

Smart Tech is all Around us – Bridging Employee Vulnerability with Organizational Active Trust-Building

Antoinette Weibel^a , Simon Schafheitle^b  and
Lisa van der Werff^c 

^a*Institute for Work and Employment Research, University of St. Gallen;* ^b*Human Resource Management Group, Department of High Tech Business and Entrepreneurship, School of Behavioural, Management, and Social Sciences, University of Twente;* ^c*Irish Institute of Digital Business, DCU Business School, Dublin City University*

ABSTRACT Public and academic opinion remains divided regarding the benefits and pitfalls of datafication technology in organizations, particularly regarding their impact on employees. Taking a dual-process perspective on trust, we propose that datafication technology can create small, erratic surprises in the workplace that highlight employee vulnerability and increase employees' reliance on the systematic processing of trust. We argue that these surprises precipitate a phase in the employment relationship in which employees more actively weigh trust-related cues, and the employer should therefore engage in active trust management to protect and strengthen the relationship. Our paper develops a framework of symbolic and substantive strategies to guide organizations' active trust management efforts to (re-)create situational normality, root goodwill intentions, and enable a more balanced interdependence between the organization and its employees. We discuss the implications of our paper for reconciling competing narratives about the future of work and for developing an understanding of trust processes.

Keywords: active trust management, artificial intelligence, datafication, employee vulnerability, future of work, workplace surveillance

Address for reprints: Simon Schafheitle, Human Resource Management Group, Department of High Tech Business and Entrepreneurship, School of Behavioural, Management, and Social Sciences, University of Twente, Drienerlolaan 5, Enschede 7500AE, The Netherlands (s.d.schafheitle@utwente.nl).

This is an open access article under the terms of the [Creative Commons Attribution](#) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

In recent years, technological intensity in the workplace has dramatically increased (Cascio and Montealegre, 2016), inspiring a lively debate on this ongoing transformation and its possible merits and drawbacks (Flensburg and Lomborg, 2021). Organizational scholars Bodrožić and Adler (2022) argue that the interplay of business decisions and societal regulations will determine whether this technology will be used to serve or exploit humanity. In the field of leadership, De Cremer (2022) notes that we should be wary when introducing AI management-assistance systems in organizations, as doing so may harm responsible leadership practices and dehumanize employees. Meanwhile, scholars in human resource management (HRM) call for identifying a priori conditions that organizations must develop to guide the use of self-learning technology and prevent negative effects such as psychological contract breaches and untenable employee stress (e.g., Budhwar et al., 2022; Meijerink and Bondarouk, 2023). The scholarly discussion of the trade-off between the merits and drawbacks of datafication technology mirrors the ongoing societal debate, which ranges from utopian to dystopian perspectives. For example, in *Homo Deus*, Harari (2016) outlines a future in which humans succeed in gaining god-like abilities with the help of intelligent technology. In contrast, in her seminal work on surveillance capitalism, Zuboff (2019) outlines a technology-fuelled commodification of reality which, she argues, will lead to technology subjugating humanity.

In a nutshell, we can discern two dominant and somewhat contradictory narratives in this discourse. The first frames technological advancement as a catalyst for efficiency improvement that threatens employee well-being through work intensification and dehumanization (e.g., Leicht-Deobald et al., 2019). In contrast, the second narrative presents technology as a means of reducing human labour on repetitive tasks while increasing the meaningfulness and importance of human potential in organizations, thereby creating a more positive employee experience (e.g., Spreitzer et al., 2017). Our paper rejects such a deterministic perspective and outlines how the potentially negative effects of datafication technology can be mitigated or even reversed. Specifically, we argue that employees' trust in their employing organization can provide a mechanism for navigating the proliferation of such technology in a way that allows organizations to benefit from technological advancement and leads employees to feel empowered rather than threatened. Such positive effects depend on how organizations actively contribute to positive trust expectations and experiences in a workplace permeated by datafication technology. Our paper explores how employees' trust in their employer is potentially jeopardized by the use of datafication technology and how organizations can react by more actively managing core trust issues.

Datafication is the transformation of social interactions into numerical, statistically analysable data, which then serve as a resource for optimizing and transforming organizational processes and services to generate value (Murray et al., 2022). Datafication technology can existentially alter workplace experiences (Ball, 2021) and, as we will discuss in this paper, jeopardize employees' trust in their employer. Drawing on dual-process theories of social cognition (e.g., Hodgkinson and Sadler-Smith, 2018), we argue that datafication can create small, erratic surprises in the work environment, highlighting two types of employee vulnerability (Misztal, 2012). First, *discontinuity vulnerability* may be triggered because datafication creates uncertainty regarding the employer's expectations of employees in the 'new' workplace,

as well as how the employer will behave in the future. Second, *socio-emotional vulnerability* may become more salient because datafication technology can be used to turn humans into ‘cogs in the machine’, undermining employees’ dignity and increasing exploitation. This highlighted vulnerability motivates a more systematic processing of trust in which employees consciously search for and (re-)evaluate trust cues instead of relying on the more stable and often subconscious automatic trusting (Lewicki and Brinsfield, 2011).

Thus, we develop a framework of active trust management strategies for organizations to create a new narrative through symbolic strategies and enable employee–employer interest alignment through substantive strategies. To be effective, we argue that these active trust strategies must be tailored to match the core trust issues that employee vulnerability makes salient, specifically: (1) a lack of shared normality perceptions, (2) unclear goodwill intentions, and (3) pronounced employee dependency. Consequently, we delineate how complementary active trust strategies can be implemented as we integrate insights from the fields of technology design and the future of work to build a practical theory of active trust management.

Our paper makes three contributions to the literature. First, we contribute to the growing literature on the future of work and, in particular, the ongoing debate about the benefits and disadvantages of the datafication of work (e.g., Neufeind et al., 2019). We aim to reframe this debate by integrating insights from diverse literature and developing a framework for organizations to embrace the benefits of technology-permeated workplaces in an employee-centric manner. By responding proactively to heightened employee vulnerability through active trust management, organizations can not only preserve employees’ trust but also strengthen it as they embrace technological advancement. This theoretical contribution has significant practical implications as we provide concrete and actionable suggestions for organizations and highlight areas where scholars and practitioners could collaborate to further develop the evidence base.

We also contribute to the understanding of trust in the workplace in two significant ways. The first is our theorization of the role of vulnerability as an underexplored but central mechanism that has become increasingly crucial in debates concerning our understanding of trust. Specifically, while advancements in trust theory have demonstrated the potential for trust to be driven by systematic (Mayer et al., 1995) or automatic processes (Baer et al., 2018; McEvily, 2011), we have yet to resolve how and when individuals engage in automatic or systematic trusting (Dirks and de Jong, 2022). We offer one way to conceptualize the switch in emphasis from automatic to systematic trusting by distinguishing different types (and threats) of vulnerability. Second, we extend a small but growing body of literature that recognizes trust as an active process. Our model of active trust management extends previous work in this field (e.g., Child and Möllering, 2003; Gustafsson et al., 2021; Williams, 2007) to differentiate between symbolic and substantive strategies and further integrates theories on vulnerability (Misztal, 2012) to argue that effective active trust strategies should be tailored to the core trust issues created by context-specific employee vulnerabilities.

EMPLOYEES’ TRUST IN THE EMPLOYER

Trust is defined as a willingness to be vulnerable based on positive expectations of another party (Rousseau et al., 1998). In this definition, willingness is a volitional process

that involves a choice or decision to accept vulnerability in a relationship (Van der Werff et al., 2019). Thus, vulnerability refers to ‘vulnerability to each other’ (Heimer, 2001, p. 43) and the perceived risk that the other might take advantage of a situation of dependence and incomplete knowledge (Kramer, 2006). In employment relationships, vulnerability exists in both directions, though our primary focus in this paper is on the vulnerability of employees as the trustor and the employing organization as the trustee.

Trust in the employer is paramount for every employee, as they encounter dependencies and uncertainties regarding their employer that are often ambiguous and difficult to judge (Gustafsson et al., 2021; Weibel et al., 2016). In practical terms, employees’ wellbeing and advancement, in addition to their pay, promotion, and tenure, depend on whether their employer honours written and unwritten promises and establishes a system that enables employees to thrive (Bachmann, 2001). Employees render themselves vulnerable precisely by accepting that their wellbeing and advancement depend on the employer’s behaviour. Thus, previous trust scholars have portrayed employees as putting themselves at the mercy of their employer (Deutsch, 1958) and accepting the risk that they may be exploited (Tzafrir and Dolan, 2004).

Employee Trust Processing

Employees trusting their employer is a highly complex and context-dependent process, and we argue that perceptions of vulnerability and their interplay with trust must be framed as a dynamic process. We draw on parallel competitive models of dual processing (Hodgkinson and Sadler-Smith, 2018) and propose that the process of employees trusting their employer involves a passive ‘consumption’ of trust cues intertwined with a more active search for cues that the employee will not be exploited by their employer. While dual-process accounts of trust have been used previously (e.g., Baer et al., 2018), the interplay of these systems and the understanding of how we switch between them is still nascent. We propose that the relative dominance of the passive or active trusting system hinges on the perception of vulnerability central to the definition of trust itself. Although both processes are always present in an employment relationship, perceptions of vulnerability typically operate subconsciously and, therefore, trust is likely to be processed more automatically. In contrast, disruptions and ongoing small surprises in the context of a relationship may highlight vulnerability and motivate the more systematic processing of trust (Gustafsson et al., 2021).

We propose that trustors generally rely on automatic processing in established employment relationships for two reasons. First, trust in the employer is complex, as it depends on the characteristics of various organizational members as well as the properties of abstract systems (Weibel et al., 2016). Hence, we expect employees to cope with this complexity by predominantly trusting their employer in a more low-key, automated fashion rather than spending the time and considerable cognitive effort needed to rationally weigh all available information to make conscious trust decisions. The everyday monitoring of trust cues in this context is likely to be less effortful and more unconscious and automatic (Gavetti and Levinthal, 2000). Second, organizations are designed to offer a long-term narrative, and trust in the employer is developed through specific institutional arrangements, policies, and practices (Bachmann

and Inkpen, 2011). These institutions enable situational normality beliefs (McKnight et al., 1998) because they stabilize expectations in a more abstract entity such as ‘the employer’. We argue that this more enduring characteristic encourages automatic trust to become the default.

Employee Vulnerability in the Employment Relationship

Since vulnerability is central to trust, there is an increasing awareness that vulnerability must be analysed in greater depth to understand how the employee’s automatic and systematic processing of trust in the employer unfolds (Nienaber et al., 2015; Searle et al., 2017). To further conceptualize employees’ vulnerability toward their employer, we draw on Misztal (2012) to distinguish between two types of vulnerability in employment relationships: discontinuity vulnerability and socio-emotional vulnerability.

Discontinuity vulnerability reflects our inability to predict what the future holds. Employees cannot predict how their employment relationship will develop, as their knowledge about the future behaviour of their employer and the future working context will always be limited. For instance, new technological advances may create different skill requirements or change the organization’s interpersonal fabric. This type of vulnerability has been explored by Luhmann (1989), who understands trust primarily through its complexity-reduction function, which enables humans to navigate the world despite discontinuities. The interplay between familiarity and trust is at this conceptualization’s core: ‘trust has to be achieved within a familiar world, and changes may occur in the familiar features of the world which will have an impact on the possibility of developing trust in human relations’ (Luhmann, 2000, p. 94). Hence, changes in (dis-)continuity and (un-)familiarity will invariably reframe the trust situation.

