

Seven Elements of Phronesis: A framework for Understanding Judgment in Relation to Automated Decision-making

Dina Koutsikouri
Department of Applied IT
Gothenburg University
dina.koutsikouri@ait.gu.se

Lena Hylving
Department of Informatics
University of Oslo
lenaandr@ifi.uio.no

Susanne Lindberg
School of Information Technology
Halmstad University
susanne.lindberg@hh.se

Jonna Bornemark
Centre For Studies in Practical Knowledge
Södertörn University
jonna.bornemark@sh.se

Abstract

This conceptual paper aims to explore judgment in the context of automated decision-making systems (ADS). To achieve this, we adopt a modern version of Aristotle's notion of phronesis to understand judgment. We delineate seven elements of judgment which provide insights into what humans are better at, and what AI is better at in relation to automated decision-making. These elements are sources of knowledge that guide action including not-knowing, emotions, sensory perception, experience, intuition, episteme, and techne. Our analysis suggests that most of these attributes are not transferable to AI systems, because judgment in human decision-making requires the integration of all which involves considering the contextual and affective resources of phronesis, and the competence to make value judgments. The paper contributes to unpack human judgment capacities and what needs to be cultivated to achieve 'good' AI systems that serves humanity as well as guiding future information systems researchers to explore human-AI judgment further.

Keywords: Artificial intelligence, Judgment, Phronesis, Automated Decision-making.

1. Introduction

Although there are mixed experiences on the use of AI in relation to the replacement of humans, an emerging research agenda focuses on the potential to enhance human-AI collaboration by examining the human advantage rather than the superpowers of AI (Berente et al, 2021). The research is beginning to unpack what happens when human meets the machine in various disciplines including for example radiology (Gichoya et al, 2018; Lebovitz et al 2022), predictive

policing (Waardenburg et al, 2020), algorithmic matching (AI for hiring), and studies that examine the joint performance of human-AI to establish the merits and pitfalls of working with AI (Fügener et al, 2021). The insights derived from these empirical studies is that the professionals who work with machines will outperform that machine or human working. What has slipped under the radar of much academic attention is the role of human judgment to reinforce AI capability and vice-versa. It highlights the need to explore what is uniquely human (in relation to machines) to better understand how human and AI can enhance each others complementary strenghts, and propensity to 'act synergistically' (Jarrahi, 2018, p. 577).

For example in decision-making, AI is based on calculating capacity and rational logic (i.e. $1+1=2$), while humans are guided by judgment (e.g. aspects of not-knowing, sensibility, emotions, and experiential knowledge); asking questions such as: Is this reasonable? Are the categories right? What is important in this specific situation? To what are we blind? (Bornemark, 2018)

Indeed, automated decision-making has caught wide attention across disciplines including law, public and administration (Waardenburg et al, 2021). In organizations across sectors more and more people make decisions with the help of algorithms. We see, for example, that robots make decisions about financial assistance (Eubanks, 2018) and that parts of the employment service's assessment support have been automated (Gahl et al, 2020). Digital data can often give us a faster picture, and in many cases a more detailed understanding of what is going on and the required changes needed to improve a situation. The use of AI for decision-making can thus give a sense of efficiency, quality, and impartiality (Neuroth et al, 2000;

Vinichenko et al, 2020). However, studies (e.g., Hylving and Lindberg, 2021) show that challenges arise when we attach greater importance to the amount of data than to crucial details for creating solutions with quality. Over-reliance on large amounts of data over human judgment reflects an expectation that automated decision-making (ADS) allows for more objective decisions than those made by humans who may be influenced by prejudice, conflicts of interests or fatigue (Lepri et al, 2021).

As AI technologies are being rapidly deployed into the workplace it is therefore important to explore both human judgement and AI since autonomous systems lack the capabilities related to aspects of emotions, sensibility and subjectivity that are at stake in ‘acting wisely’ in a given situation, and which are the very core of decision-making (Irwin, 1999). This prospect raises fundamental philosophical issues about the role of human judgment (during decision making processes) in relation to the capability of AI in ADS. The boundary becomes blurred between what human judgement is, and what the purpose of this competence is now that technology such as ADS exists to make decisions.

In this paper, we draw on Aristotle’s perspective of phronesis we understand as a kind of judgment, to compare humans and automated decision making in terms of what humans ‘can do’, and what is seen to be mastered by both man and machine. The overarching aim is to characterize the sources of knowledge that are brought to bear in exercising judgment, and their ‘function’ in relation to the capacities displayed in AI-assisted technologies. For the purposes of achieving this, we pursue the following research question: *What are the elements of judgment that are (currently) not transferable to automated decision-making?* Our guiding assumption is that the human judgment is something different from how, where, and when “judgment” is applied in ADS.

We propose the contours of a framework that assists the comparison of the unique characteristics of automated systems and AI with human intelligence (HI) and verbalize their relations in the context of decision-making. By showing the resources of phronesis we contribute with a vocabulary to express what cannot be automated in decision-making, and that human involvement in an increasingly digitalized society needs to be cultivated to flourish and steer the direction of technology (Tegmark, 2017) with the purpose to make decisions. Unpacking phronesis lays bare that human judgment is different from AI, and that awareness of what is different can amplify both humans and machines in the pursuit of responsible AI (Dignum, 2019) so that ADSs serves decision-making for the greater good.

