

Article

An AI-Driven Socio-Technical Framework for Performance Management in Teleworking Environments

Yasmine Wafa ¹ and Justin Longo ^{2,*}¹ Johnson Shoyama Graduate School of Public Policy, University of Saskatchewan, Saskatoon, SK S7N 5B8, Canada; yasminewafa23@gmail.com² Johnson Shoyama Graduate School of Public Policy, University of Regina, Regina, SK S4S 0A2, Canada

* Correspondence: justin.longo@uregina.ca

Abstract

The shift to teleworking, defined as technology-enabled work arrangements in which employees perform organizational tasks remotely outside traditional office settings, has exposed the limitations of traditional performance management systems, including the lack of direct oversight, micromanagement risks, communication barriers, and employee isolation and well-being. These systems often rely on physical presence or intrusive surveillance rather than outcome-based evaluation. This paper asks how AI-driven performance management can be designed to address the documented challenges of teleworking while safeguarding employee autonomy, fairness, and well-being. The study integrates a comprehensive literature review on AI capabilities with empirical evidence from a sequential mixed-methods study of Canadian public servants, comprising machine learning analysis of over 205,000 tweets, document analysis of federal and provincial teleworking policies, a survey of 176 public servants analyzed using logistic regression, and semi-structured interviews with Government of Canada employees. Grounded in socio-technical theory and the Theory of Planned Behavior, the findings reveal that organizational support, workplace socialization, and attitudes are stronger predictors of teleworking success than digital skills or monitoring, while isolation functions as a measurable risk factor. These empirical patterns are mapped to specific AI capabilities to produce a socio-technical framework organized around three interdependent layers: technological, organizational, and human-centered. The paper contributes an empirically grounded alternative to purely speculative treatments of AI in performance management, offering design requirements derived from what teleworkers actually experience rather than from technological possibilities alone. While the framework is analytically grounded in empirical evidence, behavioral theory, and existing AI capabilities, it has not yet undergone full technical or longitudinal organizational validation. Accordingly, it should be understood as a theoretically and empirically informed design artifact intended to guide future implementation and evaluation efforts. It is worth acknowledging that the study's key limitations include a Canada-specific public sector sample, modest survey and interview sizes, and the exploratory nature of several proposed AI capabilities; future cross-sectoral, comparative, and longitudinal research is needed to validate and extend the framework.



Received: 9 March 2026

Revised: 25 May 2026

Accepted: 2 June 2026

Published: 8 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](#)[Attribution \(CC BY\)](#) license.

Keywords: artificial intelligence (AI); teleworking; performance management; ethical AI; socio-technical systems; employee well-being; Canadian public service; theory of planned behavior

1. Introduction

Teleworking refers to flexible work arrangements in which employees perform organizational tasks remotely from traditional office locations while relying on digital communication and collaboration technologies (B. Wang et al., 2021). Contemporary teleworking environments are increasingly characterized as digitally mediated, knowledge-intensive, and hybridized organizational systems that rely on formal coordination mechanisms, platform-based collaboration, and technology-enabled accountability structures (Eurofound and the International Labour Office, 2017; B. Wang et al., 2021). Although the underlying work is often knowledge-intensive in character (Alvesson, 2004), these environments are also commonly embedded within highly regulated organizational contexts—including public-sector, unionized, and policy-constrained settings—where teleworking and performance management practices are shaped by formal accountability requirements, labor regulations, and institutional governance structures that distinguish them from less regulated private-sector environments (Boyne, 2002; Mergel et al., 2019). While often associated with remote work, teleworking is conceptually broader and includes hybrid and digitally mediated work arrangements that rely on organizational connectivity rather than physical co-location. Despite its strategic advantages, teleworking is interpreted variably across research and practice, with related concepts including telecommuting, hoteling, flexi-places, and virtual workplaces (Bailey & Kurland, 2002; De Vries et al., 2019; Brynjolfsson et al., 2020), reflecting the conceptual diversity of remote work arrangements but collectively pointing to a broader shift in how work is organized through digital means.

Importantly, teleworking is not treated in this study as an isolated organizational practice, but as one of the most visible manifestations of digital transformation. Digital transformation refers to the strategic integration of digital technologies into organizational processes, structures, communication systems, and decision-making practices. In this sense, teleworking represents a structural shift from physically co-located work to digitally mediated coordination, where organizational outcomes are increasingly achieved through technology-enabled interaction rather than physical proximity. By enabling employees to fulfill professional responsibilities beyond traditional office settings, teleworking has transformed organizational structures, reduced geographical constraints, and improved flexibility and work–life balance (Dutta & Mishra, 2025). Initially adopted selectively, teleworking became widespread during the COVID-19 pandemic, demonstrating its viability as a long-term organizational strategy. In today’s digital economy, teleworking is therefore both enabled by and constitutive of digital transformation, reshaping workforce management, operational efficiency, and organizational culture (Baki et al., 2023).

Despite its strategic benefits, teleworking presents substantial challenges for organizations, particularly in performance management, which requires sustained employee engagement, fair evaluation, and effective oversight. In the post-COVID era, teleworking has encountered political and cultural pushback as well (Braesemann et al., 2022). Performance management is a systematic process where organizations set expectations, monitor progress, evaluate outcomes, and support employee development. Traditional performance management systems were largely designed for co-located environments, where performance visibility, direct supervision, and informal interaction were readily available. However, in teleworking environments, these assumptions no longer hold. Reduced visibility, asynchronous communication, and the absence of physical oversight weaken traditional mechanisms of evaluation and feedback (Bernstein et al., 2020; Ouchi, 1979). As a result, organizations face increasing difficulty in assessing performance without relying on direct observation.

In response to these challenges, many organizations have adopted digital monitoring and algorithmic tools to restore visibility in remote work settings. However, such approaches often

introduce unintended consequences, including increased surveillance, reduced psychological safety, and diminished trust between employees and managers. This creates a fundamental tension between visibility and autonomy in remote work environments.

Based on the literature, several alternative approaches have emerged that reflect a broader shift in organizational theory toward autonomy, intrinsic motivation, and outcome-based evaluation. Trust-based management models reduce reliance on surveillance and hierarchical control in favor of relational accountability, professional norms, and managerial discretion, particularly in high-trust organizational contexts (Hu et al., 2023). Rather than eliminating control, these models reconfigure it around clearly defined goals, mutual expectations, and interpretive managerial judgment, which become especially important in teleworking environments where direct oversight is limited.

Closely related is the Results-Only Work Environment (ROWE) approach, which evaluates employees based on outputs rather than time or physical presence. ROWE represents a more radical form of outcome-based management in which employees are granted autonomy over when, where, and how work is performed, provided that agreed-upon results are achieved. Research suggests that such models can enhance autonomy, motivation, and work–life integration by decoupling performance from temporal and spatial constraints (Kniffin et al., 2021). However, evidence also shows that ROWE does not eliminate control but redistributes it through peer accountability, self-regulation, and continuous alignment with organizational goals. This creates an autonomy–control paradox in which increased flexibility is accompanied by heightened expectations for responsiveness, coordination, and performance delivery. In practice, employees often engage in additional invisible work such as availability management, productivity signaling, and coordination effort. Moreover, the effectiveness of ROWE depends heavily on the measurability of outputs, which is particularly challenging in knowledge-intensive and public-sector roles where performance is multidimensional. Consequently, organizations often supplement ROWE with formal metrics, coordination routines, and managerial oversight mechanisms.

These dynamics are reflected in broader teleworking performance frameworks, which extend beyond simple output measurement to incorporate a more multidimensional understanding of performance. Recent research emphasizes that effective telework performance systems should integrate not only productivity indicators, but also dimensions such as communication quality, collaboration effectiveness, employee well-being, and adaptability in digitally mediated environments (B. Wang et al., 2021). In this perspective, performance is not treated as a static outcome but as an emergent property of ongoing interactions among individuals, teams, and digital infrastructures. This interpretation is further supported by research on social influence and network dynamics, which suggests that behavior and performance in digitally mediated environments are shaped not only by individual effort but also by interaction patterns within communication networks, where peer effects and information flows can amplify or constrain collective outcomes (Salganik, 2018).

Accordingly, teleworking frameworks increasingly rely on continuous feedback mechanisms, digitally mediated coordination practices, and adaptive goal-setting processes to sustain alignment over time. At the same time, these frameworks highlight the critical role of organizational support structures, including leadership practices, access to digital tools, and clear communication protocols, in shaping performance outcomes. Without such structures, high levels of autonomy may lead to fragmentation, reduced visibility, and coordination breakdowns. Furthermore, telework performance is influenced by contextual and individual factors such as home working conditions, digital competencies, and role characteristics, introducing variability that traditional performance systems are often not designed to accommodate.

As a result, contemporary teleworking performance frameworks aim to balance flexibility with structured accountability, combining outcome-based evaluation with mechanisms that ensure transparency, consistency, and fairness across distributed teams (Kniffin et al., 2021; Jarrahi et al., 2021; Kellogg et al., 2020). When implemented in supportive organizational contexts with strong leadership alignment, these approaches can improve engagement, reduce turnover intentions, and enhance perceived fairness. However, they also highlight a central challenge: the need for systems that sustain performance visibility, coordination, and equity at scale without reverting to intrusive monitoring or undermining trust-based relationships.