Employees’ willingness to accept discontinuity vulnerability is linked to perceptions of situational normality and their need to rediscover the familiar in the unfamiliar (McKnight and Chervany, 2000). Renewed situational normality enables employees to form generalized, stable, and reliable expectations regarding their employer’s competency and goodwill (McKnight et al., 1998). Misztal (2012) also argues that discontinuity vulnerability can be mitigated by collectively renegotiating the background assumptions that shape familiarity and thus (re-)creating situational normality beliefs. In addition, she argues for making credible promises that actors, despite new and unknown opportunities, will still act in a trustworthy fashion. Based on this account, promises made by the employer must be credible and ‘institutionalized’ to some degree.

Socio-emotional vulnerability is linked to our ‘human condition’, as it unfolds ‘because we are dependent on others for our self-realization, and with this for the development of self-confidence, self-respect, and self-esteem’ (Misztal, 2012, p. 218). As human beings, employees depend on their employers for decent work; this means, at a minimum, not being exploited or harmed and being treated with dignity regardless of one’s organizational role (Braganza et al., 2020). However, in many workplaces today, the unwritten contract between employees and their employer extends far beyond this minimal condition, and employees also expect meaningful work, positive relationships (De Hauw and De Vos, 2010), and the fulfilment of their full range of socio-emotional needs.

In trust research, socio-emotional vulnerability is reflected in moral (Baier, 1986) and identification-based trust (Lewicki and Bunker, 1995). According to moral trust, employers are mandated to take care of their employees and embed genuine goodwill and integrity into their processes and practices precisely because employees are socio-emotionally (inter-)dependent on their employers. As such, demonstrating organizational goodwill and honouring the moral mandate to care and show integrity is important for employees' willingness to be socio-emotionally vulnerable. In addition, identification-based trust suggests that both employers and employees are united in the form of community and a social 'we' identity, which has been characterized as 'mutual understanding so that one can act for the other' (Lewicki et al., 2006, p. 1007). Identification-based trust thus captures far-ranging socio-emotional aspects of the employment relationship. At the very least, such trust implies that employees are willing to accept socio-emotional vulnerability based on perceptions of interdependence rather than dependence vis-à-vis their employer (Lewicki et al., 2006). Table I consolidates our distinction of employee vulnerabilities according to the core trust issues that they create.

THE RELATIONSHIP BETWEEN DATAFICATION TECHNOLOGY, TRUST PROCESSING, AND VULNERABILITY

Datafication Technology and Trust Processing

Datafication is increasingly ubiquitous, and most aspects of our daily lives can now be rendered into data and analysed with sophisticated technology (Gasser and Almeida, 2017). The impact of datafication technology unfolds in the workplace via two interrelated technological features: (1) augmented data gathering and (2) sophisticated data interpretation (Schafheitle et al., 2020).

Table I. Matching employee vulnerability with core trust issues

	<i>Employee vulnerability</i>		
	<i>Discontinuity vulnerability</i>	<i>Socio-emotional vulnerability</i>	
Core trust issues	(Re-)Creating Situational Normality	Demonstrating Goodwill Intentions	Addressing Imbalances in Dependence
Primary focus of active trust management	(Re-)Creating situational normality so that 'things seem in proper order', promising successful interaction, conveying feelings of personal ease, and mitigating possible hazards (e.g., Lewis and Weigert, 1985)	Credibly demonstrating the organization's moral obligation toward its employees as well as its concern for protecting their interests, even at the expense of its own (e.g., Weibel et al., 2016)	Installing interdependence between the organization and its employees, 'showing skin in the game' and actively raising the organization's vulnerability toward its employees (e.g., Cohen and Isaac, 2021)

Augmented data-gathering capacities allow for fine-grained employee tracking and digital tracing, even entering workplace spheres that were previously reserved for human sensory perception. For instance, these capacities are widely used in smart ID badges, wearable GPS, face-recognition CCTV cameras, and bio-radio-frequency ID chips. As such, datafication technology can be used to make almost every move of employees visible (Stanton and Stam, 2006), while it does not inherently make the employer's expectations more transparent to the employee. This asymmetry suggests a shift in the dependency balance of the employment relationship, such that the employee might become more dependent in the relationship with the employer. Furthermore, *sophisticated data-interpretation capacities* largely rely on unsupervised/reinforcement learning algorithms, which allow for rapid and automatic analyses and interpret vast volumes of employee data in a manner that was not previously feasible (Nilsson, 2014). Some procedures, for instance, evaluate employee sentiment with standardized scripts, leaving little room for idiosyncrasy and the interpretation of more fine-grained and context-sensitive aspects (for further discussion, see Sewell et al., 2012; Stanton and Stam, 2006). Other interpretation capacities are used to support HR decisions with limited human involvement (Parasuraman et al., 2000), reducing opportunities for leaders to implement systems that consider individual workers' circumstances (Skinner and Searle, 2011).

Taken together, we argue that these two datafication technology features and their likely impact on the employee–employer relationship will motivate the employee to shift to a more systematic and conscious consideration of the employer's trust-related cues. Datafication technology's (socio-)technological materializations destabilize automatic trust by delivering small and erratic surprises, signalling that 'things might not go as expected' or conveying the message, 'you are a mere cog in the machine for us'. Theoretically, we would expect these cues to trigger a change in emphasis for trust processing, whereby systematic processing would supplement automatic trust (Hodgkinson and Sadler-Smith, 2018; Searle et al., 2017). Empirical evidence suggests that even in low-stakes situations, information that is unexpected and inconsistent with prior beliefs can motivate people to use more arduous and systematic processing (Trémolière and De Neys, 2014). Ongoing employment relationships are far from low stakes, and misplaced trust can have considerable costs for an employee (Skinner et al., 2014). Thus, as vulnerability becomes more salient, employees will be motivated to be more deliberate and accurate in their trust decisions. Situations where accuracy is primed as being more important than time pressure encourage more reliance on deliberative and systematic processing (Guo et al., 2017).

Datafication Technology and Employee Vulnerability

For datafication technology to influence employee vulnerability, its technological capacities – that is, data gathering and interpretation – must be experienced by employees. Typically, these are experienced via three socio-technological materializations: (1) their opaque functioning, (2) their 'function creep' and dynamic nature, and (3) their prescriptive capabilities (Schafheitle et al., 2020). Below, we elaborate on these three materializations of datafication technology and explain how they highlight employee vulnerabilities.

Table II. Matching socio-technological materializations of datafication technology with employee vulnerability

<i>Materializations of datafication technology</i>	<i>Predominant vulnerability</i>	<i>Examples</i>
Opaque functioning	Discontinuity vulnerability	Food deliverers from gig economy firms report that their work is assigned to them by algorithms that gather and analyse their behaviour and performance in real-time. It remains unclear what personal data is used, how it is weighed and analysed, and how algorithms recognize or reward employees' work efforts with bonuses or maluses (Barratt et al., 2020)
Function creep dynamism	Discontinuity vulnerability	Las Vegas casinos employ pattern recognition algorithms in surveillance cameras to enhance overall safety via licence plate screenings, visitor flow tracking, and facial analyses to identify 'blocked' gamblers and wanted persons and to detect fraud and theft (Analytics Insight, 2021). Cameras also analyse card dealers' smiles as a part of the casino's overall performance metrics since smiling was said to increase gamblers' money-spending behaviour (Madrigal, 2013)
Prescriptive capacity	Socio-emotional vulnerability	Amazon holds two patents for smart wristbands that use prescriptive capabilities to forecast the optimal 'pick and pack' velocity of their stock workers. As a result of algorithmic performance forecasting, wristbands physically vibrate to make warehouse workers pick up items faster. Consequently, employees feel dehumanized (Walsh, 2019)

Table II displays the links between the technological manifestations and employee vulnerabilities and provides illustrative examples.

First, datafication technology is inherently *opaque* to employees as it often relies on proprietary and undisclosed algorithms (Orlikowski and Scott, 2014). Additionally, its perceived opacity is enhanced because it often operates 'intelligently' due to an ability to 'get better at what it does' simply through its application (Nilsson, 2014). Consequently, datafication technology presents itself as a black box even to those trained to understand it (Burrell, 2016). We propose that the opacity of datafication technology heightens employee discontinuity vulnerability because employees no longer understand cause and effect in areas that matter for the terms of their exchange relationship with their employer. For instance, employees may no longer understand how a datafication-based performance evaluation is composed and hence, may not be able to understand what is expected from them or what reward is offered by the employer. This ambiguity may be exacerbated as leaders can also fall short because the datafication technology's opacity deprives leaders of their 'mastery of observation' (Hirschi, 2017) and complicates their ability to communicate the employer's expectations to employees.

Second, datafication technology is often adjusted dynamically as new ways of use emerge through its ongoing application. This characteristic is termed *function creep*

(Ball, 2010) and is also characterized as incomplete informed employee consent (Sutrop and Laas-Mikko, 2012), incomplete knowledge of monitoring (Stanton and Weiss, 2003), and inappropriate use (Holton and Fuller, 2008). Function creep refers to how the self-learning nature of datafication technology leads to novel data-gathering or analysis opportunities that are subsequently exploited by organizations without sufficient communication and planning. Function creep further highlights employees' discontinuity vulnerability as it can quickly undermine any possible basis for clear expectations regarding reciprocal promises and obligations.

Third, datafication technology has descriptive, predictive, and prescriptive capabilities. As the most 'intelligent' aspect, *prescriptive capabilities* can even override human decision-making (Krumeich et al., 2016). This distinguishes their effect from descriptive and predictive capabilities, which merely assist human decision-making based on a 'static' assessment or forecast of future alternatives (Leicht-Deobald et al., 2019). Prescriptive capabilities potentially introduce dual structures in which employees may experience two 'supervisors' (the machine and the human) whose directives may appear contradictory, ambiguous, or less individualized (Schafheitle et al., 2021). Thus, prescriptive capabilities significantly impact the socio-emotional vulnerability of employees and have the greatest potential to render an employee a 'cog in the machine', dehumanizing the workplace. In this vein, prescriptive capabilities increase employees' socio-emotional vulnerabilities in at least two ways. First, they can negate employees' socio-emotional needs by expecting employees to act as reliable partners to machines or even subordinating them to machines (Giermindl et al., 2022; Stinglhamber et al., 2023). Second, the prescriptive capabilities of datafication technology can prevent human leaders from acting as intermediaries who protect employees' socio-emotional needs (De Cremer, 2022). While human leaders can theoretically mitigate dehumanization and de-contextualization, it is unclear whether they will do so in practice, as the balance of responsibilities between leaders and machines will also be realigned, and the leaders themselves may experience dehumanization (Farndale et al., 2011).

A THEORETICAL FRAMEWORK: ACTIVE TRUST MANAGEMENT AND VULNERABILITY

Using a dual-process perspective, we have argued that datafication technology heightens employee vulnerability and its salience and thereby triggers a stronger reliance on systematic trust processing. When employees are more actively and consciously seeking out and weighing trust cues, organizations must take a more active stance to manage employees' trust. Active trust management involves actively working to build or preserve trust and shifts the focus of the trust literature away from the trustor, as a relatively passive consumer of trustworthiness signals, to the trustee (Child and Möllering, 2003; Gustafsson et al., 2021). Active trust, therefore, is an interactive and co-created process where the trustee, here the employing organization, has a role to play in assisting the employee to navigate salient uncertainties and newly stressed vulnerabilities as they make decisions about their willingness to accept that vulnerability (see also Williams, 2007).