The paper is organized as follows: first, we provide a brief literature overview of AI and automated systems where automated decision-making represents the

technological agency in human-AI decision-making. We then describe the concept of phronesis, and the elements involved in carrying it out and how these elements are implicated (if, at all) in automated decision-making. Finally, we discuss our analysis in view of the guiding research question establishing what seems to be best handled by ADSs and humans respectively and present a tentative framework for understanding judgment in automated decision-making.

2. AI and Autonomous Systems

We witness today how Artificial Intelligence (AI) and autonomous systems are increasingly being deployed and entangled in our personal and professional life, from before we are born (Dias & Torkamani, 2019) until we die (Lu, 2019; Wiederhold, 2019) and beyond (Clark, 2014). They are intertwined in everyday activities ranging from dating and mating (Slater, 2013; Sumter, Vandenbosch, & Ligtenberg, 2017) to epidemiologic research (Mayer-Schönberger & Cuckier, 2013). AI is rapidly becoming a pervasive aspect of the present (Russell, 2019) and while it is impossible to predict exactly how it will develop or on what time line, it is clear that the technology that enables AI is continuously advancing as well as becoming increasingly affordable for the many (Fountain et al, 2019). How this shapes our world is yet unknown (Baskerville et al, 2020) however, important to explore because it is the dominant technology of the future (Russell 2019). As such, it is essential to explore and understand, to drive the frontier of technology and knowledge forward (Berente et al., 2021).

AI and autonomous systems are concepts that include an immense number of different types of technologies and systems. Here we focus on one specific type of AI system, namely automated decisions systems (ADS). ADS aim to aid or replace human decision making based on rules and statistical and mathematical algorithms, often in context-sensitive settings where details make a difference, and where the application of judgment and practical wisdom is critical (Kolbjørnsrud, Amico, & Robert, 2016).

We have seen many examples of when automated decisions have caused unfairness, inequality, and biased outcomes (O’Neal, 2016; Eubanks, 2018; Crawford, 2021). This causes not only suffering for the people who are affected by the decision made by the systems, but also a skepticism towards technology and automated decisions.

2.1 Automated decision-making systems (ADS)

ADS can take many forms and uses; at its most general level, ADS are algorithms that are used to collect,

process and model data to make decisions or recommendations for decisions, and then in turn use these decisions to improve the system itself (Araujo et al., 2020). The role of the human varies, from nonexistent to having full autonomy. For example, the range goes from fully automated ADS that only communicate the decision, to recommender systems that the user can ignore (Araujo et al., 2020). Pre-existing and emerging biases have been an unrelenting issue for ADS (Dobbe et al., 2018). There are advocates of always ensuring that human beings have agency in automated decision making, though it is not currently done (Wagner, 2019). Others instead focus on the process of designing and implementing ADS in a value-centric way (Dobbe et al., 2018). A third stream of research is instead focused on the governance perspective, arguing for the need of ethical auditing of ADS (Mökander & Axente, 2021).

The assumption is that ADS lead to more efficient, effective, and objective decisions; yet this is not always the case (Araujo et al., 2020). We have also seen systems that involve humans to “rubber stamp” automated decisions to avoid the necessary regulations placed on fully automated systems, which have led to misplaced liability and biased systems (Wagner, 2019; Mökande & Axente, 2021). Further, studies have shown that human knowledge can decrease, making humans act as “borgs”, lacking individuality (Fügener et al., 2021).

ADS can be designed and built to emulate different human capacities and capabilities (Hussain, 2018). It is considered to outperform humans in many domains (Boström, 2017). Yet, one of the problems with ADS is that even though they are supposed to emulate human capabilities, many building blocks are malfunctioning (Eubanks, 2018; O’Neil, 2016) or missing (Smith, 2019). Indeed, advocates of the human augmented AI-configuration highlight that not only does this need to be better understood, but that it may create new perspectives of the role of the human in relation to IT in organizations (Teodorescu et al., 2021).

Here we approach phronesis and ADS from the vantage point to not only discover how practical wisdom differs from these automated systems but also how they can work together and augment each other for social good.

The intensification of implementation and usage of digital technology, including ADS, have in many ways put the human in the background. For one, it has been considered that by getting rid of ‘the human factor’, better decisions and accurate predictions can be reached (ref). Yet, this has been proven wrong in many ways (Eubanks, 2018; O’Neal, 2016).

3. Conceptual background

In this paper, we draw on the Aristotelian view of phronesis and appropriate the concept in a way that is relevant to understand how practical wisdom is embodied in how practitioners arrive at professional judgment everyday decision-making situations.

When carrying out this analysis, we also rely on philosophical thinking based on prior work by Hannah Arendt’s and her portrayal of judgment, Edmund Husserl’s (2014) description of the meaning of ‘lived experience’, Nussbaum’s (2001) description of the relevance of emotions in cognition, and Merleau-Ponty’s (1974); illustration of the role of sensory perception in sense-making and empirical studies the theory of practical knowledge conducted to understand the components of phronesis and how these are integrated in ‘solving’ and intervening in practical situations. Although much of the analysis is couched in terms of the comparison between human intelligence and artificial intelligence this, the ultimate aim is to illuminate that judgment is a type of phronesis that motivates human rather than machine action, which we describe next.

3.1 Phronesis as human judgment

While digital technology is superior in some aspects to humans, such as visuospatial processing speed and pattern recognition, it lacks in terms of reasoning, creativity, and what Aristotle called phronesis, what we in current terms call practical wisdom (Jeste et al 2020; Kase, González-Cantón, & Nonaka, 2014).

