult, but routine tasks. Specialist-In-Training 3 explained that 'it does help [in specific situations, for example] when doing the 50th diagnosis of the day [...], there might be some small nodules that I might have otherwise missed'. Specialists repeatedly reported that the AI application was particularly good at detecting pulmonary nodules. Specialist 4 explained this support as 'another focus that I might not have noticed before, then I look at it again and see ... ok is there something there now or not?' Usually, radiologists scrutinised specific areas in X-ray images that appear suspicious, thus if the AI highlighted a specific area with a recommended diagnosis, radiologists could follow known decision-making steps to evaluate the AI advice, that is, interrogate the AI application according to a human-logic guided procedure. This allowed the AI application to triangulate bias, simply by causing human agents to revisit some factors that may otherwise not have received extensive attention.

The accumulation of positive experiences when using the AI for specific tasks convinced several specialists that it is worth checking what the AI recommends. Specialist-In-Training 1 and 2 remarked that looking at the AI after writing their own diagnosis created a sense of confirmation which 'actually makes [them] more secure in [their] work'.

5.2.2 | Non-Specialists Find a New Focus on Key Indicators by Using the AI

At first, the AI supported non-specialists for quick decisions and routine tasks by directing their attention towards possible causes of concern, or alternatively reassuring them that their diagnosis was correct. This was also how specialists anticipated that non-specialists would use the system, as it could offer 'a great benefit for the non-radiologist, who can much more easily say that there is nothing. If the AI is not showing him anything; then he can send the patient home with a much better feeling' (Specialist 3). Viewed in this way, the role of the AI application was to help determine whether to escalate a diagnosis to specialists.

In practice, however, the effect of the AI recommendations on non-specialists was more transformative. Much of the reported positive feedback arose because the AI recommendations challenged non-specialists to think through their diagnosis more carefully. Internist 1 described this impact, explaining

> 'It does not necessarily make me more sure, but it makes me more awake or alert. [Also], it adds some stress, because it makes you question yourself, but I don't think that's a bad thing, because it makes you look a little closer'.

This meant the AI application became a useful learning tool for two main reasons. First, it prompted non-specialists to pay attention to seemingly simple diagnoses, as well as to assume greater responsibility for those diagnoses. Second, because the AI application was designed to emulate the diagnostic reasoning of a specialist, it highlighted the information that those specialists found most important, and why it was relevant.

This allowed the AI application to act as a form of 'teacher', propagating general diagnostic knowledge from specialists to non-specialists via the information priorities embedded in the AI recommendations.

5.3 | Human Self-Organisation

Most existing accounts of self-organisation do not differentiate between human and technical actors. However, this case context suggested that human self-organisation played a distinct and important role. Once the AI application began to demonstrate value, both specialists and non-specialists began to adapt to take advantage of the new system. The way they adapted varied, depending on how they perceived the strengths and weaknesses of the AI application. Specialists recognised that, although the system could perform clearly defined tasks quickly and reliably, the system did not learn quickly. Specialists therefore began to refocus their role within the system to concentrate on the new and unusual aspects of diagnoses. Non-specialists recognised that, although the AI application allowed them an effective way to analyse the clearly defined aspects of a diagnosis, the system missed many other contextual details that were important. Non-specialists therefore began to refocus their role on contextualization and the 'human side' of treatment.

5.3.1 | Specialists Learn Faster and More Broadly Than the AI

Over time, the perceived limitations of the AI became a way for specialists to conceptualise how and why some cases were unusual and worthy of more nuanced analysis. Specialist 3 explained that 'surely the more severely sick patients, those where multiple diseases occur together [...] there the AI has weaknesses, it simply shows many things and doesn't relate them to the context'. The specialists learned to treat AI recommendations with caution and started to crystalise the situations in which they should refer to the AI application. This was further facilitated during active revision processes which encouraged the specialists to consider the clinical context of a diagnosis, and to compare current diagnoses with previous cases. Thus, specialists and specialists-in-training were able to learn quicker than the AI could, often identifying recurring problems in the AI that could be reported to the development team. This practical reality (that human learning quickly outpaces AI learning) was widely accepted among interviewees, including the members of the development team themselves. While the AI needs to be told which information to consider, users can decide for every case which information is useful for the evaluation.

More broadly, much of the human self-organisation occurred because, while some of the more basic tasks are relatively rule-based, the more advanced learning tends to require more creativity and dynamic problem solving. Product Manager 1 summarised this as 'if you are already at a very high level, it is difficult to get even better, but if you're at a lower level, you have much more potential to improve'. As a result, many specialists found they used the system differently once they gained more experience, as they found the value of confirming their recommendations less reassuring and the ability to challenge their recommendations less compelling. These pressures forced the development team to expand the system in order to maintain those specialists' engagement.

5.3.2 | Non-Specialists Renew Their Focus on Human Judgement

While specialists were often trained to look closely for a specific range of illnesses, non-specialists were often more sensitive to the broader context of a patient. Non-specialists felt they were obliged to consider a wider range of explanations, albeit with significantly less precision. Further, non-specialists also felt it was not unusual to receive conflicting recommendations from multiple sources. For these reasons, non-specialists often perceived imprecisions in the AI application to be less severe than specialists. The Data Scientist confirmed that non-specialists were more tolerant of mistakes and inconsistencies, noting 'often radiologists don't agree with each other'. When this happens, it is up to the various human agents involved to adapt to these conflicting recommendations and find a way to balance various perspectives. Internist 1 explained when a non-specialist reads a written report, it was common for this to prompt a longer informal conversation where 'you can call the radiologist and with the radiologist you can discuss the state of the patient better'. Such discussions were facilitated during daily meetings, in which ambiguous cases were presented and reviewed to agree on a treatment approach and to facilitate social learning among specialists and non-specialists.

These norms made it more straightforward to accept the limitations of the AI for non-specialists, as they were accustomed to this process of resolving conflicting or incomplete recommendations. AI recommendations became integrated into the daily discussions, and this allowed them to be challenged by specialists if necessary. Thus, the addition of AI recommendations created more pressure on non-specialists to also apply human judgement. Often, this judgement took the form of requesting additional diagnostics, such as blood testing or more advanced imaging (e.g., computed tomography (CT) scans). Surgeon 1 explained

> 'AI just tells me, you have to look at it again, maybe you didn't see something. [...] Just look at it, is it the lung? It's not the lung, then more diagnostics, and that's very good actually, that you have to look twice'.

5.4 | AI/Developer Self-Organisation

The case context also identified a distinct AI/developer self-organisation stage. The AI application had limited computational agency to change how it operated, for example, the data it used, the parameters it considered, and the techniques it applied. Instead, most changes in the AI application were initiated and carefully controlled by the development team. However, the AI developers were nonetheless constrained by the capabilities of the AI application, the available data, and which algorithms were feasible. Similar to the human self-organisation stage, AI/developer self-organisation was responsive to tensions in the system. For example, just as the specialists and non-specialists self-organised around the perceived strengths and weaknesses of the AI application, so the AI application (under the guidance and supervision of its developers) self-organised around the perceived strengths and weaknesses of the specialists and non-specialists. In particular, AI/developer self-organisation focused on the tendency for specialists to become fatigued and to make 'human errors', and the tendency for non-specialists to seek multiple, potentially contradictory perspectives when making decisions. What distinguished AI/developer self-organisation from the human self-organisation stage was that, while the latter generally reacted to and resolved tensions and instabilities, AI/developer self-organisation included a continuous reconfiguration of internal elements which actively created tensions and instability.

5.4.1 | AI Adapts to the Short-Term Limitations of Both Itself and Specialists

Early adaptions to the AI application focused on general improvements in the recommendations. This required ongoing dialogue between specialists and the development team to explicate the reasons when the AI application was underperforming. One prominent example was the use case of pneumothorax detection; one of the most critical and well-known lung X-ray diagnoses performed by doctors. In an earlier version, the AI application provided surprisingly inaccurate recommendations for this (from a human perspective) relatively easy task. After extensive communication and collaboration, the development team discovered that the AI application was misclassifying the tubes used to treat a pneumothorax as physiological features. The algorithms were subsequently adjusted, the number of false positives was reduced, and the accuracy improved.

Data Scientist 1 explained that the large amount of training data and the black-box nature of deep learning algorithms made it difficult to discern why seemingly reliable and well-tested computational processes produced unreliable outputs in certain situations. As a result, the system had to be troubleshooted as it was deployed, and this demanded close dialogue with specialists. This dialogue allowed specific cases to be reannotated, labels to be updated, algorithms to be improved, and new training data to be added. The AI application could then autonomously adjust the filters of its convolutional layers, until the system produced more reliable recommendations. This was an attritional process, which the Medical Specialist described as 'very tedious' and based on 'a lot of trial and error'. However, it also provided a rich exchange between specialists and AI developers, as each came to understand in more detail how recommendations could be derived.