Extending the work of Gustafsson and colleagues (2021; see also Misztal, 2012), we theorize that active trust strategies should first be distinguished by whether they are symbolic or substantive (Ashforth and Gibbs, 1990; Frederiksen, 2014). Symbolic strategies

serve as representations of an employer's trust intent. Their function is to persuade employees, influence their trust expectations, and guide the day-to-day interpretation of trust inside the technology-permeated workplace. Substantive strategies facilitate and change employees' first-hand experience of trust with identifiable (inter-)actions and outcomes (Schnackenberg et al., 2019). Additionally, active trust strategies must target the different types of vulnerability inherent in the erratic surprises to the employment relationship inside technology-permeated workplaces.

The *symbolic* route to active trust involves the employer sending credible signals that employees' vulnerability will not be exploited. Symbolic strategies are linked to promoting the perception that the workplace is still a familiar context and the employer a trustworthy counterpart. As such, symbolic strategies always entail strong communication efforts (Van Der Merwe and Puth, 2014). Symbolic active trust strategies, such as 'cognitive bridging', which is understood as communicative 'efforts to connect past, present, and future' (Gustafsson et al., 2021, p. 9), can be found in previous work. We extend this work by delineating symbolic active trust management according to its possible use to address discontinuity and socio-emotional vulnerability. In the case of discontinuity vulnerability, communicative efforts must re-establish situational normality beliefs and credibly signal that the new workplace can be mastered in a similar fashion to the old one. Moreover, socio-emotional vulnerability involves two core issues of trust, which must be addressed independently to maintain employees' trust: the employer's goodwill intentions and as well as a narrative of a shared destiny and balanced interdependence.

The *substantive* route to active trust requires the employer to 'invest' in recalibrating the vulnerability of their employees (Shabana and Ravlin, 2016). Here, active trust management manifests in first-hand tangible acts and involves concrete and lasting changes to decision-making structures, processes, responsibilities, and role allocations (Ashforth and Gibbs, 1990). We propose that discontinuity vulnerability is substantially addressed when leaders act as flexible 'normality navigators' (Kellogg et al., 2020). This is enabled by training leaders to act as intermediaries between employees and their employers whilst navigating the new normality and transformation process. Socio-emotional vulnerability can be substantially addressed by allowing employees to participate in changes, that is, by investing in co-creating the use of datafication technology, as this empowers employees and supports positive work outcomes (Adler and Borys, 1996). Finally, power imbalances and changing dependencies can be redressed by formalizing employee voice in a permanent fashion (Desai, 2018).

Table III integrates symbolic and substantive active trust strategies with discontinuity and socio-emotional vulnerability. We argue that organizations wishing to manage employee-employer trust in an environment where datafication technology has proliferated should first consider the types of vulnerability that are salient and then enact symbolic and substantive strategies to maintain and build trust during disruptions.

A FRAMEWORK FOR PRACTICING ACTIVE TRUST MANAGEMENT IN A WORKPLACE PERMEATED BY DATAFICATION TECHNOLOGY

We expand on the cells of our 2×3 active trust management matrix (see Table III) with practical recommendations presented in Tables IV and V. In doing so, we strive

Table III. Integrating active trust management with employee vulnerability in technology-permeated workplaces

		Vulnerability		
		Discontinuity vulnerability	Socio-emotional vulnerability	
Active trust management strategy		Core Trust Issue 1: (Re-) Creating Situational Normality	Core Trust Issue 2: Rooting Employer Goodwill Intentions	Core Trust Issue 3: Installing Inter-Dependence
	<i>Symbolic Strategies: Creating a Credible Narrative</i>	Comprehensive communication regarding the new situation	Signalling employer goodwill to employees	Signalling that ‘we are all in the same boat’
	<i>Substantive Strategies: Interest Alignment</i>	Investments to build predictability in the workplace environment	Committing to and institutionalizing employee-friendly outcomes	Institutionalizing interdependence between employees and the employer

for three key goals. First, we aim to integrate literature from diverse, multidisciplinary scholarly fields to build new theoretical insights into how trust in the employer can be managed in the context of increasing technological advancements and their deployment in the workplace. Second, we aim to enhance the potential impact of our paper on organizational practice by providing actionable examples of strategies that managers and organizations can use to enact active trust management. Finally, we aim to inspire future management research in this field by highlighting where additional evidence-based insight is needed to guide future practice. We do not intend to provide an exhaustive list of actionable recommendations, mainly because of the rapid speed of technological advancement (Miao et al., 2017; Schafheitle et al., 2020). Instead, we offer an exemplary framework for understanding how active trust strategies can address employee vulnerabilities to support employment relationships in a workplace permeated by datafication.

Core Trust Issue 1: (Re-)Creating Situational Normality Beliefs

Situational normality beliefs are salient promoters of trust in the absence of complete knowledge (McKnight et al., 1998), as they set a scene in which ‘things seem in a proper order’, promising successful interaction, conveying feelings of personal ease, and mitigating possible hazards (Lewis and Weigert, 1985). In this vein, symbolic strategies of active trust include comprehensively communicating the new situation to employees. Moreover, organizations can make tangible, substantive investments to align both parties’ interests so that the new situation is experientially normal for employees. Through such substantive investments, organizations can enable first-hand experiences of situational normality that help employees feel that the new situation will not be detrimental.

Table IV. Practical recommendations for implementing symbolic active trust management in technology-permeated workplaces

Vulnerability			
Discontinuity vulnerability			
Symbolic strategies: Creating a credible narrative	Core Trust Issue 1: (Re-)Creating Situational Normality	Socio-emotional vulnerability	Core Trust Issue 3: Installing Inter-Dependence
	Summary	Core Trust Issue 2: Rooting Employer Goodwill Intentions	
	Create a context that promises successful interaction, conveys feelings of personal ease, and mitigates possible hazards.	Credibly demonstrate the organization's moral obligation toward employees as well as its concern for protecting their interests.	Signal interdependence between the organization and its employees.
	Recommendations: 1. Conduct digital literacy training for employees covering competencies in information and data literacy, communication and collaboration, digital content creation, and safety and problem-solving 2. Purchase technology with self-explanatory and justifiable designs to support: 2.1. Intuitive interpretation 2.2. Feedback options 2.3. Trusted third-party auditing/certification	Recommendations: 1. Deploy enabling technology design via: 1.1. Enabling employee participation 1.2. Offering learning-oriented feedback Offering some employee control over technology	Recommendations: 1. Enable a learning culture, by: 1.1. Creating a shared belief that errors in a technologically complex context are unavoidable but can be transformed into something positive 1.2. Accepting human errors 1.3. Normalizing learning from errors 1.4. Enabling trialability
Further sources for active trust management practitioners			
	On digital literacy training, see Nisa and Setiyawati (2019). On self-explanatory design, see Mohseni et al. (2021)	On deploying an enabling technology design, see Vorn and Combs (2022) and Leach et al. (2001)	On enabling a learning culture, see Frese and Keith (2015)

Table IV. (Continued)

Vulnerability		
Discontinuity vulnerability		
Socio-emotional vulnerability		
Inspiration for developing the evidence base		
<p>In general, there is not enough evidence on the use and effects of digital literacy training in the workplace (see Nikou et al., 2022). Thus, we pose the following questions for future research:</p> <ol style="list-style-type: none">1. Considering the self-learning nature of AI, what would continuous improvement of digital literacy look like?2. What are the salient long-term effects of digital literacy?3. Concerning (self-) explainable AI systems, how are purchasing decisions for or against such systems made?4. To what extent do the decision-makers' digital literacy and the organization's employment philosophy matter in such purchasing decisions?	<p>Deploying an enabling technology design has a long tradition in research focusing on less smart technology (see Ravid et al., 2020). Recently, research has gained momentum in the use of AI-based 'smart' technology and self-learning algorithms (e.g., Vorm and Miller, 2020). We delineate the following questions for developing the evidence base:</p> <ol style="list-style-type: none">1. What does effective prioritization of enabling design options look like, and does this differ with regard to different effectiveness criteria, such as trust, acceptance, or speaking up?2. What role do context and technological capabilities play in determining the threshold of 'enough' transparency?3. How do the needs and priorities of both employees and managers affect the use of enabling design options?4. What do employees' mental frames look like, and how do they influence an enabling design (Vorm and Miller, 2020)?	<p>Based on the recent review conducted by Frese and Keith (2015), the positive impact of a learning culture is predominantly investigated in a controlled setting and seldomly in an AI-based 'smart' technology context. Thus, we pose the following questions to guide future research:</p> <ol style="list-style-type: none">1. What is the optimal context for a positive learning culture to develop? In particular, what are the effects of leadership climate and performance management (Dahlin et al., 2018)?2. How is a positive learning culture related to a proactive and psychologically safe climate, and what is the contribution of each to the perception that 'we are all in the same boat' (Baer and Frese, 2003)?

Table V. Practical recommendations for implementing substantive active trust management in technology-permeated workplaces

Vulnerability			
	Discontinuity vulnerability		Socio-emotional vulnerability
Substantive strategies: interest alignment	Core Trust Issue 1: (Re-)Creating Situational Normality		Core Trust Issue 2: Rooting Employer Goodwill Intentions
	Summary		
	Make tangible, substantive investments to align the organization and employees' interests so that the new situation is experientially normal.		Make goodwill investments, so that 'we care' becomes rooted in decision-making structures and the overall strategy discourse.
	Recommendations:		Recommendations:
	1. Define new leadership roles that explicitly explain how leaders inter-mediate between technology and employees		1. Install formal employer-employee co-creation mechanisms, super users, and early adopters in all phases of the implementation process
	2. Develop leadership in the following areas:		2. Be sensitive to various organizational levels and deliberately iterate between more central and decentral units for either broadening or converging perspectives (Kellogg, 2022)
	2.1. Compassionate and emotionally competent leadership		
	2.2. Responsible leadership		
	Further sources for active trust management practitioners		
	On the relationship between AI and leadership, see De Gremer (2020)		On the various approaches to how employees can be formally included in the implementation of technology, see Auernhammer (2020). On investments for mitigating ethical threats, see Lucivero (2016)
		Core Trust Issue 3: Installing Inter-Dependence	
		Make long-lasting and tangible investments to substantially increase the organization's vulnerability toward its employees.	
		Recommendations:	
		1. Formally involve stakeholders in joint decision-making processes through:	
		1.1. Technology impact assessments	
		1.2. Issue-oriented common initiatives	
		1.3. Issue-driven councils	
		1.4. Permanent work councils and mission/ethics boards	
		On institutionalizing stakeholder voice, see Kujala et al. (2022) and Delgado et al. (2021)	

Table V. (Continued)

Vulnerability		
Discontinuity vulnerability	Socio-emotional vulnerability	
Inspiration for developing the evidence base		
Leadership training and development must be supported by a context that enables new forms of leadership to emerge. Thus, we delineate the following questions for developing the evidence base:	AI-based smart technology differs in its functioning logic and impact, which means that we do not know much about an effective enabling implementation in this new context. Thus, we delineate the following questions for developing the evidence base:	Organizational acts of showing vulnerability are barely addressed in the broader management literature. Based on the above reasoning in this paper, we delineate the following questions for developing the evidence base:
1. What is the impact of HR systems (e.g., the selection of leaders or incentives for leaders), employment philosophy, and leadership culture on leaders' behaviours (see e.g., Gurdjian et al., 2014; McCauley and Palus, 2021)?	1. What skills are needed to meaningfully guide and structure a formal co-creation process?	1. How can an ethical technological assessment that effectively balances (conflicting) views from developers, implementers, and users be implemented as an internal discourse within the organization (see Kiran et al., 2015)?
2. How do leaders make sense of their changing role requirements, especially in the automation/augmentation context (e.g., Raisch and Krakowski, 2021)?	2. What do effective strategies for harmonizing conflicting interests look like?	2. How can organizations organize stakeholder involvement, given the tension between idealized vision and practical constraints (Delgado et al., 2021), and ensure that stakeholders view their involvement as qualitatively valuable (Moratis and Brandt, 2017)?
3. How do leaders balance emerging dilemmas, such as those resulting from conflicting responsibilities vis-à-vis different stakeholder groups?	3. Is there 'too much of a good thing' when it comes to the design and implementation of co-creation?	3. How can vulnerability investments be meaningfully and unambiguously rooted in organizational routines and processes (Nishii et al., 2008)?