The Aristotelian notion of phronesis reflects the aspects that people in practice situations would refer to as being ‘the reasonable thing to do’, in relation to the particulars of the specific situation (Shotter & Tsoukas, 2014). It is commonly translated into ‘wisdom’ but here we refer to it as embodied judgment which refers to the human capacity to draw from lived experiences as well sensibility (Bornemark, 2020).

In Arendt’s philosophy of judgment there are two sides, that sometimes have been read as contradictory (Arendt, 1992, Yar, 2000). On one hand she describes the person with judgement as an actor and on the other side as a spectator. But Marshall has pointed out that both these two sides are needed in a sound judgement. We need to be able to act, even if we don’t know everything in the situation, in that way the person with judgement is an actor, on the other side, precisely as we don’t know everything, we need to keep listening and be open to what we don’t know, i.e. keep a spectating position.

Phronesis can in relation to this be considered a kind of judgement that focuses on the ‘here and now’, and the concrete situation at hand. Thus, it is therefore often referred to as situational knowledge (Shotter & Tsoukas, 2014) In practice, we often ask ourselves what we are supposed to do in a particular situation, and in relation to the person we are interacting with. It is about bringing into the equation timing and knowing what and when to do something. It is closely associated to action and to act even when not every aspect of the situation is known.

It is closely connected to judgment, reflection, values, open-mindedness, and the pursuit of for the common good (Flyvbjerg, 2006; Nonaka & Toyama, 2007; Silvia Vaccarezza, 2018). Aristotle argued that phronesis is the most important knowledge type or ‘intellectual virtue’ for actors to possess (Ameriks & Clarke, 2000).

The action-oriented nature of phronesis makes it crucial to form sound ethical judgements (Begley, 2006). Phronesis is one’s capacity to direct action (Polansky, 2000) and is “a reasoned and true state of capacity to act with regard to human goods” (Aristotle, 2009). It begins from the apprehension of ‘what should be’ and whether a particular action should be done in a particular circumstance (Baggini and Fosl, 2007).

Aristotle asserts that the highest form of human well-being is the life controlled by reason (Beauchamp, 1991). Thus, sound judgment plays an important role in being wise, and it refers to both theoretical reasoning (the apprehension of what the truth is) and practical reasoning (the apprehension of whether a particular action should be done) as well as ethics and moral action.

Bornemark (2020) suggests phronesis may be the ‘most forgotten’ knowledge type/concept in contemporary organizations and society. She suggests that reliance on knowledge that can be generalized (episteme) and easily transferred is increasingly favored since it offers a promise to make our lives and enterprises more efficient and controllable.

3.2 Elements of human judgment

Calculating rationality describes how we turn to rules of abstractions and generalization to organize and make sense of our world. Bornemark’s (2018) argument is that too much of this type of rationality makes us loose contact with ourselves, others and the specific situation in a way that disables us to develop judgment. Instead, we rely on external parameters to objectively guide our action. Practices built on phronesis, on the contrary, emphasize the subjective, emotional, temporary and our ability to ‘not know’, but to learn to cope with insecurity, instability, anxiety and find ways to act in

such terrains (Shotter and Tsoukas, 2014). Bornemark (2018; 2020) posits that practices based on a calculating rationality, with rules and abstractions, and the opposite, where creativity and the not-known are in focus, are interdependent yet both are needed. But, in modern societies capacities connected to phronesis has been suppressed, overlooked, and seen as state of lack of better knowledge (Kristjánsson et al 2021).

This paper delineates seven key components of phronesis that shed light on the importance of human agency in automated decision-making. These components include *not knowing* (openness to possibilities), *emotions*, *sensory perception*, *experiences* (one’s own and others), *intuition* (seeing a solution without deliberate thinking), *episteme* and *techne*, and can be seen as fluid capacities that enables attention to what the situation requires rather than what can be defined, structured, organized, and systemized.

Human judgment, and to judge well requires a continuous movement between these components in an evaluative way, asking: what is important in this particular situation? How can the components be weighed in and what potential conflicts/paradoxes do they entail? Sometimes this happens through a conscious reflection process and sometimes it emerges intuitively (i.e., like a pre-language process). When decisions must be made quickly the people often arrive at judgment without conscious reflection regarding the particulars of the situation. However, neither of these thought processes are calculative, mathematical (rationalist orientation) since the capacity to evaluate and grasping important features are central in coming to a judgment (Shotter and Tsoukas, 2004). Such capacities are dependent upon an emotional and embodied sensing.

On this basis, now that organizations are faced with the challenge of developing ethical, responsible automated decision-making systems with embedded artificial intelligence, phronesis, or human judgment, as a knowledge form takes on a new urgency. Next, we take a closer look at the elements that we consider central in human judgment.

Not-knowing. When professionals make decisions the aspect of not-knowing is ever present (Souto, 2019). This form of knowledge relates to answering questions like: *Where does this problem (situation) begin and where does it end? What is important? What is at stake?* For example, when an injured person goes to an emergency room for medical care, it is the job of the nurse to gather information and assess the patient’s health information using evidence informed tools including posing questions to make an appropriate assessment. This includes considering that the situation also contains dimensions of not-knowing. It means that

while it is possible to consider information that is possible to capture from an initial analysis of the patient's health, there will always be aspects about a person's health status that are not possible to sort and gather. A skilled nurse therefore is acutely aware that what may seem a routine situation may well hide something more serious, and that it is important to remain sensitive to the possibility of not-knowing as part of deciding on patient health status. This entails the propensity to display humility and continuous listening.