Critically, over time, this dialogue also revealed weaknesses in the ways that specialists performed diagnoses. They observed that specialists are subject to fatigue and diminishing attentiveness in ways that the AI application was not. Specialists also acknowledged this, with Specialist 4 explaining that 'for the complex diagnosis it is less of a benefit for the radiologists, but with combinations of simple tedious tasks it can be very helpful, if it provides intermediary steps'. This focus on fatigue meant the challenge for the AI changed from replicating human decision-making to folding in around the weaker points of human decision-making, without interfering with the human dynamism required for complex, dynamic problem solving. For example, the development team responded by including explainability features, such as highlighting visualisations and a confidence score, with the goal of helping human agents extract helpful embedded information from the AI application. This approach also alleviated some of the concerns around relying on the AI application, as it ensured that the specialists, as human agents, remained at the centre of the diagnosis. Specialist-In-Training 1 summarised,

> 'I actually don't think that the radiologist will be completely replaced, well I think, that it will get better and better. [...] I think that we will always need someone who at least takes a look and says that everything is ok'.

5.4.2 | AI Adapts to the Long-Term Limitations of Itself and Non-Specialists

The wide range of (sometimes conflicting) inputs typically received by non-specialists allowed the AI application to self-organise into the role of 'just another perspective'. The inaccessibility of specialists at certain times meant the AI application was not replacing the role of specialists, but rather it helped non-specialists to proceed when diagnoses were relatively straightforward, and to escalate to specialists when diagnoses were more ambiguous. The AI application adapted to this role by ensuring it was weighted to prefer false positives over false negatives, meaning its impact on non-specialists was typically to elicit more caution, rather than less. This did not mean that non-specialists were not influenced by the reliability of the system; only that the ways in which they made decisions allowed them to be more tolerant of imprecise AI recommendations than specialists. This 'just another perspective' role for the AI application meant that, in contrast to specialists, non-specialists were often impressed with the performance of the AI application, especially compared with earlier versions. For example, Internist 2 reflected that, 'I think now it is a little bit more reliable and more detailed than at the beginning... That encourages me to always look at the AI and which result it formed'.

The ability for non-specialists to generate value from the AI, even when its recommendations were not always reliable, create the opportunity for the AI application to expand its capabilities and interactions. In particular, non-specialists highlighted the opportunity to expand the system from just looking at X-rays to also looking at other imaging technologies, such as CT and magnetic resonance imaging (MRI). As Surgeon 1 explained,

> 'I would use artificial intelligence in everything if it was available. It tells you look twice, so why not?'

Non-specialists' appetite for more AI recommendations provided impetus for the development team and specialists, who began to see the knock-on effect on patient outcomes. The Medical Specialist in the development team remarked that, 'with the system that we have now, I hope and think that this already helps [non-specialists] and that we can change something for the quality of care for patients'. In response the development team expanded the product's capabilities towards a multi-modal AI by creating a product platform, which included different AI extensions for various use cases. This development was still ongoing at the time of study and not yet implemented at our case site.

5.5 | Phase Transition

Similar to many previous accounts of emergence, phase transition played an important role in the emergence of augmented intelligence in the case context. Phase transition occurred roughly in parallel for both specialists and non-specialists, as each group found a way to integrate the AI application into a new system of augmented intelligence. For specialists, the integration of the AI application for specific tasks not only reduced their participation in those tasks, but also created a new scrutiny on other tasks that could be similarly integrated with the system. Simultaneously, the specialists maintained scrutiny over the recommendations produced by the AI application, as despite positive feedback generated from the detection of pulmonary nodules, they remained sceptical due to their unclear understanding of the exact workings of the AI and the frequent diagnostic errors made by the AI application. For non-specialists, the integration of the AI application was part of a trade-off between the accuracy and availability of recommendations. While they observed many issues with the AI application, the non-specialists viewed this recommendations as part of a larger assemblage of diagnostic inputs. They thus felt increasingly comfortable and confident with their use of the AI application so that their use of the application became routinised.

5.5.1 | Specialists' AI Use for One Specific Task Causes Formalisation of Additional Tasks

Over time, as the AI application demonstrated high levels of reliability and the dialogue between specialists and the development team ensured sufficient mutual understanding, specialists began to routinise the system and formalised it as part of diagnosis practices. The risk averse nature of specialists meant that the system was integrated gradually and within targeted domains, rather than being applied to the more complex and uncertain tasks. As the system demonstrated value, and the development team and relevant specialists generated a mutual understanding, the AI application was embedded as an extra precaution. Specialist-In-Training 3 explained this tentative, selective, and risk averse routinisation,

> 'In total it is balanced, that for the small things that I might have missed [without the AI application], I write a sentence more for when it is wrong. After all that is fine, because it is not a large time effort, [...] it is presented automatically, I don't have to click or load anything, [but if I had to] then I wouldn't use it'.

Thus, specialists also found the system useful when it did not confirm their diagnosis, as this encouraged those specialists to challenge their thinking, even if they generally ended up sticking with their own diagnosis. As such, a non-confirming AI recommendation did not necessarily lead to insecurity or doubt, because specialists were confident enough to justify their decision based on previous experience, professional knowledge, and their human-logic guided decision making process.

5.5.2 | Non-Specialists' Routinisation of the AI Creates New Structures

The routinisation of the AI application, and the formation of intelligence that resulted from understanding its information priorities, appeared to empower non-specialists to interact with patients, specialists, and the AI application differently. Non-specialists became more confident that they could preempt some of the feedback they would receive from specialists. They also became more confident ignoring or overruling the AI recommendations, as they increasingly felt they understood the limitations of the system. Internist 3 summarised this, explaining that when discrepancies arise, they 'usually follow the clinical representation of the patient [and] just continue without the X-ray itself'. This meant they could slip back into more conventional processes when the AI recommendations did not seem suitable.

The effect of this was that non-specialists did not necessarily become faster with their own diagnosis for the tasks where they had support from the AI application. However, the overall process became faster because non-specialists felt they could make progress, even when specialists were not available. For example, Internist 1 reflected that, 'The system is not a big part of my normal day [but]... it is a big part of my night shifts'. Thus, the changes became embedded within larger systemic changes in rhythm at the medical centre and affiliated hospitals, as the AI application did not suffer from the same limitations in terms of working hours as typical specialists.

More broadly, the routinisation of the AI application thus meant these non-specialists became better informed, more focused, and more capable of progressing a diagnosis without waiting to engage in dialogue with a specialist. This change allowed them more autonomy from those specialists, provided the AI application could fill in some of the gaps in the system. Non-specialists also became better at differentiating between the types of routine diagnoses that could be processed reliably by the AI application, and the types of information it prioritised, and those that were more challenging and required alternative information or discussion.

5.6 | New Emerging Tensions

The integration of the AI application addressed some of the causes of previous instabilities, particularly previous frustrations regarding the availability of specialists and the amount of time spent on seemingly routine diagnoses. However, the integration of the AI application was also linked with new emerging tensions. For specialists, these new tensions surrounded the additional demands to triangulate their diagnoses with the recommendations of the AI application, and to consider and often correct erroneous AI recommendations. For non-specialists, new emerging tensions were linked to their growing dependency on the AI application and the concern that this dependency could start to black box the logic behind the diagnoses, especially if it became less normalised to discuss these diagnoses with specialists.

5.6.1 | Specialists Experience an Increased Triangulation of AI Recommendations Across Tasks

Specialists' focus on return of effort introduced a new and interesting tension into the use of the AI application. In order to be routinised, the system had to be intuitive for specialists and integrate easily into their workflow, without forcing them to change their basic approach. However, the resulting simplicity of the system meant it became predictable, with the result that the perceived value to specialists diminished over time. Specialists used it to check and improve their diagnoses, treating any divergences as an opportunity to critically reflect upon their own processes and those of the AI application. Once they had done this, the frequency of divergences became fewer and the need to refer to the AI application decreased. For example, Specialist-In-Training 1 lamented that,

> 'If it stays at this level I really don't think [we will keep using it] because the more you are experienced in your field, the more you already know. At a first glance on your X-Ray you already know, so here is the problem [...] so we don't need extra help... if it adds something to your daily work, then [maybe yes, we keep using it]'.

This meant the system needed to adapt so that specialists could extract value without relying on continuous dialogue with the development team or effort-intensive interactions with the AI application. This reality of ongoing commitment to dialogue conflicted with one of the initial motivations for many specialists, which was to use the AI application to free up time. However, the improved decision making that occurred because of the continuous dialogue also increased these specialists' respect for the AI application. This led to a crystallisation of the role of the AI application among specialists; a low-effort tool for triangulating their diagnoses. Specialist-In-Training 3 explained,

> 'If we are well staffed, I take those extra seconds [...] it is rather a small percentage, where you see something, that was not relevant to the acute situation, but that might have a relevance later on [...] by the end of the day you are also happy about the small benefits ... maybe it makes you a bit slower after all [...] but if you're a little bit better in the end, compared to without AI, then I think it is worth something [...]'.