Symbolic (re-)creation of situational normality. Recently, scholars have begun to acknowledge digital literacy as a critical precondition for employees to interact with datafication technology meaningfully. Terminology to describe such efforts ranges from digital capability building (De Marco et al., 2019) and employees' ability to learn from trial and error (Beck and Wade, 2004) to training for technology-workflow integration (Kellogg et al., 2020). Digital literacy training imparts basic technical knowledge about technological functioning and enables employees to better understand how to navigate technology-intensive workplaces. Additionally, digital literacy training provides employees with self-management and interpersonal competencies that sensitize them to technology-induced challenges and support them in developing practical solutions to re-establish a sense of normality (Trenerry et al., 2021).

First, employees should be trained to use the datafication technology deployed in their organization. As Hsi (2007) shows, employees' initial skills and experiences using technology enable them to develop additional skills independently. As the data volume and analytical sophistication increase, employees should also be trained in basic computational and statistical skills. This will empower them to critically question technology-driven predictive or even prescriptive outputs (Leicht-Deobald et al., 2019) and ensure that employees can assess for themselves whether they need additional training and development (Scholtz, 2019). Second, soft skills should form an important part of the training. In a technology-permeated environment, employees likely need training to support their self-efficacy and resilience to cope with the, at least temporary, work intensification necessary to master their new environment (Colbert et al., 2016). The new situation and salient vulnerability can also jeopardize relationships. (Yee, 2014), for instance, describes how interacting with technology involves balancing the interests of peers, leaders, and technological limits. Hence, digital literacy training should provide guidance on how to maintain interpersonal bonds in this new context.

Furthermore, technology design options that increase visibility into how analyses are performed and why particular outcomes occur are important for credibly communicating situational normality, and organizations should pay particular attention to such options during the selection, purchase, and implementation phases. A large body of multidisciplinary literature has approached these design options from various perspectives. We draw on recent machine-learning and human-machine interaction scholarship (Biran and Cotton, 2017; Kulesza et al., 2015) to distinguish between the explainability and justifiability of technology. Design options related to technology's explainability disclose the process of data gathering and analysis to explain 'what has been done in what way' to arrive at an outcome. In contrast, justifiability explains 'why' the outcome has occurred. Ideally, technology design provides explainability and justifiability in a way that minimizes the discrepancy between one's mental model and the actual function represented (Kulesza et al., 2015). Explainability and justifiability can be achieved by organizations by (1) choosing datafication technology that is inherently interpretable and (2) relying on reactive and interactive designs.

Inherently interpretable datafication technology uses 'shallow rule-based' models that are interpretable by default because they guide interaction through a predefined set of explicit rules (Rudin et al., 2013). For instance, Si and Zhu (2013) have trained hierarchical compositional models in which the algorithmic identification of a cat

begins with whether it has a pair of ears, and only the next step concerns whether they are pointed or round. From an applied perspective, such rules can mimic variance-based regression techniques or Boolean models and are thus easier and more intuitive to understand (Wang et al., 2015). However, in some instances (e.g., unsupervised/reinforcement learning), the ‘black box’ character of the technology disallows proximal interpretation (Biran and Cotton, 2017), and explanatory information can only be disclosed externally. In such cases, reactive or interactive design options can be utilized with feedback mechanisms that increase explainability and justifiability via textual or spoken feedback, instant visualization (e.g., nomograms, prototypes, or explain vectors, Biran and Cotton, 2017), localization (i.e., comparison of counterfactuals), or by example (i.e., recalling and comparing previous outputs, Vellido, 2020). Alongside these feedback mechanisms, beliefs in situational normality can also be communicated through third-party certifications (Adler et al., 2018) and interactive design options that enable a two-way exchange between the user and the algorithm in the form of explanatory debugging (Kulesza et al., 2015).

Substantive (re-)creation of situational normality. In line with the broader technology management and leadership literature, we suggest (financial) investment in leaders to increase their capacity to attend to employee needs amidst datafication. This can be done in two ways: (1) by redefining leaders’ roles as ‘courageous intermediaries’ and ‘compassion agents’ and (2) by subsequently investing in leadership training and development.

First, building on the initial insights of Kellogg et al. (2020), we argue that leaders must fulfil their new role as intermediaries between technology-driven decisions and employees’ contextually grounded needs. Such intermediation is important, as employees need guidance in making sense of novel organizational routines and adjusting interpersonal relationships to maintain performance (De Cremer, 2020). Enabling consistency is an important part of the leadership role, especially in states of organizational ambiguity or uncertainty (Balogun and Johnson, 2005). Consistency becomes even more significant as decisions based on datafication technology optimized toward a single pre-defined goal are implemented within the greater organizational context where goals are not always congruent. For instance, datafication technology could predict that it is less expensive to fire an employee than to retrain them, but leaders might counteract such a decision if it contradicts organizational values concerning employee retention and development. Trade-offs between competing organizational goals are not easy to program into software when one goal does not always take precedence over another. The most accurate and efficient decision is not necessarily the best one, and exercising this situated judgement is a leader’s responsibility.

Second, algorithms are unable to model intuition, empathy, care, or social recognition (Lee et al., 2015), and datafication technology can – at least until employees have adjusted – create stressful situations that increase employees’ psychological load (Shih et al., 2013). We suggest that compassionate leadership, which involves noticing or even anticipating employees’ suffering, empathetically feeling employees’ pain, and performing actions intended to ease this suffering (Worline and Dutton, 2017), is crucial to recreating situational normality within the organization. Appropriate training should therefore

prepare leaders to offer emotional support, give employees time and flexibility, and provide resources for coping. In addition, responsible leadership is needed, based on the normative concept of employers 'being responsible for' employees and the need to discern conflicts of interest and seek a fair solution that avoids harmful consequences (Pless and Maak, 2011).

Core Trust Issue 2: Rooting Organizational Goodwill Intentions

We define organizational goodwill intentions as the means by which an organization credibly demonstrates its moral obligation toward employees as well as its concern for protecting their interests, even at the expense of its own (Weibel et al., 2016). In this context, symbolic strategies comprise a credible narrative emphasizing an employer's care as a prevailing organizing principle, true to the motto, 'if employees are doing well, then the company is doing well, too'. Practically, this is achieved via the enabling design of datafication technology (Adler and Borys, 1996) to ensure that employees are valued as human beings with creative skills and resources. In addition, organizations can also make substantive goodwill investments that root their duty of care in decision-making structures and the overall strategic discourse.

Symbolic rooting of organizational goodwill intentions. Datafication technology, designed and deployed in an enabling fashion, explicitly values employees' needs and interests as equivalent to business interests and efficiency gains (Adler and Borys, 1996). We organize our recommendations around the three key characteristics of enabling design: employee participation, instant feedback, and opt-in/opt-out mechanisms.

First, offering employees opportunities to participate should render the technology more attuned to the employees' needs and interests. This can be enabled by allowing employees to voice their feelings, opinions, and perceptions when interacting with the systems (Hoff and Bashir, 2015). For instance, the inclusion of a suggestion form that invites employees to indicate their preferences vis-à-vis technology was shown to be effective in an electronic performance management setting (Alge, 2001). Experimental evidence also suggests that the opportunity to review, challenge, and add to technology-generated performance data can be helpful (Zweig and Scott, 2007). Critically, the employee voice pathway should not be organizational lip service but should show employees how their opinions have been taken into account (Aiello and Kolb, 1995).

Second, datafication technology that enables employee learning is more likely to be seen as a signal of an organization's goodwill intent. The design of technology that provides employee feedback is of particular importance. Constructive, learning-oriented feedback entails a balance of quantitative and qualitative content, while an overemphasis on quantitative (i.e., obvious and easy-to-measure) performance indicators can have detrimental effects (Carayon, 1994). Evidence from electronic performance monitoring further suggests that constructive feedback should be specific and without threats (Baron, 1993), include positive reinforcement (McNall and Roch, 2009), and be regular and immediate so that employees can prepare themselves and form stable expectations regarding their learning (Holman et al., 2002). Feedback from technology is also likely

to be perceived as more enabling if it is not tied to performance appraisal or financial incentives (Jeske and Santuzzi, 2015) and if employee performance comparisons are used sparingly (Sapegina and Weibel, 2017).

Third, maintaining possibilities for employees to opt in or out of being the target of datafication technology adds to the enabling narrative. The wider management perspective variously termed this strategy ‘employee control over technology’ (Alder and Ambrose, 2005), ‘authorization of information disclosure’ (Ball, 2010), and, from a legal perspective, ‘informed employee consent’ (Wildhaber and Lohmann, 2018). Generally speaking, ongoing possibilities to opt in or out invite employees to experiment with datafication technology, deliberately evaluate various cause–effect scenarios, and critically assess how such technology is embedded in particular legal or cultural contexts. Enabling opt-out opportunities is straightforward in practical terms, but organizations must ensure that the separation of work and leisure is respected and that employees bear no financial or reputational harm from opting out (Douthitt and Aiello, 2001).

Substantive rooting of organizational goodwill intentions. Credible investments in establishing organizational goodwill intentions manifest in allowing employees to co-create datafication technology. This refers to a joint and formal process in which technological boundaries are negotiated and bargaining power and influence are shared between employees and the employer (Weibel, 2007). Co-creation extends beyond providing employees with opportunities to participate in the enabling gestalt of technology and represents a means of granting employee care in a formal and enforceable way to ensure that employees’ needs are considered. Possibilities for co-creation are part of a novel movement in technology and HRM scholarship (Kim et al., 2020) and have previously been implicitly addressed as an assurance of informational self-determination (Chen and Park, 2005), appreciation of employee contribution (Sewell and Barker, 2006), and employee control (Alge, 2001). Practically speaking, the co-creation of datafication technology can take the form of employee super users or early adopters but should begin before the implementation phase to align employee–employer interests before unintended negative effects become salient (Susser, 1988).