Likewise, being able to discern *when* to take appropriate action even when facing an indeterminate situation (Shotter and Tsoukas, 2014). Coming to a judgment then involves consciousness about one's own finite ability to know all aspects of a situation and hence perception of phronesis.

In this way it relates to the wicked problems described by Rittel and Webber (1973). They posit that wicked problems regard such questions as where the situation begins and where it ends, what the ultimate goal is and what the central problem is (1973).

AI and ADS's do not have the capacity to relate to aspects of not-knowing as it cannot change its ultimate goals, and thus cannot emulate what is involved in this attribute of phronesis. For automated decision-making to function, the system requires data, known data, and it requires already set goals (Russell, 2019). An ADS can either use pre-organized data or learn from previously used data in decision-making.

Emotions. Another central component of phronesis are emotions. Emotions routinely affect how and what we see. Emotions are often perceived as subjective and therefore something that is not relevant as information for action in professional contexts (Schwartz & Sharpe, 2010). Further, emotions are frequently associated with personal opinions and concerns (Bornemark, 2018, 2020). Foregrounding emotions as a knowledge source acknowledges the non-optical factors of being in a situation. The flow of sensory perceptions stemming from our environment is not neutral information. It carries vital information on *what* is important in a situation, and therefore provides a motivating influence on what to do next. Schwartz and Sharpe (2010, p. 71) suggest that it is emotion that compels us to act. Hence, emotion alerts us to something that demands our attention, and that action is required. In the same vein, Nussbaum (2001) posits that emotions are part of a cognitive process and thus is about information processing. In a professional capacity the key is to attune to one's own as well as others' emotions and acknowledge them as relevant information carriers for potential actions.

Strong emotions can shut down the ability to being sensitive to context and responding to the situation as it

is now (Schwartz and Sharpe, 2010). Thus, the propensity to attune to but also relate to the layers of emotions in an impartial way is a prerequisite for developing judgment. Emotions are also necessary to make value judgements, emotions point out what is important, and what we consider good and bad. Emotions thus point out the direction in a lived experience (Bornemark 2018; Descartes 1989)

We argue in this paper that current forms of autonomous systems and AI are not alive and does not possess emotions and hence does not have the capacity for emotional involvement that is required for acquiring and developing the type of expertise displayed by for example 'caring' and 'wise' professionals, judges, police officers and teachers (e.g., Constantinescu and Crisp, 2022). But in phronesis emotion is a central component.

Sensory perception. One the knowledge sources that for example nurses at the emergency room (ER) heavily rely on is sensory perception. It is through the sensory organs, seeing, sensing, hearing, smell, and taste, we perceive information about a certain situation. Without our bodies we are unable to relate to the particulars of the situation (its uniqueness). However, while sensations help to collect what is important in a situation, they are much richer than we can fathom and verbalize; one sensation can carry infinite opportunities to distill valuable knowledge (Bornemark, 2020). The challenge is to remain open to the incoming sensations despite the influence of prior knowledge and prejudice. The Covid-pandemic brought to light the importance of sensory perception in terms of foregrounding the limitations of digital platforms in education. Teaching online emphasized listening while placing less importance on other sensory impressions compared to physical on-campus teaching where it included most sensory organs and richer sensations to be experienced. Sensory perception is not only a collection of data, but also intertwined with meaning and emotions and part of our sense-making, where not everything is of equal importance (Merleau-Ponty, 1974). According to Bornemark (2020) professional judgment also includes the capacity to take both once own and others sense perception into account, the sense perception of for example a colleague can be of crucial value to work out situations and knowing what goals to pursue.

Automated decision-making systems might record sound or temperature, but this is not the same thing as sense perception as it is disconnected from lived sense making and valuation. The introduction of automated decision systems in health care for example show that the capacity of AI is to build competence through data (from sensors), however it cannot mobilize resonance in

terms of forming meaningful connections with patients (and their social circumstances) (Lebovitz et al, 2021).

Lived experience. In hermeneutical phenomenology, phronesis and the capacity to act is closely intertwined with understanding. And Gadamer sets phronesis at the center of hermeneutical experience (Bobb, 2020, Gadamer 2004). We would also like to see it the other way around: A central element to phronetical knowledge is lived experience. And here we can build upon a Husserlian phenomenology in which the stream of lived experience is foundational (Husserl 2014). This knowledge source develops over time and experiencing ‘layers and layers of concrete situations’, increases the propensity for appropriate action. The attainment of phronesis then relies on experiences, and since each experience is unique, quantity matters. That is, opportunities for numerous experiences support the development of phronesis. Experience paves the way for dealing with horizons of not-knowing by including them in the equation rather than seeing uncertain elements as ‘noise’ to be avoided. To leverage on individual and collective experience in organizations, however, relies on professional knowledge exchange. Professional experience also present challenges in terms of upholding existing habits and behaviors which may not serve the greater good. An important part of cultivating this type of knowledge is to cultivate a culture of professional knowledge exchange to remedy to mitigate against biases and narrow perspectives taking root.

But learning from experience is not a process of generalization, it is not a mathematical exercise. Rather earlier experiences, once own and others, functions as a background material in relation to which the contemporary situation can be reflected and understood in a richer way (Bornemark, 2020).

The case of machine learning (ML) shows that AI is capable of learning from past decisions and to pick up patterns that human cognition does not see (eg., Gichoya, 2018; Lebovitz et al, 2021). But as machine learning is not a lived stream of experience it is not connected to emotions, not-knowing, valuations and “wicked problems.” Here AI and human judgement can complement each other (Jarrahi, 2018)).