Because the development team adapted to this changing role for the AI application by ensuring AI recommendations were presented automatically, it created a sense of expectation that specialists would consider them. Specialist 4 remarked that 'now, because the AI advice is available and also presented automatically, [she] has to look at it, even though it most likely up to 80% has marked something that is not pathological'. This new role meant the system was continuously challenging the decision-making outcomes and decision-making criteria of specialists. This safeguarded against human fatigue, while also pushing specialists to explain the reasons when they make recommendations at odds with the AI application. Specialist 4 described the impact this triangulation had upon diagnoses, remarking:

> 'If it is more precise and these false positives are not marked any longer, then definitely it makes life easier, because you look at it again with focus [to identify something that you haven't recognized before] or your assumption that there actually is something suspicious is supported. But for [it to work] that [way] it would need to perform better'.

This also fuelled further cycles of development, as specialists demanded the integration of more human-like decision making processes, for example, incorporating previous images or additional demographic data about the patients, so that they are enabled to engage further in the more complicated and ambiguous cases. For example, Specialist-In-Training 1 noted 'it would be interesting to see, which parameters it is really using to detect pathologies', while Specialist-In-Training 3 complained that 'the biggest crux is actually that [the AI] cannot compare previous examinations and can thus not detect whether something is constant or new. [...] That is how misinterpretations happen'. The development team recognised the potential of integrating previous images to improve accuracy and human-logic guided explainability and were working on adjusting the system accordingly at the time of study.

5.6.2 | Anticipated Dependency on the AI Application Concerns Non-Specialists

Over time, non-specialists came to increasingly rely on the AI recommendations, and this reliance was perceived to lighten the workload of specialists and allow them to focus on the most complex or critical cases. However, this growing reliance also presented a new tension in the augmented intelligence system. Non-specialists were empowered to take on more decision-making, and to rely on their discretion when interpreting information. However, the more effective the AI became, the less critical the non-specialists became of its recommendations. Specialist 3 described how 'things are misconstrued, or you rely too much on [the AI] and don't engage with the topic sufficiently yourself, [in other words] that the doctor on-duty relies too heavily on the AI and doesn't visualize it himself and only checks what is offered'.

The potential for deteriorating critical reflection was viewed as a minor threat among present users but perhaps as a more significant threat for newer non-specialists in the future. Internist 2 expressed his concern that 'those coming after me will not learn to diagnose by themselves anymore because they rely on the AI'

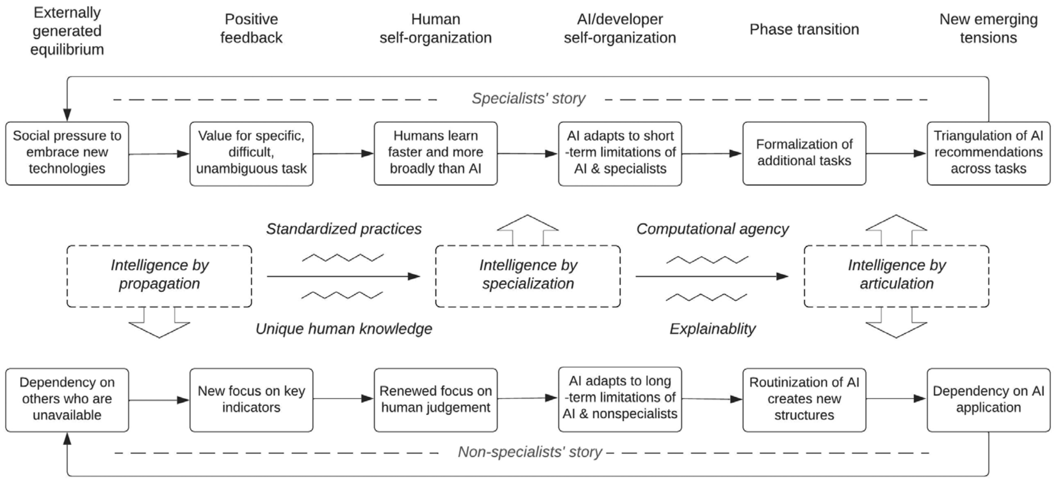

FIGURE 4 | A model of emerging augmented intelligence.

before noting that 'but then again if the AI is good enough and available, that does not really matter'. This growing dependency conjured mixed feelings in non-specialists, in part, because many would have preferred to have simply received a single, non-contradictory recommendation from specialists to begin with. They consequently did not fear losing their skills for independent evaluation, because they felt they should probably not have been required to develop those skills if proper support was available. This meant each successive improvement in the accuracy in the system also threatened to make its mistakes more consequential, as non-specialists could begin to assume it was more reliable resulting in less critical reflection.

6 | Discussion

This study investigated how new forms of intelligence emerge in augmented intelligence systems. The ability for these systems to enable new forms of intelligence is important, given existing evidence that integrating human and artificial intelligence can be harmful to each, individually. We applied an emergence perspective to guide our investigation, with particular focus on the key stages of emergence described in existing literature. We observed two distinct stories of emerging augmented intelligence in our case context. These two stories contained six similar stages, which mapped partially, but imperfectly to the four general stages of emergence described in existing literature.

First, we observed that self-organisation had to be decomposed into 'human self-organisation' and 'AI/developer self-organisation'. Existing descriptions of self-organisation suggest that systems generally self-organise towards stability (Bergmann Lichtenstein 2000). However, we observed a qualitative difference in the human self-organisation that occurred in the levels around the AI application, and the AI/developer self-organisation that occurred within the interactions with the AI application itself. While the latter was still human-driven in many respects, as it was directed and supervised by the development team as 'humans-in-the-loop' (cf. Grønsund and Aanestad 2020), the AI/developer self-organisation was nonetheless subject to an independent internal logic; a logic which inverted the relationship between attentiveness and interaction. While the human actors needed to conserve attention by selectively interacting, the AI application became increasingly attentive as it was afforded extensive, continuous interaction. Thus, while human self-organisation often resolved tensions, AI/developer self-organisation often created them.

Second, we observed an additional stage, which we called 'new emerging tensions'. The contrasting nature of human self-organisation and AI/developer self-organisation meant the AI application tended to subsume more and more responsibility over time. This appeared to create new internal tensions at multiple levels each time the system went through a phase transition, assuming these phase transitions led to a deeper integration of the AI application. By substituting interactions with the specialists for interactions with the AI application, each phase transition challenged both the position and role of specialists within the system. By substituting non-specialists' dependency on diverse human expertise for a dependency on algorithmic accuracy, each phase transition challenged those non-specialists' need to triangulate and critique recommendations. Finally, by expanding the responsibility of the AI application, each phase transition challenged the ability to allocate responsibility and accountability to human agents, which remains an important concern for both doctors and patients (Ghassemi, Oakden-Rayner, and Beam 2021). Thus, each phase transition created new instabilities between the specific healthcare system and adjacent systems, even (perhaps especially) if it resulted in better health outcomes.

Interestingly, despite these tendencies of the system which appear to elevate the AI application over time, we observed that the emergent new forms of intelligence originated within the human actors, rather than within the AI application. This did not mean the AI application was not responsible for new forms of intelligence; rather, it appears that the AI application catalysed the emergence of human intelligence in the system. The following sections discuss the underlying stages by which systemic intelligence was increased, as well as explicate the generating mechanisms that drove this process (Pentland 1999). Figure 4 presents an overview of this elaborated process theory.

6.1 | The Underlying Stages by Which Augmenting AI Increased Systemic Intelligence

Our findings revealed two stories by which systemic intelligence was perceived to improve over time. Not only do these stories provide a rich account of specialists' and non-specialists' experiences; they also share a common underlying *fabula* of events and relationships that led to these gains.

The first, and perhaps most intuitive stage, was *intelligence by propagation*. The augmented intelligence system allowed specialists' knowledge to spread to non-specialists on a day-to-day basis, grounded in those non-specialists' experience with specific contexts. Non-specialists could then make more intelligent decisions regarding a diagnosis, even when specific specialists were unavailable. This stage also afforded these non-specialists some degree of training, as prolonged exposure to the AI application meant they became more capable of anticipating AI recommendations and the reasons behind them. Intelligence by propagation was therefore mainly a means of increasing intelligence among the non-specialists in the system.