Core Trust Issue 3: Installing Interdependence

Perhaps the most salient trust issue to address when managing trust in technology-permeated workplaces is to install interdependence. To address this issue, employers must ‘show skin in the game’ and raise the organization’s vulnerability to their employees (Cohen and Isaac, 2021). When datafication technology changes power and dependence structures due to increased transparency (Bernstein, 2017), widened accessibility of information (von Krogh, 2018), and the erosion of organizational boundaries (Cascio and Montealegre, 2016), installing interdependencies represents a credible commitment to preserving reciprocal trust (Sako, 2006). A symbolic way of installing interdependence between the employer and employees is the creation of a narrative that ‘we are all in the same boat’. Eventually, the organization can also make long-lasting and tangible ‘leap-of-faith’ investments to substantially increase its vulnerability toward employees.

Symbolic installation of interdependence. A strong learning culture includes a set of salient and shared values, beliefs, and management practices related to open communication about errors, sharing relevant knowledge regarding such errors, and facilitating helping behaviours and measures to anticipate, detect, and handle errors efficiently (Van Dyck et al., 2005). Establishing a strong learning culture is relevant to a narrative of interdependence as it embraces the idea that the use of technology might not reach its fullest potential immediately and that allowing time for learning and adjusting is crucial. Additionally, a learning culture might also convey that ‘we are in the same boat’ with regard to being vulnerable vis-à-vis new technology. Practically, a strong learning culture manifests in encouraging employees to interact with technology deliberately and critically, and communicating that inconsistencies, problems, or errors should be seen as learning opportunities and will not be punished. On a daily basis, such a narrative will serve as a blueprint for employees to experiment and learn by doing when interacting with datafication technology (Reardon, 2010). Technology management literature argues for ‘trialability’, whereby innovation or new technology can be experimented with daily and without major hurdles (Vargas et al., 2018). A very recent approach to integrating trialability into corporate culture is by conducting rituals, such as so-called ‘f***-up-nights’ (Wolf, 2020), that provide both leaders and employees with an occasion to celebrate failures.

Substantive installation of interdependence. We propose that active trust strategies can include the institutionalization of broad stakeholder involvement, that is, a formal process of joint decision-making concerning the design, use, and implementation of datafication technology (Desai, 2018), to mitigate emerging power asymmetries and unequal dependencies. This renders the organization more vulnerable to employees in two ways. First, such procedures formally grant employees an active voice and the right to jointly define the boundaries of datafication technology to harmonize interests and co-create the rules of the game in a continuous dialogue. Second, they enable employees to unionize or join forces with supporting external stakeholders and broaden their area of competence to neighbouring technology-affected domains. These aspects recalibrate interdependence in the relationship (Kiran et al., 2015).

Another established form of balancing interdependence is the ‘technology impact assessment’, a broad, formalized practice intended to identify potential adverse effects of datafication technology on a variety of targets (Kiran et al., 2015, p. 7). Such an approach can take the form of a checklist, where a list of pre-defined, mainly ethical issues are assessed (IEEE, 2019). A more inclusive option involves a multistakeholder discussion regarding the possible impact of technology on work and issues such as (employee) identities, relations, norms, and values (see also Tursunbayeva et al., 2022). For our purpose, we are mainly interested in practices that seek to institutionalize the stakeholder voice in the form of partnerships, meaningful participation, and delegation (Friedman and Miles, 2006). The strongest form of voice exists when stakeholders are granted permanent positions in issue-driven councils or governance bodies. For example, Dundon et al. (2004) discuss several stakeholder institutions, ranging from bottom-up problem-solving groups and audit circles to voluntary but formal partnership agreements and joint committees or internal worker councils (see also Delgado et al., 2021; Moratis and Brandt, 2017).

DISCUSSION

Datafication technology is expected to change the workplace as we know it, enabling less biased and more fact-based decision-making, particularly pertaining to people management. Suppliers of such technology make strong claims about its potential performance-enhancing impact. If its potential is fully realized, datafication technology may be used to better allocate, control, and plan human resources. However, ongoing discourse about the future of work has produced two contradictory narratives regarding this technology. The first is focused on how new technological developments set the scene for the second machine age, enable a new ‘Taylorization’ of the workplace, and render humans ‘cogs in the organizational machine’ (Leicht-Deobald et al., 2019). A second, equally strong narrative emphasizes that digitalization is linked to positive changes in the experienced workplace and employment relations (Neufeind et al., 2019). In this article, we bring these two narratives together. We argue that datafication technology’s impact on employee experience and trust in their employer is dependent on how organizations acknowledge and actively respond to these changes.

Theoretical Contributions

Managing employees’ trust in their employer is a worthwhile endeavour in any organization (for an overview of diverse main, mediating and moderating effects of trust, see Fulmer and Gelfand, 2012; Weibel et al., 2016). This task is especially important in a technology-permeated workplace where extensive employee data is gathered and applied to HRM and workload management. In developing a framework for understanding how trust in the employer can be managed in this technology-permeated context, we make three important contributions to the literature. The first is a contribution to the literature on the future of work, while the remaining two relate to the trust literature.

As the more optimistic narrative regarding the future of work argues, organizations still depend on their employees’ ingenuity and willingness to contribute. However, even if datafication technology were deployed merely to augment employee productivity, trust would still be essential to avoid employee resistance (Kellogg et al., 2020) and fulfil ethical obligations to employees and society more generally. Our paper contributes to the future of work debate by integrating diverse literature to develop a framework of active trust management that includes a series of practices through which organizations can reap the benefits of a technology-permeated workplace while keeping employee interests central. We argue that efforts to address employee vulnerability can allow organizations not only to protect their relationships with employees but also to strengthen these relationships amidst ongoing uncertainty and change. In doing so, we build on the work of Child and Möllering (2003) and Gustafsson et al. (2021) to consider the role that both parties play in managing trust and to offer context-specific strategies for tackling trust issues in a technology-permeated workplace. In addition to providing guidance on effective organizational strategies, our framework aims to identify promising areas for practitioners and scholars to collaborate in building an evidence base. Bridging the science–practitioner gap is particularly critical in this

area, as organizational practices will need to continually adapt to remain effective given the rate of technological advancement.

Beyond our contribution to the future of work literature, our paper also makes two theoretical contributions to a nuanced understanding of trust. First, we extend previous trust theory to further explore vulnerability, a core but underexplored aspect of trust (Nienaber et al., 2015; Searle et al., 2017). Drawing on Misztal's (2012) work on vulnerability allows us to expand our current understanding of trust as a process of information processing and decision-making and contribute to the growing literature on active trust strategies. Although a small body of work has recognized the potential for trust to be automatic (e.g., Lewicki and Brinsfield, 2011; McEvily, 2011), the dominant models of trust (e.g., Mayer et al., 1995) portray trust as a more systematic weighing of trust cues, and the understanding of how and when these different trust processes operate is still poor (Dirks and de Jong, 2022). We argue that while both systems are likely to operate in tandem, a reliance on the systematic, conscious processing of trust is likely to be the exception rather than the norm, particularly concerning more abstract trust referents, such as an employing organization. We frame vulnerability as the key to understanding how trustors shift from automatic to active processing of trust cues. Considering the two types of vulnerability is also critical to tailor active trust strategies to the specific relational context rather than adopting a 'one-size-fits-all' approach to active trust management.

Our second contribution to the trust literature builds on this last point, as our framework extends previous work on active trust strategies to distinguish between symbolic and substantive actions in addressing core trust issues. Inherent in this distinction, as well as in the distinctions created by a focus on different types of vulnerability, is an implicit escalation of commitment or investment. Looking at Tables IV and V, the level of investment increases from the left of the table to the right, as well as when we move from symbolic (Table IV) to substantive strategies (Table V). The strategy requiring the most investment from an organization, and that which offers the greatest opportunity for managing trust actively, is installing interdependence by demonstrating employer vulnerability and thus levelling the playing field. The distinction between symbolic and substantive strategies has parallels in the process models of trust-building which refer to the transformation of the quality of trust across different relationship stages (Lewicki and Bunker, 1996). These models depict trust based on collective identity and a commitment to shared values as higher quality, though more difficult to achieve. Thus, our framework describes how active trust management could be escalated, leading to a continuous interplay between the symbolic and substantive strategies of active trust management geared toward the identified core issue of salient vulnerability.

Future Research Directions

In addition to the opportunities for practitioner-scholar collaboration presented in Tables IV and V, examining the introduction of datafication technology through the lens of trust opens new avenues for understanding the future of work. Related empirical work has tended to focus on trust in technology itself (e.g., Lee and See, 2004)

or, if directed at trust in the organization, the viewpoint of consumers (e.g., Hengstler et al., 2016). Our paper suggests a new direction for understanding the future of work in terms of the relationship between the employee and the employer. While our focus is on the employing organization as the referent of trust, we recognize a growing body of literature on the importance of multilevel and multi-referent trust (Gillespie et al., 2021). Trust in the employer exists within the wider context of trust in other key referents, including the datafication technology itself, institutional regulators that attempt to govern technology use, and the myriad developers, suppliers, and service providers who support the technology (van der Werff et al., 2021). The potential for cross-level influences within these multilevel trust systems (Lumineau and Schilke, 2018) suggests that future research ought to consider how features that influence and shape trust in technology interact with our central concern in this paper, trust in the employer. It is important to note that trust in other referents, including the technology itself, could reduce surprises and shorten the amount of time for which systematic trust processing is dominant or even necessary.

The technology acceptance model (King and He, 2006) and the unified theory of acceptance and use of technology (Venkatesh et al., 2003) identify a range of variables from the ease of use of technology to social influences which predict attitudes and behaviours related to technology use. Further research is needed to examine how employees' trust in their organization is affected by the acceptance of datafication technology and how trust influences employee resistance and employees' inclination to game datafication technology. We propose that studies of employee experiences of vulnerability during the introduction and deployment of datafication technology are likely to improve our understanding of the processes influencing the success or failure of such technology. Given the nascent nature of our understanding of employee vulnerability, qualitative research designs will be particularly useful at this stage (Edmondson and McManus, 2007), followed by repeated designs with short data intervals during potentially trust-hazardous phases.

Another important trust referent in the relationship between technology, employees, and the employer is the leader. In this paper, we have proposed that a substantive investment in leadership development is key to enabling a collective renegotiation of situational normality and thus preserving trust in the employer. However, further research should analyse the changing role of leaders in a technology-permeated workplace and its consequences for trust. For instance, Schafheitle et al. (2021) have argued that leaders in such workplaces face a 'two-leader situation', as prescriptive algorithms increasingly automate leadership functions. This new triadic arrangement may blur responsibilities and create contradictions between human leaders and algorithms regarding decision-making and power.

Finally, our argument that trust in the organization is often a predominantly automatic process introduces methodological challenges related to capturing some of these processes through traditional self-reporting approaches. To investigate automatic trust processes, we recommend that researchers consider a wide range of methodological approaches. Qualitatively, this may include ethnomethodological approaches, which have already been applied to some degree in the field of human-machine interaction studies (e.g., Dourish and Button, 1998). Furthermore, research studies may utilize

technological advancements to consider taking more physiological measurements in the field. We can also look to other fields of research where researchers have been grappling with the challenges of capturing subconscious or implicit processes for some time. In social psychology, for instance, debates around the most appropriate measures of implicit attitudes and heuristics, such as stereotypes, continue (O'Shea et al., 2016).