Intuition. Within the realm of phronesis, the element of intuition denotes the unconscious knowledge process that also serves as a guiding light for professional conduct. In brief, this definition of intuition can be illustrated using the fictional example an experienced fisherman looking for fish in new waters: The fisherman takes his boat to the sea to find a good fishing spot. While at sea he feels the wind direction, and its velocity, tides, air, and water temperatures and from this information determining where to locate the boat to drop

the line. The fisherman arrives to this conclusion without conscious processing, rather the accumulated experience from prior pattern-recognition of the conditions of the sea (experience, sensory perception, emotions, and subjective knowledge) is at stake here. In this way, the fisherman accumulates his experience over time to be able to use intuition as way to inform his decision. In other words, humans acquire and develop their ability for pattern recognition based on a lot of different types of sensory data and previous experience that together make up intuition. Such knowledge has also been understood as a tacit knowledge, a knowledge that is contextual and embodied (Polanyi, 1966).

In the same vein, machines (e.g., ADSs) can be described to possess the capacity for ‘unconscious’ or rather a-conscious pattern recognition simply because of its capacity to process an immense amount of information. Indeed, ADS draws upon a gigantic amount of quantitative data to perform pattern-recognition (refs). We argue that intuition is in part implicated in the intelligent machines, however, it does not display phronesis. From a phronetic perspective, while human and machines display propensity for pattern recognition, human intuition captures far more aspects (and properties) of social context, emotions, and meaning than the machine can. This points to the potential of ADS to increase awareness regarding ‘new’ emerging patterns (e.g., issues, risks, opportunities) in the horizon that may influence outcomes.

Episteme. Epistemic knowledge is according to Aristotle not part of phronetic knowledge. But today theoretical knowledge and/or evidence-based materials need to be included in professional phronetic knowledge, for example of a doctor, a judge in court, or structural engineer in construction (Schawartz & Sharpe, 2010). Phronetic knowledge encompasses both episteme and techne (described next) and, applying ‘judgment’ helps to discern (and weighing up) whether to follow episteme or techne (i.e., following the prescriptive manual). When we sideline or forget phronetic knowledge and solely focus on epistemic knowledge we run the risk of promoting work practices and routines that lead to poor performance and quality outcomes for organisations and institutions. Episteme reflects general knowledge, and phronetic use of it entails discerning when and how to apply rules and principles that corresponds with details of each situation (e.g., emotions, subjective, experiential, and embodied knowledge). Overreliance on scientific knowledge and ‘what the computer says’ (Collins, 2018) present a danger that we lose sight of aspects of not knowing that are inherent in complex problems and situations in professional practice. Following this, the rationality underpinning AI is based on a kind of epistemic

knowledge, and this is where AI displays its strength and agency in supporting decision-making. Knowledge as represented in *techne* can be readily transferred to an automated systems to support decision-making.

Techne. *Techne* represents the collection of prior experiences in relation to producing something (Aristotle & Collins, 2011). It means that to become an expert one must develop many lived experiences. A baker improves by baking many breads; however, the baker cannot transfer the specific details of making the ‘the perfect loaf’. Thus, the lessons learned through baking many breads represents *techne* knowledge. In this context, *phronetic* knowledge is at stake when an apprentice is unable to read a bread recipe, and the baker subsequently responds to the situation in a manner that promotes confidence rather than shame. By understanding the role of *techne* in professional enterprises there is a possibility to broaden the view of what knowledge is. That it is more than *episteme* (scientific knowledge and general across contexts) and *techne* (knowledge that can be transferred and learned), but they are both components of judgement in professional enterprises.

In sum, while AI has the capacity to gather recipes and manuals (for how things can be executed) to provide recommendations for action and decisions, it lacks the capacity to also ‘feel’ the situation through sensory perception to come to a judgment.

Elements of phronesis	Human decision-making	Automated Decision-making
Not-knowing	Relating to and being aware of not knowing all aspects of a situation.	Uses only known data. Unable to change ultimate goals.
Emotions	Informs <i>what</i> is important in a situation, and therefore provides a motivating influence on what to do next.	Does not possess emotions and hence does not have the capacity for emotional involvement that is required for human-centered decisions.
Sensory perception	Remaining open to the incoming sensations despite the influence of prior knowledge and prejudice.	Disconnected from lived sense-making and cannot create meaningful connections
Lived experience	Using experiences (self and others) to mitigate against	Experiences based on already processed data, with or without

	biases and prejudice.	biases. Ability to ‘reflect’ on passed experiences based on solely previous data usage.
Intuition	Using of different types of sensory data and previous experience and can identify patterns that are of essence for a particular situation, even if it cannot be fully articulated.	Capacity to process immense amount of data to see patterns
Episteme	Theoretical knowledge that is used in a particular setting.	Knowledge that is represented in the data can be used for a specific purpose
Techne	Using prior lived experiences that are systematized.	Using manualized, standardized and/or formalized knowledge.

Table 1. Seven elements of phronesis in human and autonomous decision-making.

4. Discussion

This conceptual paper examines judgment in the context of automated decision-making systems (ADS). As such it both affirms the importance of judgment (*phronesis*) in ADS while revealing important differences between human decision-making and automated decision-making and the role of *phronesis* in contexts where decisions are made. In the subsequent paragraphs, we will highlight several implications resulting from this theoretical exploration.