Despite these advances, empirical evidence highlights persistent implementation barriers. Large and complex organizations, particularly those operating in regulated, unionized, or policy-constrained environments, face structural constraints that limit full adoption. These include difficulties in standardizing outcome metrics across heterogeneous roles, maintaining equity in performance evaluation, and ensuring auditability and compliance in decision-making processes. As a result, these approaches are often implemented in hybrid forms, with organizations reverting to layered control mechanisms or supplementary evaluation criteria to preserve comparability and accountability, thereby diluting intended autonomy gains. Overall, existing teleworking and performance management approaches can be grouped into three broad categories: surveillance-based systems that prioritize visibility and control, trust-based and outcome-oriented models that emphasize autonomy, and hybrid frameworks that attempt to balance both. While surveillance-based approaches provide high levels of monitoring, they often undermine trust and employee well-being. Trust-based models improve autonomy and engagement but struggle with measurement, comparability, and accountability in complex organizational settings. Hybrid approaches attempt to reconcile these tensions but frequently reintroduce layered control mechanisms that dilute autonomy gains. Collectively, these approaches reveal a persistent gap: the absence of a coherent framework that enables performance visibility, fairness, and coordination without relying on intrusive monitoring or overly abstract outcome measures.

This underscores a persistent gap between normative models and operational realities. This gap is further reflected in public-sector governance frameworks that formalize teleworking as a structured and conditional arrangement rather than an open-ended flexibility option. For example, the Government of Canada's telework policy specifies that telework arrangements are subject to managerial approval and must remain aligned with operational requirements, performance expectations, and service delivery needs, with continuation dependent on ongoing assessment of employee performance and organizational fit (Treasury Board of Canada Secretariat, 2019). This reflects broader institutional pressures shaping organizational practices, whereby formal rules and governance frameworks exert coercive influence that structures how teleworking arrangements are designed and evaluated, contributing to increasing similarity across organizations operating under comparable regulatory environments (DiMaggio & Powell, 1983). Even in highly flexible telework systems, performance management therefore remains embedded within formal accountability structures that emphasize oversight, equity, and managerial discretion. Importantly, while this study draws on the Canadian public service context, this setting is not treated as unique. Instead, it is used as an analytically rich example of a highly structured, policy-driven teleworking environment that reflects broader global patterns in public-sector and regulated organizational contexts.

Many current approaches to performance management rely either on intrusive surveillance—such as keystroke logging or continuous webcam monitoring, which erodes trust and increases stress—or on infrequent and vague metrics that provide limited actionable insights (Kőszegi & Rabin, 2006; Aguinis & Burgi-Tian, 2021; Kulkarni et al., 2024;

Mabaso & Manuel, 2024; Mkhize & Lourens, 2025). As a result, scholars and practitioners advocate for trust-centered, outcome-focused approaches that emphasize employee autonomy, engagement, and overall well-being (Ball, 2010; Deloitte, 2023). This creates a fundamental tension between visibility and trust, revealing a critical gap in existing performance management systems: a lack of approaches that can generate meaningful, data-informed insights without reverting to surveillance-based control.

As teleworking becomes institutionalized, this challenge extends beyond operational concerns to broader organizational and policy implications. Ensuring fair evaluation, sustained engagement, and effective support in distributed environments is not only an operational issue but also a strategic and societal one—particularly in public-sector contexts where accountability, transparency, and equity are paramount.

Emerging solutions highlight AI-driven performance management as a viable approach for addressing these challenges. AI-driven analytics have the capacity to identify performance trends, deliver personalized feedback in real time, foster skill development, and align employees' efforts with organizational goals, while simultaneously protecting privacy and supporting psychological well-being (Kalischko & Riedl, 2021; Deloitte, 2023). Recent advances in AI-enabled performance management further illustrate both the potential and complexity of data-driven evaluation systems. Emerging research shows that AI can shift performance assessment from periodic and subjective evaluations toward continuous, multi-source, and data-driven insights, improving decision accuracy and responsiveness in organizational contexts (Nayak & Jagadeeswari, 2025). At the same time, these systems do not eliminate managerial judgment but reconfigure it, requiring new forms of human–AI collaboration and raising important concerns regarding transparency and perceived fairness (Nayak & Jagadeeswari, 2025). Complementary evidence from AI-enabled HRM studies indicates that capabilities such as intelligent feedback, predictive analytics, and personalized performance support can enhance employee engagement and organizational outcomes, but only when supported by trust, organizational legitimacy, and clear governance structures (Jangbahadur et al., 2025). In parallel, research on AI-driven management in digital work environments highlights the role of these systems in enabling real-time coordination, adaptive goal-setting, and continuous performance optimization in remote and hybrid teams, while also introducing risks related to algorithmic control, work intensification, and employee surveillance (Dinh, 2026). Taken together, this emerging body of research suggests that AI-enabled performance management systems offer significant advantages for distributed work environments but must be carefully designed to balance efficiency, transparency, autonomy, and employee well-being.

Importantly, AI is not positioned here as a replacement for existing management approaches, but as an enabling layer that can augment trust-based and outcome-oriented frameworks. However, the existing literature does not clearly explain how AI capabilities can be systematically aligned with empirically observed workplace challenges and established behavioral theories. Unlike traditional oversight tools, AI can focus on outcomes rather than surveillance, monitor collaboration dynamics, and recognize qualitative contributions such as problem-solving, creativity, and teamwork.

In response to this gap, this study addresses the following research question: How can AI-driven performance management systems be designed to support employee well-being, productivity, and fairness in teleworking environments, rather than simply monitoring activity? This reframes performance management from a control-oriented perspective (“What are employees doing?”) to a support-oriented perspective (“How are employees doing?”).

This paper introduces an AI-driven, socio-technical performance management framework designed for teleworking environments. Methodologically, the study adopts a socio-technical design approach informed by a sequential mixed-methods research design. What

distinguishes this contribution from purely conceptual treatments is its empirical grounding: the framework's design requirements are derived from a sequential mixed-methods study of teleworking in the Canadian public service (Wafa, 2024), comprising machine learning analysis of over 205,000 tweets, document analysis of federal and provincial government teleworking policies, an online survey of 176 public servants, and semi-structured interviews with Government of Canada employees. These empirical findings establish what teleworkers, and their managers, actually experience, struggle with, and need—providing the evidentiary base from which the proposed AI-driven framework is built. While the empirical analysis is situated within the Canadian public service, this case is used as a theoretically informative and analytically generalizable context, rather than as a basis for statistical generalization. Accordingly, the findings should be interpreted as contextually grounded insights that inform the development of a broader socio-technical framework. At the same time, this approach has important limitations that shape the scope and applicability of the proposed framework. The empirical evidence is drawn from a single national and institutional context; the public-sector setting may differ substantially from private-sector or less regulated environments; and organizational, cultural, and policy-specific factors may influence teleworking experiences in ways that are not directly transferable. Consequently, the framework's applicability to other organizational and national contexts should be understood as provisional and subject to further empirical validation across diverse settings.

While the proposed framework is analytically grounded in empirical findings, behavioral theory, and AI capabilities identified in the literature, the present study does not constitute a full technical implementation or longitudinal evaluation. Rather than claiming validated organizational effectiveness, the framework should be understood as a theoretically and empirically informed socio-technical design artifact intended to guide future development and implementation efforts. Accordingly, its practical effectiveness in addressing issues related to employee well-being, productivity, fairness, and non-intrusive performance visibility requires future pilot testing, comparative evaluation, and longitudinal validation across diverse teleworking contexts.

This study makes three key contributions. First, it advances theoretical understanding by integrating socio-technical theory and the Theory of Planned Behavior with supporting insights from SDT and JD-R to explain both system design and adoption dynamics. Second, it provides a practical contribution by proposing a structured AI-driven performance management framework tailored to teleworking environments. The framework is intended as a design-oriented and exploratory contribution rather than a fully validated operational model, thereby establishing a foundation for future implementation and evaluation research. Third, it offers a methodological contribution by demonstrating how empirical evidence, theory, and AI capabilities can be systematically integrated into artifact design.

The Canadian public-sector context is used in this study as a representative example of a highly regulated, knowledge-intensive, and hybridized teleworking environment. Similar characteristics have been identified across contemporary teleworking environments in advanced economies, where remote work arrangements increasingly depend on digital coordination, knowledge-based tasks, and formal accountability structures (Eurofound and the International Labour Office, 2017; B. Wang et al., 2021). In this sense, the Canadian public service reflects broader organizational trends associated with digitally mediated and policy-driven work environments. At the same time, the Canadian context is distinguished by particularly formalized public-sector governance structures, strong accountability requirements, and institutionalized telework policies that shape how performance management systems are designed and implemented (Treasury Board of Canada Secretariat, 2019). These characteristics provide analytically rich insight into how telework-

ing and performance management interact under structured institutional conditions. This duality allows the findings to inform a broader, transferable socio-technical framework while remaining grounded in a real-world empirical setting used to derive generalizable design principles for AI-enabled teleworking performance management systems.