CONCLUSION

Datafication technology has become increasingly prevalent in the workplace and is disrupting management processes and employees' relationships with their employers. Competing narratives regarding the future of work reach differing conclusions regarding the potential for these changes to benefit or harm employees. Our paper integrates theories on organizational trust and vulnerability to argue that the extent to which datafication technology has positive or negative effects depends on whether active trust management strategies are in place during the introduction and deployment of such technology. We develop a framework that proposes that datafication technology tests trust in the employment relationship and heightens employees' vulnerability toward their employer. The small and erratic surprises generated by datafication technology motivate employees to switch to more systematic and active trust processing and require organizations to use active trust management strategies to preserve, maintain, and, where necessary, build trust. Based on this model, we present relevant and actionable recommendations for organizations to proactively manage trust, preserve the employment relationship, and allow all stakeholders to realize the potential benefits of technological advancement.

We conclude this paper by revisiting the global debate that inspired us: what impact will smart technology have on humanity in the future? Will it liberate and improve us, as Harari (2016) suggests, or subjugate us and turn us into mere pawns of surveillance capitalism, as Zuboff (2019) warns? We believe that the recent call of the Future of Life Institute (<https://futureoflife.org/open-letter/pause-giant-ai-experiments/>) to pause large-scale AI experiments is highly justified as it emphasizes the need for careful planning and management of the development of smart technology. In this paper, we have shown that the same level of care needs to be applied to the use of smart technology in companies. Only by rising to this challenge and developing active trust management strategies in sync with the introduction of smart technology can we hope for humans to flourish despite, or perhaps even because of this innovation.

REFERENCES

- Adler, P. S. and Borys, B. (1996). 'Two types of bureaucracy: Enabling and coercive'. *Administrative Science Quarterly*, **41**, 61–89.
- Adler, P., Falk, C., Friedler, S. A., Nix, T., Rybeck, G., Scheidegger, C., Smith, B. and Venkatasubramanian, S. (2018). 'Auditing black-box models for indirect influence'. *Knowledge and Information Systems*, **54**, 95–122.
- Aiello, J. R. and Kolb, K. J. (1995). 'Electronic performance monitoring and social context: Impact on productivity and stress'. *Journal of Applied Psychology*, **80**, 339–53.

- Alder, G. S. and Ambrose, M. L. (2005). 'An examination of the effect of computerized performance monitoring feedback on monitoring fairness, performance, and satisfaction'. *Organizational Behavior and Human Decision Processes*, **97**, 161–77.
- Alge, B. J. (2001). 'Effects of computer surveillance on perceptions of privacy and procedural justice'. *Journal of Applied Psychology*, **86**, 797–804.
- Analytics Insight. (2021). *The growing use of artificial intelligence in las vegas*. Available at <https://www.analyticsinsight.net/the-growing-use-of-artificial-intelligence-in-las-vegas/> (accessed 28 March 2023)
- Ashforth, B. E. and Gibbs, B. W. (1990). 'The double-edge of organizational legitimization'. *Organization Science*, **1**, 177–94.
- Auernhammer, J. (2020). Human-Centered AI: The Role of Human-Centered Design Research in the Development of AI. Paper presented at the Synergy – DRS International Conference 2020, Held online, 11–14 August.
- Bachmann, R. (2001). 'Trust, power and control in trans-organizational relations'. *Organization Studies*, **22**, 337–65.
- Bachmann, R. and Inkpen, A. C. (2011). 'Understanding institutional-based trust building processes in inter-organizational relationships'. *Organization Studies*, **32**, 281–301.
- Baer, M. and Frese, M. (2003). 'Innovation is not enough: Climates for initiative and psychological safety, process innovations, and firm performance'. *Journal of Organizational Behavior*, **24**, 45–68.
- Baer, M. D., van der Werff, L., Colquitt, J. A., Rodell, J. B., Zipay, K. P. and Buckley, F. (2018). 'Trusting the "look and feel": Situational normality, situational aesthetics, and the perceived trustworthiness of organizations'. *Academy of Management Journal*, **61**, 1718–40.
- Baier, A. (1986). 'Trust and antitrust'. *Ethics*, **96**, 231–60.
- Ball, K. (2010). 'Workplace surveillance: An overview'. *Labor History*, **51**, 87–106.
- Ball, K. (2021). *Electronic Monitoring and Surveillance in the Workplace. Literature Review and Policy Recommendations*. Luxembourg: Publications Office of the European Union.
- Balogun, J. and Johnson, G. (2005). 'From intended strategies to unintended outcomes: The impact of change recipient sensemaking'. *Organization Studies*, **26**, 1573–601.
- Baron, R. A. (1993). 'Criticism (informal negative feedback) as a source of perceived unfairness in organizations: Effects, mechanisms, and countermeasures'. In Cropanzano, R. (Ed), *Series in Applied Psychology. Justice in the Workplace: Approaching Fairness in Human Resource Management*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc, 155–70.
- Barratt, T., Veen, A. and Goods, C. (2020). 'How algorithms keep workers in the dark'. In *BBC*, 28 August. Available at <https://www.bbc.com/worklife/article/20200826-how-algorithms-keep-workers-in-the-dark> (accessed 15 May 2023).
- Beck, J. C. and Wade, M. (2004). *Got Game: How the Gamer Generation is Reshaping Business Forever*. Boston, MA: Harvard Business School.
- Bernstein, E. S. (2017). 'Making transparency transparent: The evolution of observation in management theory'. *Academy of Management Annals*, **11**, 217–66.
- Biran, O. and Cotton, C. (2017). *Explanation and Justification in Machine Learning: A Survey*. Paper presented at the IJCAI-17 Workshop on Explainable AI (XAI).
- Bodrožić, Z. and Adler, P. S. (2022). 'Alternative futures for the digital transformation: A macro-level schumpeterian perspective'. *Organization Science*, **33**, 105–25.
- Braganza, A., Chen, W., Canhoto, A. and Sap, S. (2020). 'Productive employment and decent work: The impact of AI adoption on psychological contracts, job engagement and employee trust'. *Journal of Business Research*, **131**, 485–94.
- Budhwar, P., Malik, A., De Silva, M. T. and Thevisuthan, P. (2022). 'Artificial intelligence—challenges and opportunities for international HRM: A review and research agenda'. *The International Journal of Human Resource Management*, **33**, 1065–97.
- Burrell, J. (2016). 'How the machine "thinks": Understanding opacity in machine learning algorithms'. *Big Data & Society*, **3**, 1–12.
- Carayon, P. (1994). 'Effects of electronic performance monitoring on job design and worker stress: Results of two studies'. *International Journal of Human-Computer Interaction*, **6**, 177–90.
- Cascio, W. F. and Montealegre, R. (2016). 'How technology is changing work and organizations'. *Annual Review of Organizational Psychology and Organizational Behavior*, **3**, 349–75.
- Chen, J. V. and Park, Y. (2005). 'The role of control and other factors in the electronic surveillance workplace'. *Journal of Information, Communication & Ethics in Society*, **3**, 79–91.
- Child, J. and Möllering, G. (2003). 'Contextual confidence and active trust development in the Chinese business environment'. *Organization Science*, **14**, 69–80.

- Cohen, M. A. and Isaac, M. S. (2021). 'Trust does beget trustworthiness and also begets trust in others'. *Social Psychology Quarterly*, **84**, 189–201.
- Colbert, A., Yee, N. and George, G. (2016). 'The digital workforce and the workplace of the future'. *Academy of Management Journal*, **59**, 731–39.
- Dahlin, K. B., Chuang, Y.-T. and Roulet, T. J. (2018). 'Opportunity, motivation, and ability to learn from failures and errors: Review, synthesis, and ways to move forward'. *Academy of Management Annals*, **12**, 252–77.
- De Cremer, D. (2020). *Leadership by Algorithm. Who Leads and Who Follows in the AI Era?* Hampshire: Harriman House.
- De Cremer, D. (2022). 'With AI entering organizations, responsible leadership may slip!'. *AI and Ethics*, **2**, 49–51.
- De Hauw, S. and De Vos, A. (2010). "'Millennials'" career perspective and psychological contract expectations: Does the recession lead to lowered expectations?. *Journal of Business and Psychology*, **25**, 293–302.
- De Marco, C. E., Chesbrough, H., Lessl, M. and Heroult, M. (2019). 'Being digital v. doing digital. Empowering employees to embrace a corporate digital transformation'. *Academy of Management Proceedings*, **2019**, 18752.
- Delgado, F., Yang, S., Madaio, M. and Yang, Q. (2021). 'Stakeholder participation in AI: Beyond "add diverse stakeholders and stir"'. *arXiv preprint arXiv:2111.01122*. Available at <https://arxiv.org/abs/2111.01122> (accessed 15 May 2023).
- Desai, V. M. (2018). 'Collaborative stakeholder engagement: An integration between theories of organizational legitimacy and learning'. *Academy of Management Journal*, **61**, 220–44.
- Deutsch, M. (1958). 'Trust and suspicion'. *Journal of Conflict Resolution*, **2**, 265–79.
- Dirks, K. T. and de Jong, B. (2022). 'Trust within the workplace: A review of two waves of research and a glimpse of the third'. *Annual Review of Organizational Psychology and Organizational Behavior*, **9**, 247–76.
- Dourish, P. and Button, G. (1998). 'On "technomethodology": Foundational relationships between ethnomethodology and system design'. *Human-Computer Interaction*, **13**, 395–432.
- Douthitt, E. A. and Aiello, J. R. (2001). 'The role of participation and control in the effects of computer monitoring on fairness perceptions, task satisfaction, and performance'. *Journal of Applied Psychology*, **86**, 867–74.
- Dundon, T., Wilkinson, A., Marchington, M. and Ackers, P. (2004). 'The meanings and purpose of employee voice'. *International Journal of Human Resource Management*, **15**, 1149–70.
- Edmondson, A. C. and McManus, S. E. (2007). 'Methodological fit in management field research'. *Academy of Management Review*, **32**, 1246–64.
- Farndale, E., Hope-Hailey, V. and Kelliher, C. (2011). 'High commitment performance management: The roles of justice and trust'. *Personnel Review*, **40**, 5–23.
- Flensburg, S. and Lomborg, S. (2021). 'Datafication research: Mapping the field for a future agenda'. *New Media & Society*, 1–19. <https://doi.org/10.1177/14614448211046616>
- Frederiksen, M. (2014). 'Calculativeness and trust: A view from phenomenology'. *Journal of Trust Research*, **4**, 34–43.
- Frese, M. and Keith, N. (2015). 'Action errors, error management, and learning in organizations'. *Annual Review of Psychology*, **66**, 661–87.
- Friedman, A. L. and Miles, S. (2006). *Stakeholders: Theory and Practice*. Oxford: Oxford University Press.
- Fulmer, C. A. and Gelfand, M. J. (2012). 'At what level (and in whom) we trust trust across multiple organizational levels'. *Journal of Management*, **38**, 1167–230.
- Gasser, U. and Almeida, V. A. (2017). 'A layered model for AI governance'. *IEEE Internet Computing*, **21**, 58–62.
- Gavetti, G. and Levinthal, D. (2000). 'Looking forward and looking backward: Cognitive and experiential search'. *Administrative Science Quarterly*, **45**, 113–37.
- Giermindl, L. M., Strich, F., Christ, O., Leicht-Deobald, U. and Redzepi, A. (2022). 'The dark sides of people analytics: Reviewing the perils for organisations and employees'. *European Journal of Information Systems*, **31**, 410–35.
- Gillespie, N., Fulmer, C. A. and Lewicki, R. J. (2021). *Understanding Trust in Organizations: A Multilevel Perspective*. New York: Routledge.
- Guo, L., Trueblood, J. S. and Diederich, A. (2017). 'Thinking fast increases framing effects in risky decision making'. *Psychological Science*, **28**, 530–43.
- Gurdjian, P., Halbeisen, T. and Lane, K. (2014). 'Why leadership-development programs fail'. *McKinsey Quarterly*, **1**, 121–6.
- Gustafsson, S., Gillespie, N., Searle, R., Hope Hailey, V. and Dietz, G. (2021). 'Preserving organizational trust during disruption'. *Organization Studies*, **42**, 1409–33.
- Harari, Y. N. (2016). *Homo Deus: A Brief History of Tomorrow*. London: Harvill Secker.
- Heimer, C. A. (2001). 'Solving the problem of trust'. In Cook, K. S. (Ed), *Trust in Society*. New York: The Russell Sage Foundation, 40–118.