The notion of judgment can be summarized as a synthesizing capacity in human action. This means that judgment is closely tied to action. Since it is a form of practical wisdom judgement is oftentimes not viewed as a source of knowledge. Most of the time we rely on theoretical and analytical knowledge since it appears to give us much more direction and sense of being in control. The downside then is that we sideline what we are not able to ‘name’ or what is often seen as subjective knowledge (e.g., experiences, emotions, intuition), but that may be equally important to guide decisions.

We also need to note that judgement in professional enterprises has a strong collective quality. It is something that is built up in a tradition, and that is in constant transformation. It includes the knowledge about who to turn to in a specific situation, and it include a continuous, and collective, relation to complicated ethical questions.

ADSs have pre-defined goals, although they continuously learn and improve based on used data and outcomes. However, the pre-defined goal is still its core and the fundamental logic rests on numbers and measurability. It is based on a one-dimensional rationality which phronesis transcends with its ability to deal with ambiguities, tensions, and differences in values and worldviews (Baltes & Staudinger, 2000). Consequently, the relation of human-AI configurations and phronesis is still unexplored yet has potential to achieve social impact both in short and long term. By better understanding the building blocks of human-AI configurations, we can also better understand how and why they can fail; the “cyborgism” effect described by Fügener et al. (2021) is one example of how human-AI configurations can fail. More importantly, researching Human-AI configurations with a phronetic lens offers possibilities to set the ground for human-centered AI development and usage.

Our tentative framework (Table 1) of phronetic judgment informed by Bornemark (2020) provides an initial basis for understanding the elements of decision-making judgment that we hope will stimulate a more nuanced appreciation of what humans can do in relation to ADS. It highlights that it is the role of the human to use judgment, rather than simply overseeing the ADS. While AI surpasses the human in terms of processing speed and pattern-recognition, it does not have the qualities involved in phronesis. Applying Bornemark’s (2020) conception and reasoning of phronetic professional judgment in relation to ADS, shows that the technology is not capable of making wise judgement. Phronesis synthesizes capacities including not-knowing, sensory perception, emotion, experiences, and intuition, *techne* and *episteme* to guide appropriate action. For one, this perspective of judgement provides us with a vocabulary to begin to better articulate and discuss the role of humans and how human intelligence can augment automated decisions and vice versa.

Given the width and depth of human judgement required for *virtuous and responsible* decision making, we argue that human involvement in automated decision making is not only essential but should consist of more than “rubber stamping” (Wagner, 2019). Humans should have agency in the decision-making process when the questions are of a wicked nature or has potential impact on human beings or society, even though that would have repercussions for the efficiency and objectivity of ADS. As this paper shows, AI does not have the potential to act wisely in complex situations, where human capabilities, such as phronesis, are required.

Although much of traditional research aims to improve efficiency, and control, especially in times of digitalization (Vinichenko, 2020), efficiency, in this

context, is associated with risk to society and human beings. Thus, there needs to be a balance between efficiency and phronesis, where only the human being can make up the fulcrum (Schwartz & Sharpe, 2010). Hence, there is a need for future research to unpack both the limitations of technology, as well as the role and capabilities of the human in this relationship.

While the advantages and capacities of ADS include following rules, collecting large amounts of data, i.e., knowledge that lend itself to being measured and categorized, resources of phronetic judgment include an array of elements that are relevant to seeing the whole situation and searching for the appropriate course for action. Judgment transcends what AI can do by interpreting and dealing with ambiguities, tensions, and differences in worldviews of stakeholders (Baltes and Staudinger, 2000).

Scholars in the field of AI and ADS reinforce the urgency to develop people-focused approaches in the design, use and implementation of automated systems and technologies. On a deeper level it is a call for the need to rehumanise automation and hence emphasising human agency in human-machine interaction, in current research termed human-in-the-loop (e.g., Grønsund and Aanestad, 2020). Against this background, the paper contributes to this research agenda by shedding light on the role of human judgment in relation to AI represented in automated decision-making.

We hope that this conceptual paper can contribute to the public debate on the critical issue raised by the development and use of AI namely, to better understand judgment as an important form of human intelligence now and in a possible future. In this regard, the paper could spur discussions to inform and empower understanding of what judgment *is* and its role in human-AI interaction and decision-making contributing to research and education on responsible AI.

Our analysis emphasizes the role of judgment alongside AI-based systems to help managers and other knowledge professionals rise above the limitations of calculating rationality and enact decisions that resonates with what the situation demands (Shotter & Tsoukas, 2014). It directs us to identify when it is better to rely on evidence-based knowledge, prescriptions, or manualized procedures and when we need to lean on experiential knowledge, sensory experience and, when we need to employ both sides of the knowledge spectrum. At the core, this project contributes with knowledge regarding the interaction of artificial intelligence and human judgment, foregrounding the argument that when humans interact with machines, we have new expanded agency (Suchman, 2007).

Finally, and importantly, there is a pressing need for more awareness of the ‘human black-box’, namely the illusion that we understand humans better than

algorithmic decision-making (Bonezzi et al, 2022), and its societal consequences. In other words, much attention is placed on explaining the algorithm, but the human remains unexplainable. Understanding human judgment and how humans draw on the resources of phronesis can contribute to greater trust in humans and human-AI decisions.