The logic of the paper can be expressed as follows: we present what the evidence tells us about the real challenges and success factors for remote performance management; what the literature tells us AI can do; and how a socio-technical framework could integrate those AI capabilities to address the empirically documented needs. By combining a comprehensive literature review on AI capabilities with empirical evidence from the Canadian public service, and by integrating socio-technical theory with the Theory of Planned Behavior, this paper offers a framework that is speculative in its forward-looking design but grounded in real-world evidence. The remainder of the paper is structured as follows. Section 2 reviews the literature on teleworking, performance management, and AI-enabled systems. Section 3 presents the theoretical framework. Section 4 outlines the empirical foundation. Section 5 introduces the proposed framework. Section 6 discusses the findings and implications, and Section 7 concludes the paper.

2. Literature Review

Across the literature, three persistent gaps remain in relation to teleworking and performance management. First, concepts such as productivity, engagement, and employee well-being are inconsistently defined and measured across teleworking studies, limiting conceptual clarity and comparability of findings (Aguinis & Burgi-Tian, 2021; B. Wang et al., 2021). Second, existing research often focuses on teleworking outcomes such as satisfaction and productivity without clearly identifying the mechanisms that support effective performance management in distributed environments (Kniffin et al., 2021). Third, there is limited integration between behavioral theory, organizational design principles, and technological capabilities, particularly in AI-enabled performance management systems (Jarrahi et al., 2021; Kellogg et al., 2020). As teleworking becomes increasingly institutionalized across industries, organizations face growing challenges in sustaining productivity, accountability, coordination, and employee well-being in digitally mediated work environments characterized by autonomy, flexibility, spatial separation, and asynchronous communication. These conditions create a structural misalignment between traditional performance management systems based on physical visibility and emerging teleworking environments that require more adaptive, trust-centered, and outcome-oriented approaches.

Artificial intelligence (AI) is increasingly positioned as a potential enabler of performance management in teleworking environments. AI systems can support performance evaluation through real-time analytics, adaptive feedback, workload monitoring, collaboration analysis, and behavioral pattern detection. However, the literature does not yet clearly explain how these technological capabilities can be systematically aligned with empirically observed workplace challenges and broader organizational requirements related to fairness, transparency, employee autonomy, and well-being. This creates an important disconnect between organizational practice, behavioral theory, and technological system design in AI-enabled performance management research.

2.1. Teleworking and Digital Transformation

As defined above, teleworking refers to flexible work arrangements in which employees operate remotely from traditional office locations while leveraging digital communication and collaboration technologies (B. Wang et al., 2021). Despite its strategic advantages, teleworking is interpreted variably across research and practice, with related concepts including telecommuting, hoteling, flexi-places, and virtual workplaces (Bailey & Kurland,

2002; De Vries et al., 2019; Brynjolfsson et al., 2020), reflect the conceptual diversity of telework arrangements, but collectively point to the same underlying shift: the digital transformation of work organization and coordination.

Although teleworking has existed for decades, its adoption accelerated significantly during the COVID-19 pandemic, which forced organizations to implement telework at an unprecedented scale. Teleworking models combining remote and on-site work have since become dominant, particularly among large organizations (GitLab, 2021; Microsoft, 2021; Statistics Canada, 2021). Evidence from the literature consistently suggests that teleworking can generate positive organizational and employee outcomes. For example, meta-analytic evidence indicates that teleworking is associated with higher job satisfaction, improved work–life balance, reduced stress, and lower turnover intentions, although excessive telework may weaken interpersonal relationships and increase feelings of isolation (Gajendran & Harrison, 2007). Complementing this broader evidence base, experimental research provides causal support for these effects. A randomized field experiment conducted in a large service organization found that employees assigned to work from home experienced significant productivity gains, driven by both increased working time and improved efficiency, alongside higher job satisfaction and reduced turnover rates (Bloom et al., 2015). However, the study also highlighted that some employees chose to return to the office due to reduced social interaction, underscoring the importance of balancing flexibility with social and organizational connection. This shift exposed limitations in traditional performance management approaches, particularly in maintaining performance visibility, coordination, fairness, and employee well-being without direct supervision in distributed environments.

In this sense, teleworking should be understood as one of the most visible and organizationally significant expressions of digital transformation, rather than as an isolated work arrangement. Digital transformation is defined as the strategic integration of digital technologies across organizational processes and structures that enables teleworking by providing the technological infrastructure necessary for coordination, communication, and performance management in distributed environments. Tools such as cloud computing, workflow automation, and AI-driven analytics allow organizations to coordinate work, provide continuous feedback, and evaluate performance in ways that do not depend on physical presence (Madanchian et al., 2024). However, the effectiveness of these technologies depends not only on availability but also on how they are embedded within organizational structures, cultures, and management practices. In this sense, digital transformation is not purely technical but socio-organizational. Beyond technology, successful digital transformation requires trust-based organizational cultures and adaptive leadership that balance performance expectations with employee autonomy and well-being (Hossain et al., 2025).

Rather than operating independently, teleworking and digital transformation are mutually reinforcing. Telework drives organizations to invest in scalable infrastructure, adopt collaborative platforms like Microsoft Teams, Slack, and Zoom, and integrate AI-powered analytics for workflow optimization and decision-making (Mancl & Fraser, 2023; Microsoft, 2025a, 2025b, 2025c). However, this shift also introduces new dependencies, visibility asymmetries, and coordination challenges, indicating that digital transformation is not a neutral enabler but a structural force reshaping control, communication, and work organization. These developments highlight the broader shift from office-centric structures to digitally mediated, flexible workplaces.

These developments have direct implications for performance management. Traditional approaches based on physical presence and direct oversight are increasingly inadequate for assessing productivity, sustaining collaboration, and supporting employee well-being in remote environments. This creates a growing need for performance management systems aligned with digital workflows and human-centered principles. AI-enabled

performance management has emerged as a response, offering outcome-oriented, adaptive, and data-driven approaches. However, this shift has also exposed important limitations in traditional performance management systems, which were largely designed for co-located work and rely heavily on physical visibility, direct supervision, and synchronous interaction. In distributed digital environments, these assumptions no longer hold, creating challenges related to performance visibility, coordination, fairness, and employee well-being. Addressing these challenges therefore requires not only new management practices but also a deeper reconfiguration of how digital tools are embedded within organizational systems. As a result, AI can enhance transparency, fairness, and organizational responsiveness when combined with socio-technical principles, although its effectiveness depends on careful design and governance.

Importantly, much of the existing research on telework and digital transformation is based on Western, private-sector contexts. This reflects a broader WEIRD (Western, Educated, Industrialized, Rich, Democratic) bias in organizational research (Henrich et al., 2010), where empirical evidence is disproportionately drawn from highly developed, market-oriented environments and then implicitly treated as universally applicable. As Henrich et al. (2010) demonstrate, such samples are often outliers rather than representatives of global populations, raising fundamental concerns about the external validity of widely accepted theories and practices.

In the context of AI-enabled management, this bias is particularly consequential. Much of the literature is grounded in private-sector, platform-based, or technology-driven organizations, where performance is more readily quantifiable, competitive pressures are pronounced, and managerial discretion is relatively flexible. These conditions shape both the design of AI systems and the assumptions embedded within them—such as the prioritization of efficiency, optimization, and measurable outputs. As a result, widely cited best practices may not fully capture the institutional and regulatory complexities characteristic of public-sector settings.

Public-sector organizations operate under distinct logics, including formalized accountability structures, transparency requirements, equity considerations, and policy-driven mandates that extend beyond efficiency alone. Performance is often multidimensional, involving service quality, public value, and adherence to procedural fairness, which are not easily reducible to standardized metrics. In the context of the Canadian public service, institutional constraints, policy frameworks, and accountability requirements may therefore shape teleworking practices—and the feasibility of AI-driven performance management—in ways that differ substantially from those observed in private-sector environments.

This misalignment has important implications. First, it suggests that AI tools and management practices developed in WEIRD, private-sector contexts may embed assumptions that do not translate effectively to public-sector settings. Second, it highlights the risk of uncritically importing models that prioritize efficiency over accountability or standardization over contextual judgment. Third, it underscores the need for empirically grounded, context-sensitive approaches that account for institutional variation rather than assuming cross-context equivalence. Accordingly, this study does not treat existing AI-in-management literature as universally generalizable, but instead engages with it critically, using the Canadian public sector as a context through which to examine how these assumptions hold, where they break down, and how they may need to be adapted.

Comparative international research further highlights these contextual differences. For example, Eurofound and the International Labour Office (2017) on telework and ICT-based mobile work demonstrates that teleworking is not a uniform organizational phenomenon, but one that varies substantially across countries in terms of prevalence, regulation, intensity, and employee experience. Across European contexts, differences

in labor market structures, institutional protections, and organizational practices shape not only who engages in telework, but also how it is experienced in terms of autonomy, workload, and work–life boundaries.