This stage can be compared to the learning that emerges from a knowledge management system (Schultze and Leidner 2002), albeit the knowledge that is shared is embedded within specific recommendations. The AI application also differed from traditional knowledge management systems in how it engaged users. Knowledge management systems have historically been limited by the social costs of sharing knowledge, as the most knowledgeable people in an organisation do not always have clear incentives to share what they know (Kankanhalli, Tan, and Wei 2005). Further, the act of sharing often brings the sharer under additional scrutiny, as it can highlight inefficiencies in their practices and blind spots in their understanding (Young, Kuo, and Myers 2012). The AI application overcame these challenges because it incentivised specialists to share certain routine types of knowledge so they could offload some time consuming tasks, without necessarily exposing all areas of their professional practices to ridicule. As a result, the augmented intelligence system did not necessarily have to create new knowledge to effectively increase intelligence across the system. Rather, it increased systemic intelligence by elevating the decision-making intelligence of the least well-trained human agents of the system for the given diagnosis tasks. In other words, it made the system smarter by lifting the baseline levels of decision-making intelligence.

The second stage in which the introduction of the reconfigurable AI application increased systemic intelligence was *intelligence by specialisation*. In addition to providing more timely and consistent support to non-specialists, the AI application triggered specialists' to deemphasise routine consultations and instead engage more with complicated or ambiguous cases. While the addition of AI recommendations reduced the effort required by specialists to formulate detailed reports for routine diagnoses—a custom that appeared to offer a diminishing learning opportunity for them—it pressured them to redirect their attention towards the more difficult diagnoses. Intelligence by specialisation was therefore mainly a means of increasing intelligence among the specialists in the system.

The opportunities of AI applications to increase efficiency are well documented in existing literature, with many organisations realising value through process automation to reduce human labour (Shollo et al. 2022). However, these applications rarely discuss the potential gains in mutual learning that occur when that human labour is redirected towards more challenging problems. Instead, AI-based process automation is often described as a threat to human participation, as more and more jobs become 'computerised' and human agents become redundant (Frey and Osborne 2017). Viewed in terms of augmented intelligence and mutual learning, our case context showed how the addition of the AI for routine recommendations challenged specialists to become more specialised, and so increased the depth of knowledge in the system as a whole. This pattern is at odds with many anticipated applications of AI in contexts such as healthcare and image screening, where deep learning-based AI is expected to tackle some of the more complex problems that are beyond human comprehension (cf. Hosny et al. 2018). Instead, our findings suggest systemic intelligence gains occur when human agents address the most dynamic and challenging problems, building on the support of reliable AI recommendations for routine tasks. Specialists become more specialised and sophisticated in their work, and this feeds into more advanced algorithmic support as mutual learning increases. This learning becomes embedded in the algorithms, which allows (and demands) that specialists find new problems where they can apply their human knowledge to create additional intelligence.

The third stage in which the introduction of AI increased systemic intelligence, and arguably the most subtle, was *intelligence by articulation*. Previous research has illustrated that AI applications are more effective when human decision makers make use of 'metacognitive processes' and engage in 'AI-interrogation practices' to evaluate AI recommendations (Jussupow et al. 2021; Lebovitz, Lifshitz-Assaf, and Levina 2022). Our findings suggest this effect may be more pronounced than what is described in existing literature. We observed that, as specialists and non-specialists interacted with the AI application, they became more aware of what they knew and what they did not know. Specialists and non-specialists also appeared to become more aware of what they could and could not clearly articulate. Further, the act of articulating knowledge may also have helped individuals to advance their own understanding, as the ability to clearly symbolise and externalise knowledge is part of the enactment of higher psychological processes (Vygotsky 1980). This also occurred when AI recommendations disagreed with specialists' conclusions. Under such conditions, both specialists and non-specialists were prompted to scrutinise the situation more closely. Individuals did not appear to necessarily integrate the logic of the AI in such instances. Instead, if they applied human logic to arrive at a conclusion and the machine disagreed, they would often apply more human logic to see if their conclusion changed. Intelligence by articulation was therefore a means of increasing intelligence among both specialists and non-specialists in the system.

The relationship between human agents and the AI application thus created a new mediator of peer learning, in which the specialists benefitted from teaching the AI application, in the same way experienced clinicians benefit from teaching novices

(Weiss and Needlman 1998; Ten Cate and Durning 2007). Thus, the introduction of AI tools created new forms of intelligence, as it created an incentive to articulate and reflect upon practices that may otherwise have remained implicit. It is also possible, although we did not have the data to examine this, that the increased symbolization added to the morality of decision making, as such symbolization also allows individuals to internalise moral cues in their environment and to develop a more consistent and salient moral identity (Aquino and Reed II 2002). In other words, where specialists chose to overrule the AI recommendations for moral reasons, this may have helped verbalise those moral reasons in ways that allowed them to become normalised.

6.2 | The Generating Mechanisms of Augmented Intelligence

The parallel stories arising from the fabula, as well as the asymmetrical tensions and pressures perceived by human agents, suggest the emergence of augmented intelligence is compilation-based in nature, rather than composition-based, that is, perceptions and behaviours divide into distinct sub-systems, each contributing to larger system outcomes in different ways (Kozlowski and Klein 2000). The stories further describe a process where the integration of AI recommendations to increase systemic intelligence is constructive and exploratory, rather than prescribed within clear a priori planning. Thus, the generating mechanism for the process can be considered 'dialectical'; driven by competing forces and contradictory values (Van de Ven and Poole 1995).

Our findings highlighted two salient sets of competing forces and values. The first was the need to balance *standardised practices* with *unique human knowledge*. While this generating mechanism appeared to drive all three stages in the fabula, it was especially prominent in the progression from intelligence by propagation to intelligence by specialisation.

Existing literature describes this pressure on augmented intelligence systems to both standardise and exploit existing knowledge, and to consider and explore new uniquely human knowledge (cf. Fügener et al. 2021; Allen and Choudhury 2022). Yet, much of the existing discourse around augmented intelligence treats human agents as a single population of knowledgeable users; users whom the AI application must convince it is trustworthy. This view is summarised by Jain et al. (2021, 680), who highlighted three ways AI applications can complement human agents '(i) AI works alongside humans to accomplish peripheral tasks and generally looks to support the human expert. (ii) AI takes over when the human has high cognitive load. (iii) AI replaces humans in areas where humans have limited strengths, or the environment is toxic or when real-time response is key'. This study expands on this view by distinguishing between the highly knowledgeable human agents who often help inform the logic of the AI application and the less knowledgeable human agents who often apply the AI recommendations.

Viewed in this way, augmented intelligence becomes a means of knowledge mediation among a population of heterogeneous agents. In our case context, specialists provided the standardised diagnostic considerations that informed the AI application, and the AI application propagated this intelligence to others in the system. This spread of intelligence allowed those specialists to engage more with problems that could not be addressed by standard considerations, so enabling them to deepen their specialisation and expand the horizon of their unique human knowledge. Building on this propagation of standards and deepening specialisation, both specialists and non-specialists were better positioned to articulate their unique human knowledge, either from their deepening specialisation (for specialists) or their empowerment to engage with patients' context (for non-specialists). Mutual learning is therefore not a case of AI applications 'echoing back' knowledge to the same users who provided it. It is a means to connect continuous learning in different parts of the larger emergent augmented intelligence system. This resonates with findings by Sturm et al. (2021) who argue that domain specialists are needed to continuously reconfigure AI algorithms. However, unlike Sturm et al. (2021), our findings suggest that reconfiguring the AI application may mean allowing it to act as a mediator of human knowledge flows, rather than allowing the AI application to act as a source of knowledge.

The second set of competing forces and values was human agents' simultaneous desire to afford the AI application with the *computational agency* to take advantage of its powerful computational capabilities but also to afford the users with *explainability* to understand and explain how the AI arrives at its recommendations. As with the previous generating mechanism, this appeared to drive all three stages in the fabula. However, it was especially prominent in the progression from intelligence by specialisation to intelligence by articulation.

Competing forces and values around computational agency and explainability have been widely discussed in existing literature (cf. Zhang, Mehta, et al. 2021; Ghassemi et al. 2021). Much of the existing, composition-based discourse around augmented intelligence focuses on distinguishing between automatable tasks and tasks that require human judgement (Frey and Osborne 2017; Benbya et al. 2020). This implies that AI applications should either be designed to imitate or differentiate from the logic of human agents. Our findings suggest an alternative compilation-based dynamic; AI applications may offer a way to construct and deconstruct the accepted logic of human agents as part of an ongoing learning process. At the heart of this process is the discovery of generative conflicts and the routinisation of adapted behaviours (Putnam 2015).