- Hengstler, M., Enkel, E. and Duelli, S. (2016). 'Applied artificial intelligence and trust – The case of autonomous vehicles and medical assistance devices'. *Technological Forecasting & Social Change*, **105**, 105–20.
- Hirschi, C. (2017). 'Die automatisierung der angst. Wir brauchen mehr grundwissen in technologiegeschichte'. In *Frankfurter Allgemeine Zeitung*. Frankfurt am Main: Frankfurter Allgemeine Zeitung GmbH, 9.
- Hodgkinson, G. P. and Sadler-Smith, E. (2018). 'The dynamics of intuition and analysis in managerial and organizational decision making'. *Academy of Management Perspectives*, **32**, 473–92.
- Hoff, K. A. and Bashir, M. (2015). 'Trust in automation: Integrating empirical evidence on factors that influence trust'. *Human Factors*, **57**, 407–34.
- Holman, D., Chissick, C. and Totterdell, P. (2002). 'The effects of performance monitoring on emotional labor and well-being in call centers'. *Motivation and Emotion*, **26**, 57–81.
- Holton, C. F. and Fuller, R. M. (2008). 'Unintended consequences of electronic monitoring of instant messaging'. *IEEE Transactions on Professional Communication*, **51**, 381–95.
- Hsi, S. (2007). 'Conceptualizing learning from the everyday activities of digital kids'. *International Journal of Science Education*, **29**, 1509–29.
- IEEE. (2019). *Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems*. Available at https://standards.ieee.org/wp-content/uploads/import/documents/other/cad_v2.pdf (accessed 31 March 2023).
- Jeske, D. and Santuzzi, A. M. (2015). 'Monitoring what and how: Psychological implications of electronic performance monitoring'. *New Technology, Work and Employment*, **30**, 62–78.
- Kellogg, K. C. (2022). 'Local adaptation without work intensification: Experimentalist governance of digital technology for mutually beneficial role reconfiguration in organizations'. *Organization Science*, **33**, 571–99.
- Kellogg, K. C., Valentine, M. A. and Christin, A. (2020). 'Algorithms at work: The new contested terrain of control'. *Academy of Management Annals*, **14**, 366–410.
- Kim, S., Wang, Y. and Boon, C. (2020). 'Sixty years of research on technology and human resource management: Looking back and looking forward'. *Human Resource Management*, **60**, 229–47.
- King, W. R. and He, J. (2006). 'A meta-analysis of the technology acceptance model'. *Information & Management*, **43**, 740–55.
- Kiran, A. H., Oudshoorn, N. and Verbeek, P.-P. (2015). 'Beyond checklists: Toward an ethical-constructive technology assessment'. *Journal of Responsible Innovation*, **2**, 5–19.
- Kramer, R. M. (2006). *Organizational Trust: A Reader*. Oxford: Oxford University Press on Demand.
- von Krogh, G. (2018). 'Artificial intelligence in organizations: New opportunities for phenomenon-based theorizing'. *Academy of Management Discoveries*, **4**, 404–9.
- Krumeich, J., Werth, D. and Loos, P. (2016). 'Prescriptive control of business processes – New potentials through predictive analytics of big data in the process manufacturing industry'. *Business & Information Systems Engineering*, **58**, 261–80.
- Kujala, J., Sachs, S., Leinonen, H., Heikkinen, A. and Laude, D. (2022). 'Stakeholder engagement: Past, present, and future'. *Business & Society*, **61**, 1136–96.
- Kulesza, T., Burnett, M., Wong, W.-K. and Stumpf, S. (2015). *Principles of explanatory debugging to personalize interactive machine learning*. Paper presented at the Proceedings of the 20th International Conference on Intelligent User Interfaces, Interactive Machine Learning, Decision Making, Topic Modeling, and Robotics. Atlanta, GA: Association for Computing Machinery (ACM).
- Leach, D., Jackson, P. and Wall, T. (2001). 'Realizing the potential of empowerment: The impact of a feedback intervention on the performance of complex technology'. *Ergonomics*, **44**, 870–86.
- Lec, M. K., Kusbit, D., Metsky, E. and Dabbish, L. (2015). *Working with machines: The impact of algorithmic and data-driven management on human workers*. Paper presented at the Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.
- Lec, J. D. and See, K. A. (2004). 'Trust in automation: Designing for appropriate reliance'. *Human Factors*, **46**, 50–80.
- Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Schafheitle, S., Wildhaber, I. and Kasper, G. (2019). 'The challenges of algorithm-based HR decision-making for personal integrity'. *Journal of Business Ethics*, **160**, 377–92.
- Lewicki, R. J. and Brinsfield, C. (2011). 'Framing trust: Trust as a heuristic'. In Donohue, W. A., Rogan, R. G. and Kaufman, S. (Eds), *Framing Matters: Perspectives on Negotiation Research and Practice in Communication*. New York: Peter Lang Publishing, 110–35.
- Lewicki, R. J. and Bunker, B. B. (1995). 'Trust in relationships: A model of development and decline'. In Bunker, B. B. and Rubin, J. Z. (Eds), *Conflict, Cooperation, and Justice: Essays Inspired by the Work of Morton Deutsch*. San Francisco, CA: Jossey-Bass, 133–73.

- Lewicki, R. J. and Bunker, B. B. (1996). 'Developing and maintaining trust in work relationships'. In Kramer, R. M. and Tyler, T. R. (Eds), *Trust in Organizations: Frontiers of Theory and Research*. California: Sage Publications Inc, 114–39.
- Lewicki, R. J., Tomlinson, E. C. and Gillespie, N. (2006). 'Models of interpersonal trust development: Theoretical approaches, empirical evidence, and future directions'. *Journal of Management*, **32**, 991–1022.
- Lewis, J. D. and Weigert, A. (1985). 'Trust as a social reality'. *Social Forces*, **63**, 967–85.
- Lucivero, F. (2016). *Ethical Assessments of Emerging Technologies*. Cham: Springer.
- Luhmann, N. (1989). *Vertrauen: Ein mechanismus der reduktion sozialer komplexität*, 3rd ed. Stuttgart: Ferdinand Enke Verlag.
- Luhmann, N. (2000). 'Familiarity, confidence, trust: Problems and alternatives'. In Gambetta, D. (Ed), *Trust: Making and Breaking Cooperative Relations*. New York: Basil Blackwell, Inc., 94–107.
- Lumineau, F. and Schilke, O. (2018). 'Trust development across levels of analysis: An embedded-agency perspective'. *Journal of Trust Research*, **8**, 238–48.
- Madrigal, A. C. (2013). 'Your job their data: The most important untold story about the future'. *The Atlantic*. Washington, DC: Atlantic Monthly Group LLC, 21 November. Available at <https://www.theatlantic.com/technology/archive/2013/11/your-job-their-data-the-most-important-untold-story-about-the-future/281733/> (accessed 15 May 2023).
- Mayer, R. C., Davis, J. H. and Schoorman, F. D. (1995). 'An integrative model of organizational trust'. *Academy of Management Review*, **20**, 709–34.
- McCauley, C. D. and Palus, C. J. (2021). 'Developing the theory and practice of leadership development: A relational view'. *The Leadership Quarterly*, **32**, 101456.
- McEvily, B. (2011). 'Reorganizing the boundaries of trust: From discrete alternatives to hybrid forms'. *Organization Science*, **22**, 1266–76.
- McKnight, D. H. and Chervany, N. L. (2000). *What is Trust? A Conceptual Analysis and an Interdisciplinary Model*. Paper presented at the Americas Conference on Information Systems (AMCIS), Paper No. 382, <https://aisel.aisnet.org/amcis2000/382>.
- McKnight, D. H., Cummings, L. L. and Chervany, N. L. (1998). 'Initial trust formation in new organizational relationships'. *Academy of Management Review*, **23**, 473–90.
- McNall, L. A. and Roch, S. G. (2009). 'A social exchange model of employee reactions to electronic performance monitoring'. *Human Performance*, **22**, 204–24.
- Meijerink, J. and Bondarouk, T. (2023). 'The duality of algorithmic management: Toward a research agenda on HRM algorithms, autonomy and value creation'. *Human Resource Management Review*, **33**, 100876.
- Miao, H., Li, A., Davis, L. S. and Deshpande, A. (2017). *Towards Unified Data and Lifecycle Management for Deep Learning*. Paper presented at the 2017 IEEE 33rd International Conference on Data Engineering (ICDE).
- Misztal, B. A. (2012). 'Trust: Acceptance of, precaution against and cause of vulnerability'. In Sasaki, M. and Marsh, R. M. (Eds), *Trust. Comparative Perspectives*. Brill: Leiden, 358–79.
- Mohseni, S., Zarei, N. and Ragan, E. D. (2021). 'A multidisciplinary survey and framework for design and evaluation of explainable AI systems'. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, **11**, 1–45.
- Moratis, L. and Brandt, S. (2017). 'Corporate stakeholder responsiveness? Exploring the state and quality of gri-based stakeholder engagement disclosures of european firms'. *Corporate Social Responsibility and Environmental Management*, **24**, 312–25.
- Murray, A., Rhymer, J. and Sirmon, D. G. (2022). 'Humans and technology: Form of conjoined agency in organizations'. *Academy of Management Review*, **46**, 552–71.
- Neufeind, M., O'Reilly, J. and Ranft, F. (2019). 'Work in the digital age: Challenges of the fourth industrial revolution'. In Marengo, M. (Ed), *Partecipazione e conflitto*. London: Rowman & Littlefield, **12**, 970–9.
- Nienaber, A.-M., Hofeditz, M. and Romeike, P. D. (2015). 'Vulnerability and trust in leader-follower relationships'. *Personnel Review*, **44**, 567–91.
- Nikou, S., De Reuver, M. and Kanafi, M. M. (2022). 'Workplace literacy skills – How information and digital literacy affect adoption of digital technology'. *Journal of Documentation*, **78**, 371–91.
- Nilsson, N. J. (2014). *Principles of Artificial Intelligence*. Burlington, MA: Morgan Kaufmann Publishers.
- Nisa, A. and Setiyawati, D. (2019). *A Systematic Review of Digital Literacy Training for High School Students*. Paper presented at the Third International Conference on Sustainable Innovation 2019 – Humanity, Education and Social Sciences (IcoSIHESS 2019).