Importantly, Eurofound’s findings show that telework outcomes are strongly mediated by national policy frameworks and organizational governance structures, rather than being determined solely by technological infrastructure or managerial intent. For instance, regulatory environments that provide stronger worker protections and clearer boundaries around working time tend to mitigate some of the negative effects associated with telework, such as overwork and boundary erosion.

This suggests that teleworking outcomes are shaped as much by policy and governance structures as by technological capabilities, reinforcing the need for context-sensitive performance management approaches. In other words, the effectiveness of remote performance management systems—including emerging AI-driven approaches—cannot be separated from the institutional environments in which they are deployed.

From this perspective, Eurofound’s evidence complements concerns raised in this study regarding the transferability of AI-enabled performance management models across contexts. It reinforces the argument that performance management systems developed in highly digitalized, private-sector environments may not directly translate to public-sector settings, where regulatory oversight, accountability requirements, and public service mandates fundamentally reshape both managerial practices and employee expectations.

In summary, the literature underscores that the future of performance management in teleworking environments is not defined solely by physical location but by the integration of technology, organizational culture, and human-centered management. Rather than presenting a linear progression toward more effective digital work systems, the literature points to an evolving set of trade-offs and tensions that organizations must navigate. AI-driven, socio-technical frameworks provide a conceptual foundation for developing performance management systems that are ethical, inclusive, and effective in digitally mediated work environments, but their effectiveness ultimately depends on how these competing demands are balanced in practice.

2.2. Challenges in Performance Management for Telework

Performance management is conventionally framed as an iterative cycle comprising goal-setting, ongoing monitoring, periodic evaluation, and feedback or development conversations (Aguinis, 2009; Pulakos et al., 2015). Each stage rests on assumptions that hold most easily when employees and managers share a workspace, and each is altered—but not equally—when work moves to a distributed setting. Goal-setting in co-located environments relies on visible workflows and informal calibration of expectations. In teleworking, this must be made explicit, written, and outcome-defined, because the contextual cues that previously absorbed ambiguity are no longer available. Monitoring traditionally draws on direct observation and ambient awareness of who is doing what. These cues disappear in distributed settings, however, leaving managers to choose between intrusive digital surveillance, infrequent check-ins, or trust-based supervision (Bailey & Kurland, 2002; Mkhize & Lourens, 2025). Evaluation typically combines observed behavior, self-report, and peer input; in teleworking, observed behavior is reduced to digital traces that may privilege visibility over substance and disadvantage employees whose contributions are less easily captured in metadata (Aguinis & Burgi-Tian, 2021; Gibbs et al., 2023). Feedback and development, finally, depend on regular relational contact that sustains trust and shared interpretation; mediated communication tends to compress these exchanges into transactional updates, weakening their developmental function (Buckingham & Goodall, 2015). The recurring challenges discussed in the literature on remote performance man-

agement (e.g., lack of direct oversight, micromanagement, communication barriers, and isolation and well-being concerns) can therefore be read not as a single undifferentiated “telework problem” but as failures concentrated at specific points in this cycle, where the assumptions that previously made each step work no longer hold.

Being physically apart limits managers’ ability to gauge engagement, monitor task progress, and understand team dynamics through direct observation or informal interactions (Mkhize & Lourens, 2025). This can result in insufficient oversight, risking misalignment and underperformance, or excessive scrutiny, manifesting as micromanagement that undermines trust and autonomy. Transparent expectations and adaptive management practices are therefore essential to maintain alignment while supporting employee autonomy and engagement (Aguinis & Burgi-Tian, 2021; Gibbs et al., 2023).

Lack of Direct Oversight. Physical separation reduces managers’ ability to assess engagement, task progress, and team dynamics through observation and informal cues (Mkhize & Lourens, 2025). This can lead to insufficient oversight, risking misalignment and underperformance, or excessive scrutiny, manifesting as micromanagement that erodes trust and autonomy (Mkhize & Lourens, 2025).

Micromanagement Risks. Telework can exacerbate tendencies toward micromanagement. Managers, facing limited visibility, may overcompensate by closely scrutinizing tasks, undermining employee morale, engagement, and psychological safety (Bailey & Kurland, 2002; Mkhize & Lourens, 2025). Balancing oversight with autonomy is critical to sustaining trust and motivation in distributed teams. A related structural response to these challenges is the adoption of outcome-based performance management, which shifts evaluation from activity monitoring to outputs and goal attainment. While this reduces reliance on continuous supervision, it does not eliminate evaluation challenges; instead, it redefines what is measured and valued within performance systems. Overreliance on output indicators may still overlook important qualitative dimensions of work, including collaboration, creativity, and discretionary effort.

Communication Barriers. Digitally mediated communication introduces structural constraints. Asynchronous messaging, absence of non-verbal cues, and limited spontaneous interactions can reduce clarity, slow decision-making, and hinder trust-building (Bailey & Kurland, 2002; Mkhize & Lourens, 2025). Informal knowledge sharing and team cohesion, essential for problem-solving and innovation, are harder to maintain in remote contexts.

Employee Isolation and Well-being. Teleworking can lead to social isolation, reduced belonging, and detachment from organizational culture, negatively affecting engagement, motivation, and performance (Bailey & Kurland, 2002; Mkhize & Lourens, 2025). Prolonged isolation increases stress and burnout risk, highlighting the importance of strategies to maintain connection and support employee well-being.

Taken together, these challenges show that shifting from input-based to output-based evaluation does not resolve the core problem of performance management in telework but rather reconfigures it. Output-based systems still require careful design to ensure that important qualitative dimensions—such as collaboration, creativity, and discretionary effort—are not overlooked. Therefore, transparent expectations and adaptive management practices remain essential to balance accountability with autonomy and engagement (Aguinis & Burgi-Tian, 2021; Gibbs et al., 2023).

These challenges can be further interpreted through the lenses of Job Demands–Resources (JD-R) model, which suggests that telework can increase job demands—such as communication complexity and role ambiguity while simultaneously altering access to key resources like managerial support and social interaction (Bakker & Demerouti, 2007). From this perspective, performance difficulties arise not only from monitoring constraints but also from shifts in the balance between demands and resources. Similarly, Self-Determination

Theory (SDT) helps explain how excessive monitoring or poorly designed performance systems may undermine autonomy and relatedness, reducing intrinsic motivation and engagement (Ryan & Deci, 2000). Taken together, these frameworks highlight that effective performance management in telework requires balancing control mechanisms with support for autonomy, connection, and well-being rather than prioritizing any single dimension.

2.3. AI-Driven Performance Management: Capabilities, Solutions, and Limitations

Building on the challenges identified above, AI-driven performance management offers targeted capabilities to address the unique constraints of teleworking environments. However, these capabilities must be understood within the broader context of algorithmic management, which introduces inherent tensions between control, coordination, and worker autonomy (Kellogg et al., 2020). As Kellogg et al. (2020) argue, algorithmic management systems do not simply support managerial decision-making; they actively reconfigure the mechanisms through which control is exercised in organizations by embedding evaluation, monitoring, and coordination into digital infrastructures.

From this perspective, AI-driven performance management can be understood as part of a broader shift toward “digitally mediated control systems,” where managerial oversight is increasingly exercised through data infrastructures rather than direct supervision. This shift alters not only how work is monitored, but also how it is defined, evaluated, and experienced by employees.

By providing continuous, data-informed evaluation, adaptive feedback, and ethical oversight, AI enhances managerial capacity while supporting employee autonomy, engagement, and well-being. While AI enables continuous, data-informed evaluation and adaptive feedback, thereby enhancing managerial capacity, it simultaneously raises concerns regarding surveillance, autonomy, and the redistribution of control. In line with Kellogg et al. (2020), these tensions reflect a central feature of algorithmic management: the simultaneous expansion of managerial visibility and the potential erosion of worker discretion, as performance becomes increasingly legible through digital traces.

At the same time, these AI systems have inherent limitations and require careful ethical and organizational governance. As Kellogg et al. (2020) emphasize, algorithmic systems embed organizational priorities and power relations into their design, shaping what is measured, what is excluded, and how work is ultimately valued. This makes governance and design choices central to determining whether AI supports or constrains fair and effective performance management. The following subsections present solutions aligned with the five key dimensions identified in the literature: automated productivity tracking, sentiment analysis and employee well-being, adaptive goal-setting and personalized feedback, enhancing fairness and reducing bias, and ethical considerations and employee trust.

2.3.1. Automated Productivity Tracking

Reduced direct oversight and the risk of micromanagement, central challenges in telework, are addressed by AI-powered analytics that aggregate task completion, workflow patterns, and team interactions. These AI-driven systems offer managers real-time, comprehensive insights into both individual and team performance, allowing for timely, proactive interventions without resorting to intrusive supervision (J. Wang & Panesar, 2022; Microsoft, 2025a, 2025b, 2025c; Asana, 2025a, 2025b, 2025c; Culture Amp, 2025; BambooHR, 2025). However, this shift from direct supervision to data-driven visibility does not eliminate control; rather, it reconfigures it into more continuous and less transparent forms. By moving the focus away from continuous oversight toward data-informed decision-making,

AI fosters accountability while maintaining employee autonomy (Mkhize & Lourens, 2025), yet this balance depends on how such systems are perceived and implemented in practice.