In our case context, external forces meant that multiple entities, the specialists, the non-specialists, and the AI application (including its developers), had to work together towards shared outcomes. The first step was to find 'uncontested' diagnostic practices that could be translated from human knowledge to algorithms with minimal ambiguity or controversy (Lebovitz, Levina, and Lifshitz-Assa 2021). The spread of these practices propagated intelligence among different agents with relatively low resistance. An outcome of this process was that specialists could explore new practices which were less immediately amenable to be translated to algorithms. Specialists could therefore expand their specialised intelligence without immediate pressure to align with computational agency. As the AI application became more widely used and trusted, the desire to integrate more human knowledge grew. This prompted increasing scrutiny of the perceived incompatibility between computational approaches and human intelligence, encouraging human agents and AI developers to articulate critical considerations for their respective approaches. This not only highlighted new opportunities for augmented intelligence, but it also enabled new forms of reflection among the agents in the system (Abdel-Karim et al. 2023).

6.3 | Practical Implications

Explainable AI is a popular topic in the fields of IS and management (Arrieta et al. 2020). Explainable AI appears to be especially important in areas where the need for accountability and adherence to best practice is important, such as health (Lauritsen et al. 2020) and financial services (Asatiani et al. 2020). One approach to balancing the dual demands for computational agency and explainability is to treat explainability as a retrospective process (Lipkova et al. 2022). This approach proposes that AI applications apply computational agency to solve problems, then translate the applied computational logic into a form that allows human agents to understand it. This has also come under criticism, as the interpretability gap means humans tend to explain algorithms by imprecisely analogizing them with human-like decision processes, with the result that 'using post-hoc explanations to assess the quality of model decisions adds an additional source of error—not only can the model be right or wrong, but so can the explanation' (Ghassemi, Oakden-Rayner, and Beam 2021, 747). Our findings suggest an alternative approach to explainability; explainability for mutual learning. This focus on mutual learning changes the requirements for explainability, as the representational focus moves from building trust in AI applications to enabling new models of AI-enabled knowledge coordination. In practical terms, explainability mechanisms may therefore become less focused on highlighting which variables were more or less important for a recommendation, and more focused on *why* specific variables were more or less important. For example, specialists may wish to embed secondary explanations which can help non-specialists deepen their understanding of the phenomenon, beyond the instance in question. Similar to generative AI systems, AI applications may also consider embedding specific sources to encourage both specialists and non-specialists to relate AI recommendations to scientific research of which they may or may not be aware.

Our findings also have specific practical implications for different groups. First, system designers and administrators need to consider that augmented intelligence is a continuous and adaptive process, rather than a one-shot transition from human intelligence to semi-automated decision making. Our case context illustrated that early perceptions of value may therefore be minimal, as human agents may need to invest significant time and effort with few obvious short-term benefits. Further, our findings suggest the generating mechanism for augmented intelligence may be the interaction between specialists, the development team, and non-specialists, as cycles of emerging tensions can produce iteratively more sophisticated algorithms, human representations, and practical routines. This means that encouraging interactions among heterogenous user groups and AI developers may be key to realising new forms of intelligence. It is therefore important that the introduction of augmented intelligence systems is carefully incentivised, so as to encourage both routine engagement and curiosity. Our case context took advantage of a wider excitement about technology and AI, prompting specialists, in particular, to engage with the system out of curiosity and a desire to stay up to date with professional discourse. In other contexts where such wider excitement is not prominent, AI developers may need to stimulate excitement through workshops and other strategic initiatives.

Second, system users need to recognise that the aim of the system is not necessarily to make their job easier, or to relieve them of responsibility. Part of integrating AI recommendations is learning when they are not reliable and why, and forming a healthy scepticism about how it arrives at its recommendations. System users may benefit from approaching AI applications as a learning technology, rather than a tool for automation. They may also need to consider that incremental learning with the AI alone is not sufficient. Rather, they should treat the system as an enabler of mutual learning with different groups. One way to encourage this type of approach may be to prompt users to document and share practical experiences, particularly practical experiences where those users discovered limitations in either the AI or their own analytical processes. Such sharing of experiences may have benefits beyond the discovery of specific limitations, if it can nurture wider reflective practices around the use of AI.

Third, this study has implications for policy makers and insurers. Much of the discussion around accountability and augmented intelligence systems assumes that accountability shifts to the human agents using such systems, as these human agents must ultimately interpret recommendations and make the final decisions (Saunders et al. 2017; Fügener et al. 2021). The findings of this study suggest this simple allocation of accountability may be misleading. We observe that augmented intelligence can operate on the collective level, by circulating decision making norms across heterogeneous agents. This makes it difficult to allocate praise or blame for specific outcomes, as each decision has direct or indirect input from a range of human contributors, including the specialists who informed the logic and the AI developers who translated that logic into computation. One possibility to address the need for accountability may be to introduce formal auditing practices which monitor the evolution of augmented intelligence systems, and which ensure that emerging practices remain grounded in established values and principles.

Finally, we also observed some indicators that the system could potentially diminish systemic intelligence over time, especially if the reliability of the system reaches levels where non-specialists stop critiquing AI recommendations. Research has long highlighted this threat that human knowledge can diminish when human agents are given sophisticated decision-making tools (Arnold and Sutton 1998; Rinta-Kahila et al. 2018; Strich, Mayer, and Fiedler 2021). There are also suggestions that human knowledge representations may adapt towards computation-friendly representations, potentially losing important nuance in the process (Introna 2016). It is not clear whether these threats are significant in our study context. The potential for complacency seems like it may be counterbalanced by the dynamism of the system and the tendency towards continuous mutual learning.

However, time will tell whether non-specialists maintain their impetus for such learning, or whether the larger balance of power and assumed ‘ground truth’ (cf. Lebovitz, Levina, and Lifshitz-Assa 2021) shifts to AI applications over time, with the result that human agents are willing to concede more and more of the unique human knowledge in search of operational convenience. For example, in our case context, the judgements of the specialists were arguably perceived as the ground truth, or at least, the closest thing available to it. This meant that the AI application recommendations were considered correct or incorrect, based on how closely they mirrored those specialists' judgements. As the AI application continues to evolve and more data becomes available about long term patient outcomes, it could become possible to compare specialist and AI recommendations. At some point in the future, it is conceivable that the AI recommendations could overtake specialists in terms of diagnosis accuracy. It is not clear how this would impact mutual learning, particularly if the growing predictive power enabled by greater computational agency justifies a decreasing focus on explainability.

7 | Conclusions

This study advances our understanding of augmented intelligence by examining how it emerges in a complex real-world setting. Specifically, we focused on the emergence of augmented intelligence in a medical context, building on several existing studies of AI in this domain (van Beek, Mirsadraee, and Murchison 2015; Lebovitz 2019; Yoo et al. 2021). However, there is a range of other contexts where augmented intelligence is applied, including legal decision support (Angwin, Larson, and Kirchner 2016), in the financial market (Ge et al. 2021), and for digital platform businesses (Rai, Constantinides, and Sarker 2019). With these varied contexts in mind, we sought to theorise the underlying mechanisms by which augmented intelligence can emerge. We discovered a dialectic process that may also be useful to inform future studies of these other contexts. Perhaps more importantly, we discovered that nurturing augmented intelligence may mean that human agents allow AI applications to take responsibility for some of the more mundane tasks, rather than some of the most dynamic and loosely defined problems at the frontier of specialised knowledge.

This study has several notable limitations. First and foremost, data gathering took place over a period of less than 12 months. This period was sufficient to observe the process of emergence over multiple adaptations of the system. However, it is not long enough to observe the longer-term evolution of AI and practice. Such trends are ongoing and subsequent research should build on our findings in coming years and decades to position them within a longer timeframe. Second, we focused on a single case context, based on a single AI application and multiple physical sites. This allowed us to perform an in-depth analysis and stay close to the subtleties of our chosen environment. Yet, we expect that different contexts will display different idiosyncrasies and different dynamics. Thus, we encourage other studies to compare findings, and particularly other contexts where the distinction between specialists and non-specialists is less clearly defined. Third, we relied heavily on self-reported data, rather than behavioural observations. This was pragmatic, as many of the medical diagnoses are sensitive in nature and difficult to observe without incurring concerns over privacy and ethics. Future research may wish to include medical practitioners who can better manage the obstacles when securing research access for participant observation. Fourth, explainable AI is an important contemporary topic when discussing the acceptance of AI (Rai 2020). Our findings raise interesting questions for that academic discourse, such as how explainability interacts with dialectic learning, and whether this relationship changes over time. While we touched on this topic, there is clearly more to be learned about this dynamic. We hope our findings provide an impetus for future research to explore this further.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Research data are not shared, as the qualitative nature of data makes it difficult to effectively anonymize.

References

Abdel-Karim, B., N. Pfeuffer, K. V. Carl, and O. Hinz. 2023. “How AI-Based Systems Can Induce Reflections: The Case of AI-Augmented Diagnostic Work.” *MIS Quarterly* 47, no. 4: 1395–1424.

Abdelzaher, T., S. Wang, P. Giridhar, et al. 2017. “Chapter 11—Social Network Signal Processing for Cyber-Physical Systems.” In *Cyber-Physical Systems*, edited by H. Song, D. B. Rawat, S. Jeschke, and C. Brecher, 161–174. Amsterdam: Academic Press.