- Nishii, L. H., Lepak, D. P. and Schneider, B. (2008). 'Employee attributions of the "why" of HR practices: Their effects on employee attitudes and behaviors, and customer satisfaction'. *Personnel Psychology*, **61**, 503–45.
- Orlikowski, W. J. and Scott, S. V. (2014). 'What happens when evaluation goes online? Exploring apparatuses of valuation in the travel sector'. *Organization Science*, **25**, 868–91.
- O'Shea, B., Watson, D. G. and Brown, G. D. A. (2016). 'Measuring implicit attitudes: A positive framing bias flaw in the implicit relational assessment procedure (IRAP)'. *Psychological Assessment*, **28**, 158–70.
- Parasuraman, R., Sheridan, T. B. and Wickens, C. D. (2000). 'A model for types and levels of human interaction with automation'. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans*, **30**, 286–97.
- Pless, N. M. and Maak, T. (2011). 'Responsible leadership: Pathways to the future'. *Journal of Business Ethics*, **98**, 3–13.
- Raisch, S. and Krakowski, S. (2021). 'Artificial intelligence and management: The automation–augmentation paradox'. *Academy of Management Review*, **46**, 192–210.
- Ravid, D. M., Tomczak, D. L., White, J. C. and Behrend, T. S. (2020). 'Epm 20/20: A review, framework, and research agenda for electronic performance monitoring'. *Journal of Management*, **46**, 100–26.
- Reardon, R. F. (2010). 'The impact of learning culture on worker response to new technology'. *Journal of Workplace Learning*, **22**, 201–11.
- Rousseau, D. M., Sitkin, S. B., Burt, R. S. and Camerer, C. (1998). 'Not so different after all: A cross-discipline view of trust'. *Academy of Management Review*, **23**, 393–404.
- Rudin, C., Letham, B. and Madigan, D. (2013). 'Learning theory analysis for association rules and sequential event prediction'. *Journal of Machine Learning Research*, **14**, 3441–92.
- Sako, M. (2006). 'Does trust improve business performance'. In Kramer, R. M. (Ed), *Organisational Trust: A Reader*. Oxford: Oxford University Press, 267–94.
- Sapegina, A. and Weibel, A. (2017). 'The good, the not so bad, and the ugly of competitive human resource practices: A multidisciplinary conceptual framework'. *Group & Organization Management*, **42**, 707–47.
- Schafheitle, S., Weibel, A., Ebert, I., Kasper, G., Schank, C. and Leicht-Deobald, U. (2020). 'No stone left unturned? Towards a framework for the impact of datafication technologies on organizational control'. *Academy of Management Discoveries*, **6**, 455–87.
- Schafheitle, S., Weibel, A. and Rickert, A. (2021). 'The Bermuda triangle of leadership in the AI era? Emerging trust implications from "two-leader-situations" in the eyes of employees'. In *Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS 2020)*. Honolulu, HI: University of Hawaii at Manoa, Association for Information Systems IEEE Computer Society Press, 5473–82.
- Schnackenberg, A. K., Bundy, J., Coen, C. A. and Westphal, J. D. (2019). 'Capitalizing on categories of social construction: A review and integration of organizational research on symbolic management strategies'. *Academy of Management Annals*, **13**, 375–413.
- Scholtz, T. M. (2019). 'Big data and E-HRM'. In Thite, M. (Ed), *E-HRM: Digital Approaches, Directions and Applications*. Abingdon: Routledge, 77–88.
- Searle, R. H., Cooper, C., Den Hartog, D. N., Korsgaard, M. A., Lester, S. W., Moellering, G., Williams, M., Brower, H. H., Kautz, J. and Nienaber, A.-N. (2017). 'Exploring trust and change: New insights into vulnerabilities by exploring tipping points for trust'. *Academy of Management Proceedings*, **2017**, 11959.
- Sewell, G. and Barker, J. R. (2006). 'Coercion versus care: Using irony to make sense of organizational surveillance'. *Academy of Management Review*, **31**, 934–61.
- Sewell, G., Barker, J. R. and Nyberg, D. (2012). 'Working under intensive surveillance: When does "measuring everything that moves" become intolerable?'. *Human Relations*, **65**, 189–215.
- Shabana, K. M. and Ravlin, E. C. (2016). 'Corporate social responsibility reporting as substantive and symbolic behavior: A multilevel theoretical analysis'. *Business and Society Review*, **121**, 297–327.
- Shih, S.-P., Jiang, J. J., Klein, G. and Wang, E. (2013). 'Job burnout of the information technology worker: Work exhaustion, depersonalization, and personal accomplishment'. *Information & Management*, **50**, 582–9.
- Si, Z. and Zhu, S.-C. (2013). 'Learning and-or templates for object recognition and detection'. *IEEE transactions on pattern analysis and machine intelligence*, **35**, 2189–205.
- Skinner, D., Dietz, G. and Weibel, A. (2014). 'The dark side of trust: When trust becomes a "poisoned chalice"'. *Organization*, **21**, 206–24.

- Skinner, D. and Searle, R. H. (2011). 'Chapter 10: Trust in the context of performance appraisal'. In Searle, R. H. and Skinner, D. (Eds), *Trust and Human Resource Management*. Cheltenham: Edward Elgar Publishing, 177–200.
- Spreitzer, G. M., Cameron, L. and Garrett, L. (2017). 'Alternative work arrangements: Two images of the new world of work'. *Annual Review of Organizational Psychology and Organizational Behavior*, **4**, 473–99.
- Stanton, J. M. and Stam, K. R. (2006). *The Visible Employee: Using Workplace Monitoring and Surveillance to Protect Information Assets – Without Compromising Employee Privacy or Trust*. Medford, NJ: Information Today, Inc.
- Stanton, J. M. and Weiss, E. M. (2003). 'Organisational databases of personnel information: Contrasting the concerns of human resource managers and employees'. *Behaviour & Information Technology*, **22**, 291–304.
- Stinglhamber, F., Nguyen, N., Ohana, M., Lagios, C., Demoulin, S. and Maurage, P. (2023). 'For whom and why organizational dehumanization is linked to deviant behaviours'. *Journal of Occupational and Organizational Psychology*, **96**, 203–29.
- Susser, P. A. (1988). 'Electronic monitoring in the private sector – How closely should employers supervise their workers?'. *Employee Relations Law Journal*, **13**, 575–98.
- Sutrop, M. and Laas-Mikko, K. (2012). 'From identity verification to behavior prediction: Ethical implications of second generation biometrics'. *Review of Policy Research*, **29**, 21–36.
- Trémolière, B. and De Neys, W. (2014). 'When intuitions are helpful: Prior beliefs can support reasoning in the bat-and-ball problem'. *Journal of Cognitive Psychology*, **26**, 486–90.
- Trenerry, B., Chng, S., Wang, Y., Suhaila, Z. S., Lim, S. S., Lu, H. Y. and Oh, P. H. (2021). 'Preparing workplaces for digital transformation: An integrative review and framework of multi-level factors'. *Frontiers in Psychology*, **12**, 1–24.
- Tursunbayeva, A., Pagliari, C., Di Lauro, S. and Antonelli, G. (2022). 'The ethics of people analytics: Risks, opportunities and recommendations'. *Personnel Review*, **51**, 900–21.
- Tzafirir, S. S. and Dolan, S. L. (2004). 'Trust me: A scale for measuring manager–employee trust'. *Management Research*, **2**, 115–32.
- Van Der Merwe, A. W. and Puth, G. (2014). 'Towards a conceptual model of the relationship between corporate trust and corporate reputation'. *Corporate Reputation Review*, **17**, 138–56.
- Van Der Werff, L., Blomqvist, K. and Koskinen, S. (2021). 'Trust cues in artificial intelligence: A multilevel case study in a service organization'. In *Understanding Trust in Organizations: A Multilevel Perspective*. New York: Routledge, 307–33.
- Van der Werff, L., Legood, A., Buckley, F., Weibel, A. and de Cremer, D. (2019). 'Trust motivation: The self-regulatory processes underlying trust decisions'. *Organizational Psychology Review*, **9**, 99–123.
- Van Dyck, C., Frese, M., Baer, M. and Sonnentag, S. (2005). 'Organizational error management culture and its impact on performance: A two-study replication'. *Journal of Applied Psychology*, **90**, 1228–40.
- Vargas, R., Yurova, Y. V., Ruppel, C. P., Tworoger, L. C. and Greenwood, R. (2018). 'Individual adoption of HR analytics: A fine grained view of the early stages leading to adoption'. *International Journal of Human Resource Management*, **29**, 3046–67.
- Vellido, A. (2020). 'The importance of interpretability and visualization in machine learning for applications in medicine and health care'. *Neural Computing and Applications*, **32**, 18069–83.
- Venkatesh, V., Morris, M. G., Davis, G. B. and Davis, F. D. (2003). 'User acceptance of information technology: Toward a unified view'. *MIS Quarterly*, **27**, 425–78.
- Vorm, E. and Combs, D. J. (2022). 'Integrating transparency, trust, and acceptance: The intelligent systems technology acceptance model (ISTAM)'. *International Journal of Human–Computer Interaction*, **38**, 1828–45.
- Vorm, E. S. and Miller, A. D. (2020). 'Modeling user information needs to enable successful human-machine teams: Designing transparency for autonomous systems'. In Schmorow, D. D. and Fidopiastis, C. M. (Eds), *Augmented Cognition. Human Cognition and Behavior*. Cham: Springer, 445–65.
- Walsh, M. (2019). 'When algorithms make managers worse'. *Harvard Business Review*, 08 May. Available at <https://hbr.org/2019/05/when-algorithms-make-managers-worse> (accessed 15 May 2023).
- Wang, T., Rudin, C., Doshi-Velez, F., Liu, Y., Klampfl, E. and MacNeille, P. (2015). 'Or's of and's for interpretable classification, with application to context-aware recommender systems'. *arXiv preprint arXiv:1504.07614*. Available at <https://arxiv.org/abs/1504.07614> (accessed 15 May 2023).
- Weibel, A. (2007). 'Formal control and trustworthiness: Shall the twain never meet?'. *Group & Organization Management*, **32**, 500–17.
- Weibel, A., den Hartog, D. N., Gillespie, N., Searle, R., Six, F. and Skinner, D. (2016). 'How do controls impact employee trust in the employer?'. *Human Resource Management*, **55**, 437–62.
- Wildhaber, I. and Lohmann, M. F. (2018). *Algorithmen und recht*. Bern: Editions Weblaw.
- Williams, M. (2007). 'Building genuine trust through interpersonal emotion management: A threat regulation model of trust and collaboration across boundaries'. *Academy of Management Review*, **32**, 595–621.

- Wolf, G. (2020). “‘fuckup night”: Zur aushandlung von scheitern und erfolg’. In Kurilla, R., Kolb-Albers, K., Krämer, H. and Pitsch, K. (Eds), *Sine ira et studio: Disziplinenübergreifende annäherungen an die zwischenmenschliche kommunikation*. Wiesbaden: Springer Fachmedien Wiesbaden, 329–57.
- Worline, M. and Dutton, J. E. (2017). *Awakening Compassion at Work: The Quiet Power that Elevates People and Organizations*. Oakland, CA: Berrett-Koehler Publishers.
- Yee, N. (2014). *The Proteus Paradox: How Online Games and Virtual Worlds Change Us – And How They Don’t*. New Haven, CT: Yale University Press.
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power*. London: Profile Books.
- Zweig, D. and Scott, K. (2007). ‘When unfairness matters most: Supervisory violations of electronic monitoring practices’. *Human Resource Management Journal*, **17**, 227–47.