AI can also help capture aspects of qualitative performance that traditional metrics often miss. Natural language processing (NLP) can analyze project updates, emails, and collaborative documents to highlight contributions to problem-solving, knowledge sharing, and collaborative engagement (Tausczik & Pennebaker, 2010; Mäntylä et al., 2018). Organizational network analysis (ONA) identifies patterns of interaction and collaboration, providing indirect insights into teamwork quality and information flow (Goodings et al., 2024; Humanyze, 2025). Nevertheless, these approaches rely on indirect indicators and thus remain limited in their ability to directly assess qualitative contributions, as they can only infer aspects of performance through available data proxies. Complex attributes such as creativity, strategic thinking, and nuanced problem-solving cannot be directly measured and are instead inferred through proxies, raising questions about the validity and completeness of AI-based evaluations. Human judgment remains indispensable for interpreting AI-generated insights and integrating them into comprehensive and context-sensitive performance evaluations.

2.3.2. Sentiment Analysis and Employee Well-Being

Communication barriers, social isolation, and well-being concerns in telework are mitigated through AI-enabled sentiment analysis and well-being monitoring. NLP systems analyze textual data from emails, messaging platforms, and collaboration tools to detect emotional cues, enabling managers to identify stress, disengagement, or burnout early (Mäntylä et al., 2018; Kulkarni et al., 2024). AI-driven mental health tools, including conversational agents like Woebot and Wysa, as well as platforms such as Spring Health, Lyra Health, and Headspace Care, deliver evidence-based interventions grounded in cognitive behavioral therapy, mindfulness, and personalized recommendations (Fitzpatrick et al., 2017; Inkster et al., 2018; Callahan et al., 2024; Lee et al., 2025).

While these tools expand organizational capacity to monitor and support well-being, they also blur the boundary between support and surveillance. Scheduling and productivity tools, including Microsoft Viva Insights, Clockwise, and Reclaim.ai, optimize workloads, protect focus time, and reduce burnout risk (Clockwise, 2023; Microsoft, 2025a, 2025b, 2025c; Reclaim.ai, 2025), yet their effectiveness depends on employee trust and voluntary engagement. If perceived as intrusive, such systems may undermine psychological safety rather than enhance it. As a result, AI-based well-being interventions must be transparent, ethically designed, and ideally opt-in. Human-centered leadership and organizational culture remain critical to reinforce genuine support, trust, and social connection (Ashdown, 2018).

2.3.3. Adaptive Goal-Setting and Personalized Feedback

AI enables dynamic, adaptive goal-setting and continuous personalized feedback, addressing challenges in alignment, performance clarity, and managerial oversight. By analyzing individual progress, peer performance, and workload distribution, AI systems recalibrate goals to enhance motivation, prevent overload, and support engagement (Rockmann & Pratt, 2015; Davenport & Beier, 2020; Jarrahi et al., 2021). Continuous, context-sensitive feedback replaces static, infrequent reviews, providing actionable insights that facilitate learning and skill development (Pulakos et al., 2015; Cosa & Torelli, 2024).

However, the increased automation of feedback processes introduces risks of depersonalization. Overreliance on algorithmic outputs may diminish opportunities for meaningful dialog, thereby weakening empathy, relational judgment, and contextual interpretation. This creates a tension between efficiency and relational quality in performance management. Sustained human oversight is therefore essential to ensure that performance assessments

remain balanced, fair, and sensitive to individual circumstances (Buckingham & Goodall, 2015; Dastin, 2018; Leicht-Deobald et al., 2019).

2.3.4. Enhancing Fairness and Reducing Bias

AI has the potential to mitigate human evaluation biases and support equitable performance management. AI can minimize biases such as favoritism, halo effects, or recency bias by employing consistent, data-driven evaluation criteria (Raisch & Krakowski, 2021). Platforms like Workday, BetterUp, and Eightfold AI enable consistent evaluations across employees, while algorithmic audits help uncover systemic inequities in promotions, recognition, or compensation (BetterUp, 2025a, 2025b; Binns et al., 2018; Raghavan et al., 2020).

At the same time, the assumption of algorithmic objectivity is increasingly contested. Biases can persist through unrepresentative training data or flawed algorithm design. As a result, AI does not eliminate bias but redistributes it in less visible forms. This underscores the need for human oversight and critical evaluation to ensure fairness, contextual interpretation, and procedural legitimacy (Glikson & Woolley, 2020; Langer et al., 2021; Mehrabi et al., 2021).

2.3.5. Ethical Considerations and Employee Trust

The use of AI in performance management raises critical ethical considerations, particularly regarding privacy, consent, data security, and algorithmic accountability. Maintaining trust requires transparency about what data are collected, how they are analyzed, and how outputs inform decisions (Binns et al., 2018; Cowgill, 2018; Yanamala, 2023).

However, transparency alone may not be sufficient. Employees must have opportunities to participate in AI system selection, customization, and evaluation to reduce power asymmetries and reinforce procedural fairness (Shrestha et al., 2019). Human oversight is essential to ensure that AI complements managerial judgment rather than replacing relational and ethical decision-making. Without such safeguards, AI systems risk undermining trust even when designed to enhance fairness and efficiency. AI systems must operate within ethical and governance frameworks that uphold organizational values, protect autonomy, and sustain legitimacy (Floridi et al., 2018; European Commission, 2019; John et al., 2022).

Across these applications, a consistent tension emerges: while AI can enhance efficiency, consistency, and scalability, it may also introduce risks related to depersonalization, opacity, and over-standardization. This underscores the importance of socio-technical approaches that integrate technological capabilities with human judgment, organizational context, and ethical governance, particularly in domains where performance cannot be fully reduced to quantifiable outputs.

Despite AI advances, AI systems remain fundamentally constrained by their reliance on observable data proxies. As such, they cannot directly measure complex human qualities such as creativity, contextual judgment, ethical reasoning, or nuanced problem-solving. These limitations are not merely technical but epistemological, reflecting the difficulty of translating rich, context-dependent human behaviors into quantifiable indicators. In practice, this means that AI systems do not evaluate performance itself, but rather representations of performance constructed from available data, which may only partially capture the underlying reality.

This reliance on proxies introduces several important risks. First, it may privilege what is easily measurable over what is meaningful, leading to an overemphasis on quantifiable outputs at the expense of less visible but equally critical contributions, such as mentoring, informal coordination, or creative insight. Second, it may encourage behavioral adaptation, where employees optimize for what is measured rather than what is organizationally valuable, potentially distorting performance outcomes. Third, it may obscure contextual

factors—such as task complexity, team dynamics, or organizational constraints—that shape performance but are difficult to encode in data-driven systems.

These limitations are particularly consequential in knowledge-intensive and public-sector contexts, where performance often involves ambiguity, discretion, and value-based judgment rather than standardized outputs. In such settings, the risk is not only incomplete measurement but misrepresentation, where algorithmic evaluations may systematically overlook or misinterpret critical dimensions of work. This reinforces the need for human interpretation and highlights the risks of over-reliance on algorithmic evaluation.

The AI capabilities reviewed here represent a range of possibilities whose real-world value depends on whether they address documented, demonstrated needs—not merely theoretical gaps. At the same time, the literature reveals important limitations, contextual biases, and unresolved tensions that constrain the direct application of AI-driven management approaches. Taken together, these insights suggest that the challenge is not simply to improve measurement accuracy, but to critically assess what should be measured, how it is represented, and how algorithmic outputs are interpreted within organizational decision-making processes. This reinforces that AI should not be viewed as a standalone solution, but as part of a broader socio-technical system requiring careful alignment with organizational context and employee needs. The empirical evidence presented below in Section 4 provides exactly this grounding, establishing what teleworkers actually experience and what the data reveal about the conditions under which AI-driven performance management could make a meaningful difference.

3. Theoretical Framework

This section establishes the conceptual foundation used to design the AI-driven performance management framework proposed in this study. Its purpose is not to provide an exhaustive review of theories, but to explain how selected theoretical lenses are combined to structure (i) the system design logic of AI-enabled performance management and (ii) the behavioral mechanisms that determine its adoption and sustained use in telework environments. In this sense, the theoretical framework serves as a bridge between the empirical gaps identified in the literature review and the development of the proposed framework.

The increasing integration of AI into organizational management systems has intensified the need for theoretical approaches capable of explaining both technological design and human interaction within digitally mediated work environments (Kellogg et al., 2020; Jarrahi et al., 2021). Existing research suggests that AI-enabled management systems cannot be understood solely as technical tools, but rather as socio-organizational systems shaped by organizational structures, employee perceptions, managerial practices, and governance mechanisms (Bostrom & Heinen, 1977; Trist, 1993).