Abràmoff, M. D., P. T. Lavin, M. Birch, N. Shah, and J. C. Folk. 2018. “Pivotal Trial of an Autonomous AI-Based Diagnostic System for Detection of Diabetic Retinopathy in Primary Care Offices.” *npj Digital Medicine* 1: 39.

Agrawal, A., J. S. Gans, and A. Goldfarb. 2019. “Exploring the Impact of Artificial Intelligence: Prediction Versus Judgment.” *Information Economics and Policy* 47: 1–6.

Allen, P. M., and L. Varga. 2006. “A Co–Evolutionary Complex Systems Perspective on Information Systems.” *Journal of Information Technology* 21, no. 4: 229–238.

Allen, R., and P. Choudhury. 2022. “Algorithm-Augmented Work and Domain Experience: The Countervailing Forces of Ability and Aversion.” *Organization Science* 33, no. 1: 149–169.

Anderson, P. 1999. “Perspective: Complexity Theory and Organization Science.” *Organization Science* 10, no. 3: 216–232.

Angwin, J., J. Larson, and L. Kirchner. 2016. “Machine Bias.” ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Aquino, K., and A. Reed II. 2002. “The Self-Importance of Moral Identity.” *Journal of Personality and Social Psychology* 83, no. 6: 1423–1440.

Arnold, V., and S. G. Sutton. 1998. “The Theory of Technology Dominance: Understanding the Impact of Intelligent Decision Aids on Decision Maker's Judgments.” *Advances in Accounting Behavioral Research* 1, no. 3: 175–194.

Arrieta, A. B., N. Díaz-Rodríguez, J. Del Ser, et al. 2020. “Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges Toward Responsible AI.” *Information Fusion* 58: 82–115.

Asatiani, A., P. Malo, P. R. Nagbøl, E. Penttinen, T. Rinta-Kahila, and A. Salovaara. 2020. “Challenges of Explaining the Behavior of Black-Box AI Systems.” *MIS Quarterly Executive* 19, no. 4: 259–278.

Balijepally, V., R. Mahapatra, S. Nerur, and K. H. Price. 2009. "Are Two Heads Better Than One for Software Development? The Productivity Paradox of Pair Programming." *MIS Quarterly* 33, no. 1: 91–118.

Bandura, A. 1977. "Self-Efficacy: Toward a Unifying Theory of Behavioral Change." *Psychological Review* 84, no. 2: 191–215.

Bansal, G., B. Nushi, E. Kamar, W. Lasecki, D. S. Weld, and E. Horvitz. 2019a. "Beyond Accuracy: The Role of Mental Models in Human-AI Team Performance." *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* 7, no. 1: 2–11. https://ojs.aaai.org/index.php/HCOMP/article/view/5285.

Bansal, G., B. Nushi, E. Kamar, D. S. Weld, W. S. Lasecki, and E. Horvitz. 2019b. "Updates in Human-AI Teams: Understanding and Addressing the Performance/Compatibility Tradeoff." *Proceedings of the AAAI Conference on Artificial Intelligence* 33, no. 1: 2429–2437.

Benbya, H., and B. McKelvey. 2006a. "Toward a Complexity Theory of Information Systems Development." *Information Technology & People* 19, no. 1: 12–34.

Benbya, H., and B. McKelvey. 2006b. "Using Coevolutionary and Complexity Theories to Improve IS Alignment: A Multi-Level Approach." *Journal of Information Technology* 21: 284–298.

Benbya, H., N. Nan, H. Tanriverdi, and Y. Yoo. 2020. "Complexity and Information Systems Research in the Emerging Digital World." *MIS Quarterly* 44, no. 1: 1–17.

Berente, N., G. Bin, J. Recker, and R. Santhanam. 2021. "Managing Artificial Intelligence." *MIS Quarterly* 45, no. 3: 1433–1450.

Bergmann Lichtenstein, B. M. 2000. "Emergence as a Process of Self-Organizing-New Assumptions and Insights From the Study of Non-Linear Dynamic Systems." *Journal of Organizational Change Management* 13, no. 6: 526–544.

Brynjolfsson, E., and A. McAfee. 2017. "Artificial Intelligence, For Real." Harvard Business Review. https://hbr.org/2017/07/the-business-of-artificial-intelligence.

Burgess, A., D. McGregor, and C. Mellis. 2014. "Medical Students as Peer Tutors: A Systematic Review." *BMC Medical Education* 14, no. 1: 1–8.

Chakraborty, S., S. Sarker, and S. Sarker. 2010. "An Exploration Into the Process of Requirements Elicitation: A Grounded Approach." *Journal of the Association for Information Systems* 11, no. 4: 219–249.

Charmaz, K. 2000. "Grounded Theory: Objectivist and Constructivist Methods." In *Handbook of Qualitative Research*, edited by N. K. Denzin and Y. S. Lincoln, 509–535. Thousand Oaks, CA: Sage.

Chiles, T. H., A. D. Meyer, and T. J. Hench. 2004. "Organizational Emergence: The Origin and Transformation of Branson, Missouri's Musical Theaters." *Organization Science* 15, no. 5: 499–519.

Choi, T. Y., K. J. Dooley, and M. Rungtusanatham. 2001. "Supply Networks and Complex Adaptive Systems: Control Versus Emergence." *Journal of Operations Management* 19, no. 3: 351–366.

Compeau, D. R., and C. A. Higgins. 1995. "Computer Self-Efficacy: Development of a Measure and Initial Test." *MIS Quarterly* 19, no. 2: 189–211.

Davison, R., and M. Martinsons. 2015. "Context Is King! Considering Particularism in Research Design and Reporting." *Journal of Information Technology* 31: 241–249. https://doi.org/10.1057/jit.2015.19.

De Lima Salge, C. A., and N. Berente. 2017. "Is That Social Bot Behaving Unethically?" *Communications of the ACM* 60, no. 9: 29–31.

Dooley, K. J., and A. H. Van de Ven. 1999. "Explaining Complex Organizational Dynamics." *Organization Science* 10, no. 3: 358–372.

Doran, D., S. Schulz, and T. R. Besold. 2017. "What Does Explainable AI Really Mean? A New Conceptualization of Perspectives." ArXiv:1710.00794 [Cs]. http://arxiv.org/abs/1710.00794.

Dwivedi, Y. K., L. Hughes, E. Ismagilova, et al. 2021. "Artificial Intelligence (AI): Multidisciplinary Perspectives on Emerging Challenges, Opportunities, and Agenda for Research, Practice and Policy." *International Journal of Information Management* 57: 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002.

Eisenhardt, K. M. 1989. "Building Theories From Case Study Research." *Academy of Management Review* 14, no. 4: 532–550.

Faraj, S., S. Pachidi, and K. Sayegh. 2018. "Working and Organizing in the Age of the Learning Algorithm." *Information and Organization* 28, no. 1: 62–70.

Floridi, L., J. Cowls, M. Beltrametti, et al. 2018. "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds and Machines* 28, no. 4: 689–707.

Flyvbjerg, B. 2006. "Five Misunderstandings About Case-Study Research." *Qualitative Inquiry* 12, no. 2: 219–245.

Frey, C. B., and M. A. Osborne. 2017. "The Future of Employment: How Susceptible Are Jobs to Computerisation?" *Technological Forecasting and Social Change* 114: 254–280.

Fügener, A., J. Grahl, A. Gupta, and W. Ketter. 2021. "Will Humans-In-The-Loop Become Borgs? Merits and Pitfalls of Working With AI." *MIS Quarterly* 45, no. 3: 1527–1556.

Gaur, Y. 2015. "The Effects of Automatic Speech Recognition Quality on Human Transcription Latency." In *Proceedings of the 17th International ACM SIGACCESS Conference on Computer & Accessibility*, 367–368.

Ge, R., Z. Zheng, X. Tian, and L. Liao. 2021. "Human–Robot Interaction: When Investors Adjust the Usage of Robo-Advisors in Peer-to-Peer Lending." *Information Systems Research* 32, no. 3: 774–785.

Ghassemi, M., L. Oakden-Rayner, and A. L. Beam. 2021. "The False Hope of Current Approaches to Explainable Artificial Intelligence in Health Care." *Lancet Digital Health* 3, no. 11: e745–e750.

Goldstein, J. 1999. "Emergence as a Construct: History and Issues." *Emergence* 1, no. 1: 49–72.

Grønsund, T., and M. Aanestad. 2020. "Augmenting the Algorithm: Emerging Human-in-the-Loop Work Configurations." *Journal of Strategic Information Systems* 29, no. 2: 101614.

Gulshan, V., L. Peng, M. Coram, et al. 2016. "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs." *Journal of the American Medical Association* 316, no. 22: 2402–2410.