5. Conclusion

We suggest that understanding judgment (phronesis) is important to yield greater trust in human decision-makers as well as automated decision-making systems. Importantly, focusing on judgment reflects the role of humans and that this aspect of decision-making is a black box. Thus, being involved in the loop of AI goes beyond auditing and altering algorithms, it also requires phronesis. It emphasizes the capacity when it is better to rely on evidence-based knowledge, prescriptions and when new need to rely on intellect and peculiar human qualities such as not-knowing, sensory experience, emotions, intuition and when we need to combine these two rationalities (Bornemark, 2018; Lebovitz et al, 2021). Based on the insights from the paper we see an opportunity to conduct meaningful and relevant studies to empirically examine judgment in automated decision systems (ADS). Without understanding how people incorporate information from algorithms with phronetic judgment during decision-making processes, organizations run the risk of deploying automated decision-making systems in a way that reduces its true potential to support decisions in organizations. Finally, we argue that explainability of judgment in relation to automated decision-making systems is valuable to furthering research on human-AI decision-making and beyond.

6. References

- Ameriks, K., and Clarke, D. M. (2000). *Aristotle: Nicomachean Ethics*. Cambridge University Press.
- Araujo, T., Helberger, N., and Kruijkemeier, S. (2020) In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & Society*, 35, 611–623.
- Aristotle. 2009. *The Nicomachean Ethics* (Oxford World's classics). Translated by David Ross, Revised with an introduction and notes by Lesley Brown. Oxford University Press Oxford, UK.
- Aristotle, Bartlett, R. C., and Collins, S. D. (2011). *Aristotle's Nicomachean ethics*. University of Chicago Press.
- Beauchamp, T. L. (1991) *Philosophical ethics: an introduction to moral philosophy*, New York: McGraw-Hill.
- Begley, P.T. (2006), Self-knowledge, capacity, and sensitivity: Prerequisites to authentic leadership by school principals, *Journal of Educational Administration*, 44(6), 570-589.
- Berente N, Bin Gu, Recker J., and Santhanam R. 2021) Managing Artificial Intelligence. *MIS Quarterly*, 45(3):1433-1450.
- Bobb, C. V. (2020). The place of phronesis in philosophical hermeneutics. A brief overview and a critical question. *Hermeneia*, (25), 29-36.
- Bonezzi, A., Ostinelli, M., and Melzner, J. (2022) The human black-box: The illusion of understanding human better than algorithmic decision-making. *Journal of Experimental Psychology General*, 1-9.
- Bornemark, J. (2018) The limits of Ratio: An analysis of NPM in Sweden using Nicholas of Cusa's understanding of reason, in A. J. (Ed.) *Metric Culture: Ontologies of Self-tracking Practices*, (pp. 235-253). Emerald Publishing Limited.
- Bornemark, J. (2020) *Horisonten finns alltid kvar: Om det bortglömda omdömet*. Stockholm, Volante.
- Boström, N. (2017). *Superintelligence: paths, dangers, strategies*, (2nd ed.). Oxford University Press.
- Clark, L. (2014). *This creepy AI will talk to loved ones when you die and preserve your digital footprint*. Retrieved from <https://www.wired.co.uk/article/eterni-life-after-death-ai>
- Collins, H. (2018) *Artificial intelligence: Against humanity's surrender to computers*. Polity Press, Cambridge.
- Constantinescu, M., and Crisp, R. (2022). Can Robotic AI Systems Be Virtuous and Why Does This Matter?. *International Journal of Social Robotics*, 1-11.
- Crawford, K. (2021) *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- Davenport, T. H. (2018) *The AI advantage: How to put the artificial intelligence revolution at work*. Cambridge, MA: MIT Press.
- Descartes, R, 1596-1650. (1989) *The Passions of the Soul: An English Translation of Les Passions De L'Âme*, Indianapolis: Hackett Publishing Company.
- Dias, R., and Torkamani, A. (2019). Artificial intelligence in clinical and genomic diagnostics. *Genome medicine*, 11(1), 1-12.
- Dobbe, R., Dean, S., Gilbert, T., and Kohli, N. (2018). A broader view on bias in automated decision-making: Reflecting on epistemology and dynamics. *arXiv preprint arXiv:1807.00553*.
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Flyvbjerg, B. (2006). Matter: Power, values and phronesis. *Phronesis A Journal for Ancient Philosophy*, 370–387.
- Fountaine, T., McCarthy, B., and Saleh, T. (2019). Building the AI-Powered Organization. *Harvard Business Review*, (August).
- Fügener, A., Grahl, J., Gupta, A., and Ketter, W. 2021. "Will Humans-in-the-Loop Become Borgs? Merits and Pitfalls of Working with AI," *MIS Quarterly* (45:3), pp. 1527-1556.
- Gadamer, H. G. (2004). *Truth and method*, London: Continuum.