To achieve this, the study integrates two complementary theoretical perspectives: socio-technical theory and the Theory of Planned Behavior (TPB). Socio-technical theory argues that organizational effectiveness emerges through the joint optimization of technological and social systems rather than through technological efficiency alone (Bostrom & Heinen, 1977; Trist, 1993). Socio-technical theory provides the system-level design logic for aligning technological, organizational, and environmental components in AI-enabled performance management systems. Complementing this organizational perspective, TPB explains how attitudes, subjective norms, and perceived behavioral control influence behavioral intentions and system adoption (Ajzen, 1991). TPB complements this by explaining how individual attitudes, perceived social norms, and perceived behavioral control shape employees' and managers' willingness to adopt and engage with such systems. Together, these theories enable a multi-level understanding of both system design and human response.

Supporting perspectives from Self-Determination Theory (SDT) and the Job Demands–Resources (JD-R) model are also used throughout the analysis to explain how autonomy, competence, social connection, workload demands, and organizational resources influence employee motivation, engagement, strain, and well-being in digitally mediated work environments (Ryan & Deci, 2000; Bakker & Demerouti, 2007). While these supporting theories do not directly structure the framework architecture, they provide additional explanatory insight into the motivational and well-being dynamics associated with AI-enabled performance management systems.

Importantly, AI-enabled performance management systems are not neutral tools but embedded socio-technical arrangements that can reshape how performance is defined, measured, and governed. Research on algorithmic management suggests that digital evaluation systems may redistribute authority, increase visibility, and embed organizational priorities into technological infrastructures that shape managerial control and employee experience (Kellogg et al., 2020; Raisch & Krakowski, 2021). From a critical management perspective, they may also redistribute authority by shifting evaluative discretion from human judgment toward algorithmically generated metrics. This introduces important implications for control, accountability, and managerial discretion in digitally mediated work environments. Rather than treating this as a separate theoretical stream, Critical Management Studies (CMS) is used here as an interpretive lens to critically examine these implications, particularly in relation to power and resistance.

3.1. Socio-Technical Theory

This study uses socio-technical theory as the primary system-level foundation for designing the proposed AI-driven performance management framework. The purpose of applying this theory is not to provide a general description of organizational systems, but to specify how technological, organizational, human, and environmental elements must be aligned to ensure effective performance management in teleworking contexts.

Socio-technical theory conceptualizes organizations as interdependent systems in which social and technical components jointly shape performance outcomes (Trist, 1993). In the context of AI-enabled performance management, this is particularly relevant because performance is no longer determined solely through managerial observation, but through digitally mediated systems such as productivity analytics, adaptive feedback mechanisms, and sentiment analysis tools. However, these technical capabilities only generate value when embedded within appropriate organizational structures and human practices, including trust, motivation, engagement, and coordination mechanisms.

Accordingly, this study applies socio-technical theory to structure AI-enabled performance management systems across four interdependent subsystems. The technical subsystem refers to AI-driven tools and digital infrastructures, including automated productivity tracking, adaptive feedback systems, sentiment analysis, and fairness-supporting algorithms. The personnel subsystem captures human factors such as employee skills, motivation, trust, and attitudes toward AI-based evaluation systems. The organizational subsystem includes governance structures, performance management policies, feedback mechanisms, and decision-making processes that determine how AI outputs are interpreted and used. The environmental subsystem reflects external institutional conditions, including regulatory requirements, cultural expectations, and ethical standards governing the use of AI in workplace monitoring and evaluation (Bélanger et al., 2013; European Commission, 2019; Zuboff, 2023).

Within this structure, socio-technical theory emphasizes that effective AI-driven performance management depends on alignment across all four subsystems rather than optimization of individual technologies. Misalignment between these elements may result in re-

duced trust, resistance to system use, or unintended consequences such as over-monitoring or distorted performance signals. This is particularly important in telework environments, where reduced physical interaction increases reliance on digital representations of work.

At a structural level, socio-technical theory also clarifies how AI systems reshape organizational control mechanisms by embedding evaluation and decision-making processes within technological infrastructures. This shifts performance management from direct supervisory judgment toward data-driven representations of work, thereby altering how performance is defined, monitored, and acted upon in organizations.

From this perspective, resistance to AI adoption should not be interpreted solely as a behavioral issue, but as an outcome of misalignment between socio-technical subsystems. To further interpret these dynamics, Critical Management Studies (CMS) is used as a complementary lens to highlight how AI-enabled performance systems may reconfigure power relations within organizations. In particular, algorithmic performance systems can increase visibility and standardization in evaluation processes, which may reduce managerial discretion and shift interpretive authority toward algorithmic outputs.

As a result, changes introduced by AI systems may be experienced differently by organizational actors. Managers may perceive a reduction in discretionary judgment over performance evaluation, while employees may interpret increased data visibility as intensified monitoring. Consequently, resistance to AI-enabled performance management may reflect not only cultural or technical misalignment, but also tensions arising from shifting authority, accountability, and control structures embedded within digital performance systems.

3.2. *Theory of Planned Behavior as a Complementary Lens*

The Theory of Planned Behavior (TPB) is used in this study as a complementary behavioral lens to explain the conditions under which AI-driven performance management systems are likely to be accepted and used in practice. Its purpose in this research is not to re-explain telework behavior in general, but to provide an explanatory mechanism for adoption and sustained engagement with AI-enabled performance management systems within teleworking contexts.

TPB posits that behavioral intention is shaped by three key constructs: attitudes toward the behavior, subjective norms (perceived social pressure), and perceived behavioral control (perceived ability to perform the behavior) (Ajzen, 1991). In the context of this study, these constructs are used to explain variation in employees' and managers' willingness to engage with AI-driven performance management systems (Venkatesh et al., 2003).

TPB complements socio-technical theory by addressing a limitation of system design perspectives alone: even well-designed socio-technical systems may fail if users do not accept, trust, or meaningfully engage with them (Davis, 1989; Jangbahadur et al., 2025). While socio-technical theory explains how AI-enabled performance management systems should be structured, TPB explains whether and why actors choose to use them in practice.

In this study, TPB constructs are analytically aligned with the socio-technical subsystems through multi-level mapping. At the individual level, attitudes toward AI-driven performance management and perceived behavioral control correspond to the personnel subsystem, reflecting perceptions of usefulness, trust, digital competence, and self-efficacy (Venkatesh et al., 2012). At the team level, subjective norms reflect shared expectations, peer influence, and coordination pressures that shape collective engagement with AI-based performance systems. At the organizational level, these norms are reinforced through managerial practices, performance culture, and formal evaluation routines that define acceptable behavior. At the environmental level, subjective norms are shaped by broader institutional conditions,

including public-sector accountability requirements, regulatory frameworks, and societal expectations regarding transparency and fairness in workplace monitoring.

This mapping clarifies how behavioral and system-level perspectives are integrated in this study. Socio-technical theory defines how AI-driven performance management systems are structurally embedded within organizational contexts, while TPB explains how individuals and groups respond to these systems in terms of acceptance, resistance, or sustained use. Together, they provide a coherent multi-level explanation of both system design requirements and behavioral adoption dynamics in AI-enabled performance management (Nayak & Jagadeeswari, 2025; Dinh, 2026).

3.3. Integrating the Two Frameworks

The integration of socio-technical theory and the Theory of Planned Behavior (TPB) provides a dual analytical lens for this study. Socio-technical theory defines the design architecture of the AI-driven performance management framework by specifying how technical, organizational, personnel, and environmental subsystems must be aligned. TPB complements this by explaining the behavioral conditions under which these systems are likely to be accepted, used, and sustained, based on attitudes, subjective norms, and perceived behavioral control.

Importantly, this integration highlights that adoption of AI-driven performance management systems is not a purely technical or rational process, but one shaped by organizational context, institutional constraints, and social dynamics across system levels. Even well-designed systems may face resistance if they conflict with established norms, reduce perceived autonomy, or disrupt existing performance evaluation practices.

From a Critical Management Studies (CMS) perspective, such resistance can also be understood as reflecting deeper tensions related to control, visibility, and legitimacy in digitally mediated work environments. AI-enabled performance systems do not only support evaluation; they also contribute to defining what counts as performance by translating work activities into data-driven representations (Kellogg et al., 2020; Zuboff, 2023; Lee et al., 2025). This shift can alter established authority structures by increasing reliance on algorithmically generated metrics rather than managerial judgment.

In this context, resistance may emerge as a response to perceived changes in power and discretion. Managers may experience reduced autonomy in evaluation as performance becomes more standardized and continuously visible through digital systems. Employees, in turn, may interpret these systems as increasing surveillance or reducing interpretive fairness, even when they are designed to improve transparency and consistency. As a result, resistance is better understood not simply as behavioral reluctance, but as a response to shifts in autonomy, accountability, and control embedded in algorithmic management systems.

The empirical findings further reinforce the importance of this integration. While attitudes toward teleworking and subjective norms strongly shape behavioral intentions, perceived behavioral control (e.g., digital skills and access to technology) does not show a significant independent effect. This suggests that adoption of AI-enabled performance management systems is driven more by cultural, normative, and attitudinal conditions than by technical capability alone.