Haenssle, H. A., C. Fink, R. Schneiderbauer, et al. 2018. "Man Against Machine: Diagnostic Performance of a Deep Learning Convolutional Neural Network for Dermoscopic Melanoma Recognition in Comparison to 58 Dermatologists." *Annals of Oncology* 29, no. 8: 1836–1842.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. "Unsupervised Learning." In *The Elements of Statistical Learning*, edited by T. Hastie, R. Tibshirani, and J. Friedman, 485–585. New York: Springer.

Holland, J. H. 1995. *Hidden Order: How Adaptation Builds Complexity.* Massachusetts: Addison-Wesley Publishing Company.

Holzinger, A. 2018. "From Machine Learning to Explainable AI." In *2018 World Symposium on Digital Intelligence for Systems and Machines (DISA)*, 55–66. https://doi.org/10.1109/DISA.2018.8490530.

Hosny, A., C. Parmar, J. Quackenbush, L. H. Schwartz, and H. J. Aerts. 2018. "Artificial Intelligence in Radiology." *Nature Reviews Cancer* 18, no. 8: 500–510.

Huston, M. 2017. "Self-Taught Artificial Intelligence Beats Doctors at Predicting Heart Attacks." https://www.science.org/content/article/self-taught-artificial-intelligence-beats-doctors-predicting-heart-attacks.

Introna, L. D. 2016. "Algorithms, Governance, and Governmentality: On Governing Academic Writing." *Science, Technology & Human Values* 41, no. 1: 17–49.

Jain, H., B. Padmanabhan, P. A. Pavlou, and T. S. Raghu. 2021. "Editorial for the Special Section on Humans, Algorithms, and Augmented Intelligence: The Future of Work, Organizations, and Society." *Information Systems Research* 32, no. 3: 675–687.

Jia, N., X. Luo, Z. Fang, and C. Liao. 2024. "When and How Artificial Intelligence Augments Employee Creativity." *Academy of Management Journal* 67, no. 1: 5–32.

Jussupow, E., K. Spohrer, A. Heinzl, and J. Gawlitza. 2021. "Augmenting Medical Diagnosis Decisions? An Investigation Into Physicians' Decision-Making Process With Artificial Intelligence." *Information Systems Research* 32, no. 3: 713–735.

Kamar, E., and E. Horvitz. 2012. "Combining Human and Machine Intelligence in Large-Scale Crowdsourcing." In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*.

Kanfer, R., and P. L. Ackerman. 1989. "Motivation and Cognitive Abilities: An Integrative/Aptitude-Treatment Interaction Approach to Skill Acquisition." *Journal of Applied Psychology* 74, no. 4: 657–690.

Kankanhalli, A., B. C. Tan, and K. K. Wei. 2005. "Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation." *MIS Quarterly* 29, no. 1: 113–143.

Kauffman, S. A. 1992. "The Origins of Order: Self-Organization and Selection in Evolution." In *Series on Directions in Condensed Matter Physics*, edited by D. L. Stein, vol. 6, 61–100. Oxford: Oxford University Press.

Klein, H. K., and M. D. Myers. 1999. "A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems." *MIS Quarterly* 23, no. 1: 67–93. https://doi.org/10.2307/249410.

Kokkodis, M. 2021. "Dynamic, Multidimensional, and Skillset-Specific Reputation Systems for Online Work." *Information Systems Research* 32, no. 3: 688–712.

Kozlowski, S. W. J., G. T. Chao, J. A. Grand, M. T. Braun, and G. Kuljanin. 2013. "Advancing Multilevel Research Design: Capturing the Dynamics of Emergence." *Organizational Research Methods* 16, no. 4: 581–615.

Kozlowski, S. W. J., and K. J. Klein. 2000. "A Multilevel Approach to Theory and Research in Organizations: Contextual, Temporal, and Emergent Processes." In *Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions, and New Directions*, edited by K. J. Klein and S. W. J. Kozlowski, 3–90. San Francisco: Jossey-Bass.

Krogh, A. 2008. "What Are Artificial Neural Networks?" *Nature Biotechnology* 26, no. 2: 195–197.

Lauritsen, S. M., M. Kristensen, M. V. Olsen, et al. 2020. "Explainable Artificial Intelligence Model to Predict Acute Critical Illness From Electronic Health Records." *Nature Communications* 11, no. 1: 1–11.

Lebovitz, S. 2019. "Diagnostic Doubt and Artificial Intelligence: An Inductive Field Study of Radiology Work." In *International Conference on Information Systems*, 11. https://aisel.aisnet.org/icis2019/future_of_work/future_work/11.

Lebovitz, S., N. Levina, and H. Lifshitz-Assa. 2021. "Is AI Ground Truth Really True? The Dangers of Training and Evaluating AI Tools Based on Experts' Know-What." *MIS Quarterly* 45, no. 3: 1501–1526.

Lebovitz, S., H. Lifshitz-Assaf, and N. 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, no. 1: 126–148.

Lewin, R. 1992. *Complexity: Life at the Edge of Chaos*. Chicago, IL: University of Chicago Press.

Lipkova, J., R. J. Chen, B. Chen, et al. 2022. "Artificial Intelligence for Multimodal Data Integration in Oncology." *Cancer Cell* 40, no. 10: 1095–1110.

Longoni, C., A. Bonezzi, and C. K. Morewedge. 2019. "Resistance to Medical Artificial Intelligence." *Journal of Consumer Research* 46, no. 4: 629–650.

Luo, X., M. S. Qin, Z. Fang, and Z. Qu. 2021. "Artificial Intelligence Coaches for Sales Agents: Caveats and Solutions." *Journal of Marketing* 85, no. 2: 14–32.

McBride, N. 2005. "Chaos Theory as a Model for Interpreting Information Systems in Organizations." *Information Systems Journal* 15, no. 3: 233–254.

McCarthy, J., M. L. Minsky, N. Rochester, and C. E. Shannon. 1955. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence." *AI Magazine* 27, no. 4: 12–14.

McKelvey, B. 2001. "Energising Order-Creating Networks of Distributed Intelligence: Improving the Corporate Brain." *International Journal of Innovation Management* 5, no. 2: 181–212.

McKelvey, B. 2004. "Toward a 0'th Law of Thermodynamics: Order-Creation Complexity Dynamics From Physics and Biology to Bioeconomics." *Journal of Bioeconomics* 6, no. 1: 65–96.

Morse, L., M. H. M. Teodorescu, Y. Awwad, and G. C. Kane. 2021. "Do the Ends Justify the Means? Variation in the Distributive and Procedural Fairness of Machine Learning Algorithms." *Journal of Business Ethics* 181: 1083–1095. https://doi.org/10.1007/s10551-021-04939-5.

Munoko, I., H. L. Brown-Liburd, and M. Vasarhelyi. 2020. "The Ethical Implications of Using Artificial Intelligence in Auditing." *Journal of Business Ethics* 167, no. 2: 209–234.

Nan, N. 2011. "Capturing Bottom-Up Information Technology Use Processes: A Complex Adaptive Systems Model." *MIS Quarterly* 35, no. 2: 505–532.

O'Hear, S. 2017. "Babylon Health Partners With UK's NHS to Replace Telephone Helpline With AI-Powered Chatbot." TechCrunch. https://social.techcrunch.com/2017/01/04/babylon-health-partners-with-uks-nhs-to-replace-telephone-helpline-with-ai-powered-chatbot/.

Orlikowski, W. J. 2000. "Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations." *Organization Science* 11, no. 4: 404–428.

Patton, M. Q. 2005. *Qualitative Research*. UK: John Wiley & Sons.

Pentland, B. T. 1999. "Building Process Theory With Narrative: From Description to Explanation." *Academy of Management Review* 24, no. 4: 711–724.

Pentland, B. T., M. S. Feldman, M. C. Becker, and P. Liu. 2012. "Dynamics of Organizational Routines: A Generative Model." *Journal of Management Studies* 49, no. 8: 1484–1508.

Pentland, B. T., P. Liu, W. Kremser, and T. Haerem. 2020. "The Dynamics of Drift in Digitized Processes." *MIS Quarterly* 44, no. 1: 19–47.

Polese, F., A. Payne, P. Frow, D. Sarno, and S. Nenonen. 2021. "Emergence and Phase Transitions in Service Ecosystems." *Journal of Business Research* 127: 25–34.

Puntoni, S., R. W. Reczek, M. Giesler, and S. Botti. 2021. "Consumers and Artificial Intelligence: An Experiential Perspective." *Journal of Marketing* 85, no. 1: 131–151.

Putnam, L. L. 2015. "Unpacking the Dialectic: Alternative Views on the Discourse–Materiality Relationship." *Journal of Management Studies* 52, no. 5: 706–716.