- Gal, U., Jensen, T. B., and Stein, M. K. (2020). Breaking the vicious cycle of algorithmic management: A virtue ethics approach to people analytics. *Information and Organization*, 30(2), 100-301.
- Husserl, Edmund (2014). *Ideas for a pure phenomenology and phenomenological philosophy. First book, General introduction to pure phenomenology*. Indianapolis/Cambridge: Hackett Publishing Company
- Gronlund, T., and Aanestad, M. (2020). Augmenting the algorithm: Emerging human-in-the-loop work configurations. *Journal of Strategic Information Systems*, 29, 101-614.
- Hussain, K. (2018). Artificial Intelligence and its Applications goal. *International Research Journal of Engineering and Technology*, 5(1), 838–841.
- Hylving, L., and Koutsikouri, D. (2020) Exploring phronesis in digital innovation, In: Proceedings of the 28th European Conference on Information Systems (ECIS), An Online AIS Conference, June 15-17, 2020, 2020.
- Hylving, L., and Lindberg, S. (2022). Ethical Dilemmas and Big Data: The Case of the Swedish Transport Administration. *International Journal of Knowledge Management (IJKM)*, 18(1), 1-16.
- Irwin, T. (1999). *Aristotle: Nichomachean Ethics* (Second). Hackett Publishing.
- Jarrahi, M. H. (2018) Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making, *Business Horizons*, 61(4), 577-586.
- Jeste, D. V., Lee, E. E., Palmer, B. W., and Treichler, E. B. H. (2020). Moving from humanities to sciences: A new model of wisdom fortified by sciences of neurobiology, medicine, and evolution. *Psychological Inquiry*, 31(2), 134–143
- Kase, K., González-Cantón, C., and Nonaka, I. (2014). *Phronesis and Quiddity in Management. A School of New Knowledge Approach*. London: Palgrave Macmillan.
- Lebovitz, S., Levina, N., and Lifshitz-Assaf, H. (2021). Is AI Ground Truth Really “True”? The Dangers of Training and Evaluating AI Tools Based on Experts’ Know-What. *MIS Quarterly*, 45(3), 1501-1525.
- Lebovitz, S., Lifshitz-Assaf and Levina, N. (2022) To Engage or Not with AI for Critical Judgments: How professionals Deal with Opacity When Using AI for Medical Diagnosis, *Organization Science*, 33(1), 126-146.
- Lepri, B., Oliver, N., and Pentland, A. (2021). Ethical machines: The human-centric use of artificial intelligence. *IScience*, 24(3), 102249.
- Lu, D. (2019). AI can predict if you’ll die soon – but we’ve no idea how it works. Retrieved from New Scientist website: <https://www.newscientist.com/article/2222907-ai-can-predict-if-youll-die-soon-but-weve-no-idea-how-it-works/#>
- Mayer-Schönberger, V., and Cuckier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.
- Merleau-Ponty, M. (1974) *Phenomenology of perception*. London: New York: Routledge & K. Paul; Humanities Press.
- Mökander, J., Morley, J., Taddeo, M. and Floridi, L. (2021) Ethics-Based Auditing of Automated Decision-Making Systems: Nature, Scope, and Limitations. *Science & Engineering Ethics* 27, 44.
- Neuroth, M., MacConnell, P., Stronach, F., and Vamplew, P. (2000). Improved modelling and control of oil and gas transport facility operations using artificial intelligence. In: *Applications and Innovations in Intelligent Systems VII* (pp. 119-136). Springer, London.
- Nonaka, I., and Toyama, R. (2007). Strategic management as distributed practical wisdom (phronesis). *Industrial and Corporate Change*, 16(3), 371–394.
- O’Neil, C. (2016). *Weapons of Math Destruction: How big data increases inequality and threatens democracy*. Crown.
- Polansky, R. M. (2000). " Phronesis" on tour: cultural adaptability of Aristotelian ethical notions. *Kennedy Institute of Ethics Journal*, 10(4), 323-336.
- Polanyi M. (1966). *The tacit dimension*, 1st ed, Doubleday.
- Rittel, H.W.J., and Webber, M.M. (1973) Dilemmas in a general theory of planning, *Policy Sci*, 4, 155-169.
- Russell, S. (2019) *Human compatible: Artificial intelligence and the problem of control*. New York: Penguin.
- Schwartz, B., and Sharpe, K. (2010) *Practical wisdom: The right way to do the right thing*. New York, Riverhead Books.
- Shoter, J., & Tsoukas, H. (2014). Performing phronesis: On the way to engaged judgment. *Management Learning*, 45(4), 377–396.
- Slater, D. (2013). *Love in the time of algorithms: What technology does to meeting and mating*. Penguin.
- Souto, P. C. N. (2019). Ontological not-knowing to contribute attaining practical wisdom: Insights from a not-knowing experience in ‘samba-de-gafieira’ dance to the value of being and responding from within our practical experience and practical knowledge. *Learning, Culture and Social Interaction*, 21, 48-69.
- Suchman, L. (2007) *Human-Machine reconfigurations: Plans and situated actions* (2nd ed.), St Louis, Missouri, Elsevier.
- Sumter, S. R., Vandenbosch, L., & Ligtenberg, L. (2017). Love me Tinder: Untangling emerging adults’ motivations for using the dating application Tinder. *Telematics and Informatics*, 34(1), 67–78.
- Tegmark, M. (2017) *Life 3.0: Being human in the age of artificial intelligence*, New York: Alfred A. Knopf.
- Teodorescu, M., Morse, L., Awwad, Y., and Kane, G. C. 2021. “Failures of Fairness in Automation Require a Deeper Understanding of Human–ML Augmentation,” *MIS Quarterly* (45:3), pp. 1483-1499.
- Silvia Vaccarezza, M. (2018). An Eye on Particulars with the End in Sight: An Account of Aristotelian Phronesis. *Metaphilosophy*, 49(3), 246–261.
- Wagner, B. (2019). Liable, but not in control? Ensuring meaningful human agency in automated decision-making systems. *Policy & Internet*, 11(1), 104–122.
- Wiederhold, B. (2019). Can Artificial Intelligence Predict the End of Life. And Do We Really Want to Know? *Cyberpsychology, Behaviour, and Social Networking*, 22(5), 297–299.