Taken together, these insights have direct implications for framework design. The proposed AI-driven performance management framework must therefore extend beyond technical system capabilities to address the social, organizational, and institutional conditions that shape acceptance, legitimacy, and sustained use.

4. Empirical Foundation: Teleworking in the Canadian Public Service

This section presents the empirical foundation used to derive the design requirements for the proposed AI-driven socio-technical performance management framework. Building on the teleworking and performance management challenges identified in the Introduction and Literature Review, the analysis focuses on how performance is monitored, evaluated, and experienced within teleworking environments in the Canadian public service. Rather than reporting findings separately by method, the empirical evidence is synthesized thematically around key performance management challenges identified in Section 2.2, including visibility, communication, fairness, isolation, and employee well-being. This approach enables the integration of multiple data sources to identify recurring patterns, organizational tensions, and practical requirements relevant to AI-enabled performance management design in teleworking environments.

The empirical evidence is drawn from a sequential mixed-methods study of teleworking in the Canadian public service (Wafa, 2024). Findings are synthesized thematically with reference to the performance management challenges identified in the literature review.

4.1. Research Design and Methods

The research design is explicitly informed by the theoretical frameworks introduced earlier, which guide both data interpretation and framework requirement elicitation. Socio-technical theory serves as the overarching analytical lens by structuring the analysis around the interaction between technological systems, organizational processes, and human behavior within teleworking environments. The Theory of Planned Behavior (TPB) informs the interpretation of adoption-related perceptions, including attitudes toward AI-enabled systems, perceived behavioral control, trust, and subjective norms influencing employee and managerial acceptance. Supporting perspectives from Self-Determination Theory (SDT) and the Job Demands–Resources (JD-R) model are used to interpret themes related to autonomy, motivation, workload, organizational support, strain, and employee well-being. Together, these theoretical perspectives informed the interpretation, synthesis, and translation of empirical findings into socio-technical design requirements for the proposed AI-driven performance management framework.

The study employed a sequential explanatory mixed-methods design (Creswell & Plano Clark, 2018), comprising four complementary approaches that combine quantitative breadth with qualitative depth. In this design, quantitative analyses are used to identify broad patterns, which are then refined, contextualized, and explained through qualitative evidence. Four complementary methods were used, each contributing a distinct analytical layer to the framework development process.

First, a machine learning and big data analysis examined 205,204 tweets from 2022 using RapidMiner, applying sentiment analysis, frequency analysis, and Latent Dirichlet Allocation (LDA) topic modeling to capture broad public discourse on teleworking. Tweets were collected via the Twitter API using a purposive hashtag-based sampling strategy, restricted to users located in Canada and the 2022 time period. Data quality was enhanced through iterative hashtag refinement, preprocessing (e.g., removal of noise, missing values, and text normalization), and validation through sentiment and objectivity checks to ensure relevance and reliability of the dataset. This stage identifies macro-level patterns, including dominant concerns, perceived benefits, and widely discussed success factors, providing an initial mapping of teleworking dynamics.

Second, a document analysis assessed federal, provincial, and territorial government teleworking policies using NVivo and RapidMiner for sentiment and thematic analysis, capturing the employer and policy perspective. A structured document-based sentiment analysis approach was employed, distinguishing between positive and negative policy

orientations at the document level (Haddara et al., 2020), supplemented by frequency analysis to identify recurring policy themes. Documents were selected based on predefined inclusion criteria, including policy relevance, temporal recency, and geographic representation across jurisdictions. To ensure diversity and reduce institutional overrepresentation, exclusion criteria limited the dataset to a maximum of one document per governmental body or author. Additional materials, including parliamentary and legislative committee reports, were identified through keyword searches and analyzed using sentiment classification and thematic coding. Reliability was supported through algorithmic confidence scores generated by RapidMiner (ranging from 0 to 1), cross-validation of outputs between NVivo and RapidMiner, and the use of standardized automated classification procedures. This stage captures the institutional and policy perspective, identifying how teleworking is formally structured, governed, and evaluated within public-sector organizations.

Third, an online survey of 176 Canadian public servants (federal and provincial) employed binary logistic regression, Pearson correlation, and Chi-square analysis to test Theory of Planned Behavior constructs quantitatively. Participants were recruited through a quasi-random sampling strategy using publicly available government directories, with proportional representation across jurisdictions. Screening criteria ensured that only respondents with teleworking experience since March 2020 were included. The final sample combined responses collected via email distribution (Qualtrics) and an online professional community, with comparative checks conducted to assess consistency across subsamples and mitigate potential sampling bias. This stage identifies statistically significant factors associated with effective teleworking arrangements, providing measurable evidence to inform framework design requirements. However, given the modest sample size and the exploratory nature of the analysis, the logistic regression results are reported here as descriptive associations among variables rather than as inferential estimates of predictive odds.

Fourth, six semi-structured interviews with Government of Canada public servants (five managers or policymakers and one employee) across four federal departments provided rich qualitative depth into how teleworking policies and practices are experienced and interpreted in practice. Participants were recruited using purposive sampling from survey respondents to ensure variation in roles and perspectives. Interviews were conducted via Zoom, recorded with consent, transcribed verbatim, and analyzed using a combination of thematic analysis (NVivo) and text-mining techniques (RapidMiner). Data collection proceeded until thematic saturation was reached, whereby no substantively new themes emerged, consistent with established qualitative research guidance (Guest & Johnson, 2006). This stage is critical for understanding contextual nuances, managerial behaviors, and unintended consequences that are not visible in quantitative data.

The sequencing reflects a structured integration logic. The initial quantitative phase (machine learning analysis and survey data) identifies dominant patterns, relationships, and predictors of telework attitudes and preferences. The subsequent qualitative phase (document analysis and interviews) is used to explain these patterns by situating them within organizational policies, managerial practices, and lived employee experiences. Integration occurs at the interpretation stage, where findings are compared and synthesized across data sources to assess convergence, divergence, and complementarity.

The sequencing strategy enabled progressive refinement of findings across analytical stages, moving from broad public discourse to institutional policy perspectives, employee-level quantitative evidence, and finally in-depth qualitative insights into lived teleworking experiences. This triangulation across methods and data sources strengthens the evidentiary base from which the framework's design requirements are derived. For full methodological detail, see (Wafa, 2024).

4.2. What Teleworkers Actually Experience: Key Findings

The findings are organized thematically around the performance management challenges identified in Section 2.2. For each theme, evidence is triangulated across the four methods where possible.

4.2.1. The Performance Visibility Problem

The literature review identified lack of direct oversight and micromanagement risks as central challenges in remote performance management (Section 2.2). The empirical evidence confirms these challenges while revealing an important nuance: employees and managers alike recognize that visibility into work can be achieved without surveillance, and that organizational support matters far more than monitoring.

In the interviews, performance monitoring was cited as a key driver of return-to-office mandates. One participant noted: “People can monitor employees better than if they are working from home. I think that would be the main reason.” Yet another participant offered a counter-narrative: “People now recognize that you do not have to see people to know they are doing work. . . it forces a conversation about efficiency and effectiveness.” This tension between surveillance-based and outcome-based approaches was a recurring theme, with interviewees identifying outcome-oriented management as a critical success factor. Managers who adapted their leadership from presence-based to results-based evaluation were seen as more effective.

The survey data reinforced this pattern. Support emerged as the single strongest predictor of better teleworking arrangements, with a substantial positive association (odds ratio ≈ 6.8 in the descriptive logistic regression), indicative of a strong relationship between perceived organizational support and self-reported teleworking quality. This finding challenges surveillance-first approaches to remote performance management and suggests that the quality of managerial engagement matters more than the quantity of monitoring.

The machine learning analysis corroborated these findings at the population level. In topic modeling of over 205,000 tweets, “productivity” emerged as the primary critical success factor for teleworking (of 30 factors identified), and “support” ranked third. Public discourse centers on output and enablement, not surveillance.

Implication for framework design: There is a documented need for performance management tools that provide visibility into work progress and outcomes without resorting to surveillance—exactly what AI-driven outcome-focused analytics promises.

4.2.2. Communication, Collaboration, and Isolation

The literature review identified communication barriers and employee isolation as significant challenges (Section 2.2). The empirical evidence demonstrates that isolation is not merely a subjective complaint but a measurable risk factor, while socialization functions as a measurable protective factor.

The survey found that isolation had a statistically significant negative effect on teleworking quality (odds ratio = 0.871), while workplace socialization had a strong positive effect (odds ratio = 3.973). These are not trivial findings: socialization increased the odds of better teleworking arrangements by nearly four times, establishing it as one of the most consequential variables in the model.

The interviews added qualitative depth to these statistical patterns. Marital status emerged as an important contextual factor: married or partnered public servants had built-in socialization, while single employees reported greater isolation challenges. As one participant reflected: “I certainly feel for my single colleagues because I am happily married and have two kids. I have more than enough interaction with people.” Technology both helped and hindered communication: video conferencing was described as “better

for many employees because only one person could speak at one time” (more inclusive for those with hearing impairments), but the absence of spontaneous encounters limited serendipitous idea exchange.