Ragin, C. C. 1992. ""Casing" and the Process of Social Inquiry." In *What Is a Case? Exploring the Foundations of Social Inquiry*, edited by C. C. Ragin and H. S. Becker, 217–226. UK: Cambridge University Press.

Rai, A. 2020. "Explainable AI: From Black Box to Glass Box." *Journal of the Academy of Marketing Science* 48, no. 1: 137–141.

Rai, A., P. Constantinides, and S. Sarker. 2019. "Next Generation Digital Platforms: Toward Human-AI Hybrids." *MIS Quarterly* 43, no. 1: iii–ix.

Raisch, S., and S. Krakowski. 2021. "Artificial Intelligence and Management: The Automation-Augmentation Paradox." *Academy of Management Review* 46, no. 1: 192–210.

Rinta-Kahila, T., E. Penttinen, A. Salovaara, and W. Soliman. 2018. "Consequences of Discontinuing Knowledge Work Automation-Surfacing of Deskilling Effects and Methods of Recovery." In *Proceedings of the 51st Hawaii International Conference on System Sciences*, 5244–5253.

Rudolph, J., N. Fink, J. Dinkel, et al. 2021. "Interpretation of Thoracic Radiography Shows Large Discrepancies Depending on the Qualification of the Physician—Quantitative Evaluation of Interobserver Agreement in a Representative Emergency Department Scenario." *Diagnostics* 11, no. 10: 1868.

Rudolph, J., C. Huemmer, F.-C. Ghesu, et al. 2022. "Artificial Intelligence in Chest Radiography Reporting Accuracy: Added Clinical Value in the Emergency Unit Setting Without 24/7 Radiology Coverage." *Investigative Radiology* 57, no. 2: 90–98.

Samtani, S., H. Zhu, B. Padmanabhan, Y. Chai, H. Chen, and J. F. Nunamaker Jr. 2023. "Deep Learning for Information Systems Research." *Journal of Management Information Systems* 40, no. 1: 271–301.

Saunders, W., G. Sastry, A. Stuhlmueller, and O. Evans. 2017. "Trial Without Error: Towards Safe Reinforcement Learning via Human Intervention." In *AAMAS'18: Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, 2067–2069.

Schultze, U., and D. E. Leidner. 2002. "Studying Knowledge Management in Information Systems Research: Discourses and Theoretical Assumptions." *MIS Quarterly* 26, no. 3: 213–242.

Shah, A. R., J. N. Adkins, D. J. Baxter, et al. 2010. "Chapter 1—Applications in Data-Intensive Computing." *Advances in Computers* 79: 1–70.

Shao, Z., S. Yuan, and Y. Wang. 2020. "Institutional Collaboration and Competition in Artificial Intelligence." *IEEE Access* 8: 69734–69741.

Shollo, A., K. Hopf, T. Thiess, and O. Müller. 2022. "Shifting ML Value Creation Mechanisms: A Process Model of ML Value Creation." *Journal of Strategic Information Systems* 31, no. 3: 101734.

Stankovic, J. A. 2015. "Foreword." In *Social Sensing—Building Reliable Systems on Unreliable Data*, edited by D. Wang, T. Abdelzaher, and L. Kaplan, xv–xvi. Amsterdam: Elsevier.

Strich, F., A. S. Mayer, and M. Fiedler. 2021. "What Do I Do in a World of Artificial Intelligence? Investigating the Impact of Substitutive Decision-Making AI Systems on Employees' Professional Role Identity." *Journal of the Association for Information Systems* 22, no. 2: 304–324.

Sturm, T., J. P. Gerlach, L. Pumplun, et al. 2021. "Coordinating Human and Machine Learning for Effective Organizational Learning." *MIS Quarterly* 45, no. 3: 1581–1602.

Talburt, J. R., and Y. Zhou. 2015. "Chapter 5—Update and Dispose Phases – Ongoing Data Stewardship." In *Entity Information Life Cycle for Big Data*, edited by J. R. Talburt and Y. Zhou, 65–87. Amsterdam: Morgan Kaufmann.

Tanriverdi, H., and S.-Y. Lim. 2017. "How to Survive and Thrive in Complex, Hypercompetitive, and Disruptive Ecosystems? The Roles of IS-Enabled Capabilities." In *International Conference on Information Systems*, Seoul.

Ten Cate, O., and S. Durning. 2007. "Dimensions and Psychology of Peer Teaching in Medical Education." *Medical Teacher* 29, no. 6: 546–552.

Teodorescu, M., L. Morse, Y. Awwad, and G. Kane. 2021. "Failures of Fairness in Automation Require a Deeper Understanding of Human-ML Augmentation." *MIS Quarterly* 45, no. 3: 1483–1500.

Tomaszewski, J. E. 2021. "Overview of the Role of Artificial Intelligence in Pathology: The Computer as a Pathology Digital Assistant." In *Artificial Intelligence and Deep Learning in Pathology*, edited by S. Cohen, 237–262. Amsterdam: Elsevier.

Turing, A. M. 1950. "Computing Machinery and Intelligence." *Mind – A Quarterly Review of Psychology and Philosophy* LIX, no. 236: 433–460.

Urquhart, C., H. Lehmann, and M. D. Myers. 2010. "Putting the 'Theory' Back Into Grounded Theory: Guidelines for Grounded Theory Studies in Information Systems." *Information Systems Journal* 20, no. 4: 357–381.

van Beek, E. J., S. Mirsadraee, and J. T. Murchison. 2015. "Lung Cancer Screening: Computed Tomography or Chest Radiographs?" *World Journal of Radiology* 7, no. 8: 189–193.

Van de Ven, A. H., and M. S. Poole. 1995. "Explaining Development and Change in Organizations." *Academy of Management Review* 20, no. 3: 510–540.

Van den Broek, E., A. Sergeeva, and M. Huysman Vrije. 2021. "When the Machine Meets the Expert: An Ethnography of Developing AI for Hiring." *MIS Quarterly* 45, no. 3: 1557–1580.

Verganti, R., L. Vendraminelli, and M. Iansiti. 2020. "Innovation and Design in the Age of Artificial Intelligence." *Journal of Product Innovation Management* 37, no. 3: 212–227.

Vygotsky, L. S. 1980. *Mind in Society: The Development of Higher Psychological Processes*. Massachusetts: Harvard University Press.

Wagner, S. 2016. "Natural Language Processing Is no Free Lunch." In *Perspectives on Data Science for Software Engineering*, edited by T. Menzies, L. Williams, and T. Zimmermann, 175–179. Amsterdam: Morgan Kaufmann.

Waldrop, M. M. 1992. *Complexity: The Emerging Science at the Edge of Order and Chaos*. Vol. 45, 83. New York: Simon & Schuster.

Wang, D., A. Khosla, R. Gargeya, H. Irshad, and A. H. Beck. 2016. "Deep Learning for Identifying Metastatic Breast Cancer." ArXiv, abs/1606.05718.

Warkentin, M., A. C. Johnston, and J. Shropshire. 2011. "The Influence of the Informal Social Learning Environment on Information Privacy Policy Compliance Efficacy and Intention." *European Journal of Information Systems* 20, no. 3: 267–284.

Weick, K. E. 2007. "The Generative Properties of Richness." *Academy of Management Journal* 50, no. 1: 14–19.

Weiss, V., and R. Needlman. 1998. "To Teach Is to Learn Twice: Resident Teachers Learn More." *Archives of Pediatrics & Adolescent Medicine* 152, no. 2: 190–192.

Wilson, J., and P. R. Daugherty. 2018. "Collaborative Intelligence: Humans and AI Are Joining Forces." Harvard Business Review. https://hbr.org/2018/07/collaborative-intelligence-humans-and-ai-are-joining-forces.

Yoo, H., S. H. Lee, C. D. Arru, et al. 2021. "AI-Based Improvement in Lung Cancer Detection on Chest Radiographs: Results of a Multi-Reader Study in NLST Dataset." *European Radiology* 31, no. 12: 9664–9674.

Young, M. L., F. Y. Kuo, and M. D. Myers. 2012. "To Share or Not to Share: A Critical Research Perspective on Knowledge Management Systems." *European Journal of Information Systems* 21, no. 5: 496–511.

Zhang, S., N. Mehta, P. V. Singh, and K. Srinivasan. 2021. "Frontiers: Can an Artificial Intelligence Algorithm Mitigate Racial Economic Inequality? An Analysis in the Context of Airbnb." *Marketing Science* 40, no. 5: 813–820.

Zhang, Z., Y. Yoo, K. Lyytinen, and A. Lindberg. 2021. "The Unknowability of Autonomous Tools and the Liminal Experience of Their Use." *Information Systems Research* 32, no. 4: 1192–1213.

Zhou, J., and F. Chen. 2019. "Towards Trustworthy Human-AI Teaming Under Uncertainty." In *IJCAI 2019 Workshop on Explainable AI (XAI)*.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.