In the Twitter analysis, “communication” ranked 25th of 30 critical success factors—relatively low, suggesting it may be taken for granted in public discourse rather than seen as a primary concern. Meanwhile, the document analysis found that every jurisdiction identified communication and collaboration as both a challenge and a success factor for teleworking, underscoring its dual nature.

Implication for framework design: The evidence demonstrates that isolation is a measurable risk factor, socialization is a measurable protective factor, and communication technology is necessary but insufficient. AI-driven sentiment analysis and well-being monitoring could detect isolation and disengagement signals before they become crises.

4.2.3. Fairness, Equity, and Policy Consistency

The literature review discussed AI’s potential to enhance fairness and reduce bias in performance evaluations (Section 2.3.4). The empirical evidence reveals that fairness concerns in telework extend well beyond evaluation bias to encompass systemic policy inconsistencies, gendered impacts, and a striking disconnect between employer rhetoric and employee experience.

In the interviews, fairness was a pervasive concern. The lack of uniform teleworking policies across federal departments created inter-departmental competition for employees. As one participant described: “None of us have the same stories. None of us have the same obligations.” Departments used telework arrangements as recruitment inducements—“What is the teleworking agreement if I am going to your department? Is it full-time? Okay, I will work with you. If not, I am not coming”—creating perceived inequity across the public service.

Gender disparities in teleworking impacts were documented through the interviews. Women bore disproportionate domestic burdens during telework. One participant reported: “All the studies show that women, even those working full time out of the house, do more work at home than their spouses.” A male manager corroborated: “Most of my female co-workers took the brunt of that new role and expectations. Many men returned to the office or had a designated space, for example, in the house, whereas women had to do this in the kitchen or wherever they fit.”

The survey found that education (odds ratio = 2.0) and marital status significantly predicted teleworking quality, suggesting that existing demographic advantages amplify teleworking benefits unevenly. Meanwhile, the document analysis revealed that government teleworking policies were 97–99% positive in sentiment across all jurisdictions—a striking disconnect from the mixed challenges reported by actual employees in the survey and interviews.

Implication for framework design: AI-driven evaluation tools could standardize performance criteria across departments and locations, reduce subjective bias, and flag emerging inequities in workload distribution or recognition patterns—but only if designed with explicit awareness of these documented equity gaps.

4.2.4. Well-Being, Work–Life Balance, and the “Always-On” Problem

The literature review identified sentiment analysis and well-being monitoring as key AI capabilities for telework (Section 2.3.2). The empirical evidence reveals that well-being in telework is not a binary outcome but a dynamic balance that shifts with individual circumstances, organizational policy, and the broader political environment.

The survey found that work–life balance had a positive but modest effect on teleworking quality (odds ratio = 1.13). The interviews provided crucial context for this modest statistical effect by revealing the complexity of well-being dynamics. Teleworking reduced some stressors—commuting, office noise, self-consciousness about breaks—but created new ones, including blurred boundaries, “always-on” culture, and lack of social interaction. One participant captured this duality: “It has been much better for my mental health. My physical health is definitely worse just because I do not really do any exercise.”

The political dimension of well-being also emerged. Mandated return-to-office policies undermined well-being gains built over three years of telework. As one participant described: “We started to develop a way to carry on our work online as much as possible and now we are destroying the culture that has been built.” In the Twitter analysis, “wellness support” ranked 10th and “work–life balance” ranked 12th among 30 critical success factors, indicating these are recognized needs in public discourse. The document analysis found that multiple jurisdictions cited work–life balance as both a benefit and a challenge of teleworking.

Implication for framework design: AI-driven well-being monitoring must be sensitive to this complexity—adaptive, opt-in, and attentive to the specific stressors documented here, rather than relying on crude binary indicators.

4.2.5. The Role of Attitudes, Norms, and Organizational Culture

The theoretical framework (Section 3) posited that TPB constructs would influence telework adoption. The empirical evidence not only confirms this but produces a surprising finding that has direct implications for framework design.

To examine these relationships, a binary logistic regression model was employed, appropriate for predicting a dichotomous outcome (teleworking vs. office preference) (Pallant, 2020). The model satisfies key assumptions, including independence of errors, absence of multicollinearity, and a linear relationship between predictors and the logit of the outcome. A Stepwise Forward (Conditional) approach was used to iteratively identify the most significant predictors while minimizing redundancy.

The survey’s TPB analysis yielded the equation: $\text{Logit} = \beta_0 + 6.293 \text{ Attitudes} + 56.008 \text{ Norms} + 0 \text{ PBC} + \epsilon$. Attitudes ($\beta = 6.293$) and social norms ($\beta = 56.008$) were powerful predictors of teleworking preference. When expressed as odds ratios, these coefficients indicate substantial multiplicative increases in the odds of preferring telework as attitudes and norms become more favorable (odds ratios > 1). The estimated effects are statistically significant ($p < 0.001$), with 95% confidence intervals indicating robust associations; however, the magnitude of the effects—particularly for social norms—should be interpreted with caution, as unusually large odds ratios may reflect scaling effects, model specification choices, or potential overfitting. These results nonetheless reinforce the dominant role of socio-cultural factors in shaping telework preferences.

Model diagnostics indicate a strong overall fit. The omnibus test of model coefficients was statistically significant ($p < 0.001$), confirming that the model provides a better fit than a null model.

Strikingly, perceived behavioral control—encompassing digital skills and technological access—showed no independent effect in multivariate analysis, despite showing bivariate correlations. This means that once attitudes and norms are accounted for, having the technical skills and tools to telework does not independently predict whether an employee prefers teleworking.

It is also important to acknowledge potential sources of bias. The survey sample may be subject to self-selection bias, as individuals with stronger views on teleworking or greater engagement with digital work practices may have been more likely to participate.

This could inflate the observed strength of attitudinal and normative effects. Accordingly, the findings should be interpreted as indicative of strong relationships within the sample rather than definitive estimates of population-level effects.

The interviews reinforced the importance of attitudes: positive attitude was described as “one of the most important factors determining perception and effectiveness” of teleworking. The interview data on organizational culture revealed that teleworking can disrupt established organizational norms and identity but also presents an opportunity to redefine norms and values using technology. The Twitter analysis showed overall positive public sentiment (32% positive versus 4.6% negative, when determinable), but 62% of tweets had indeterminate sentiment—suggesting a large, undecided population whose attitudes could be shaped.

Implication for framework design: The AI-driven framework must address not just technical capabilities but the attitudinal and cultural conditions for adoption. Participatory design, transparency, and demonstrated value are preconditions—not afterthoughts. The empirical findings map systematically to framework design requirements, as shown in Table 1.

Table 1. Mapping of empirical findings onto framework design requirements.

Empirical Finding	Framework Design Requirement
Performance visibility problem: managers need insight without surveillance	AI-driven outcome-focused analytics with transparency safeguards
Isolation as measurable risk; socialization as protective factor	Proactive well-being monitoring with opt-in sentiment analysis
Fairness gaps across departments, demographics, and geography	Standardized, bias-audited evaluation criteria with equity dashboards
Complex well-being dynamics: not binary, con-text-dependent	Adaptive, personalized well-being support with privacy by design
Attitudes and norms drive adoption more than technical skill	Participatory design, demonstrated value, human-centered governance

5. Toward a Socio-Technical, AI-Driven Performance Management Framework

Building on the empirical evidence from Section 4 and the AI capabilities reviewed in Section 2.3, this section proposes an integrated framework for AI-driven performance management in teleworking environments. This framework is developed as a design-oriented synthesis that translates empirical findings and theoretical constructs into actionable requirements for performance management system design in teleworking contexts. Each element of the framework is explicitly tied to documented challenges and evidence-based design requirements. The framework is explicitly derived from empirically identified challenges including performance visibility, employee well-being, perceived inequities, and the central role of organizational culture and translates these into design requirements that are both organizationally actionable and grounded in currently available or near-term AI capabilities (e.g., dashboard analytics, natural language processing, and rule-based decision-support systems), rather than speculative or fully autonomous technologies.

Importantly, in this study, performance management is understood as a socio-technical process involving the setting of expectations, monitoring of work progress, evaluation of outcomes, and feedback provision in digitally mediated teleworking environments. Teleworking refers to work conducted outside traditional office settings through digital communication and collaboration technologies. This positioning ensures that the proposed framework reflects

realistic implementation conditions within contemporary public-sector environments, while remaining consistent with a socio-technical perspective on system design.

More specifically, the framework serves as a bridge between (i) empirical evidence on how teleworking is experienced and evaluated in practice, and (ii) theoretical explanations (socio-technical theory and Theory of Planned Behavior) of how performance systems must be designed and accepted to function effectively. To clarify the relationships among the empirical challenges in teleworking performance management, theoretical foundations, AI-enabled capabilities, governance mechanisms, and intended organizational outcomes, Figure 1 presents an integrated AI-driven socio-technical performance management framework.

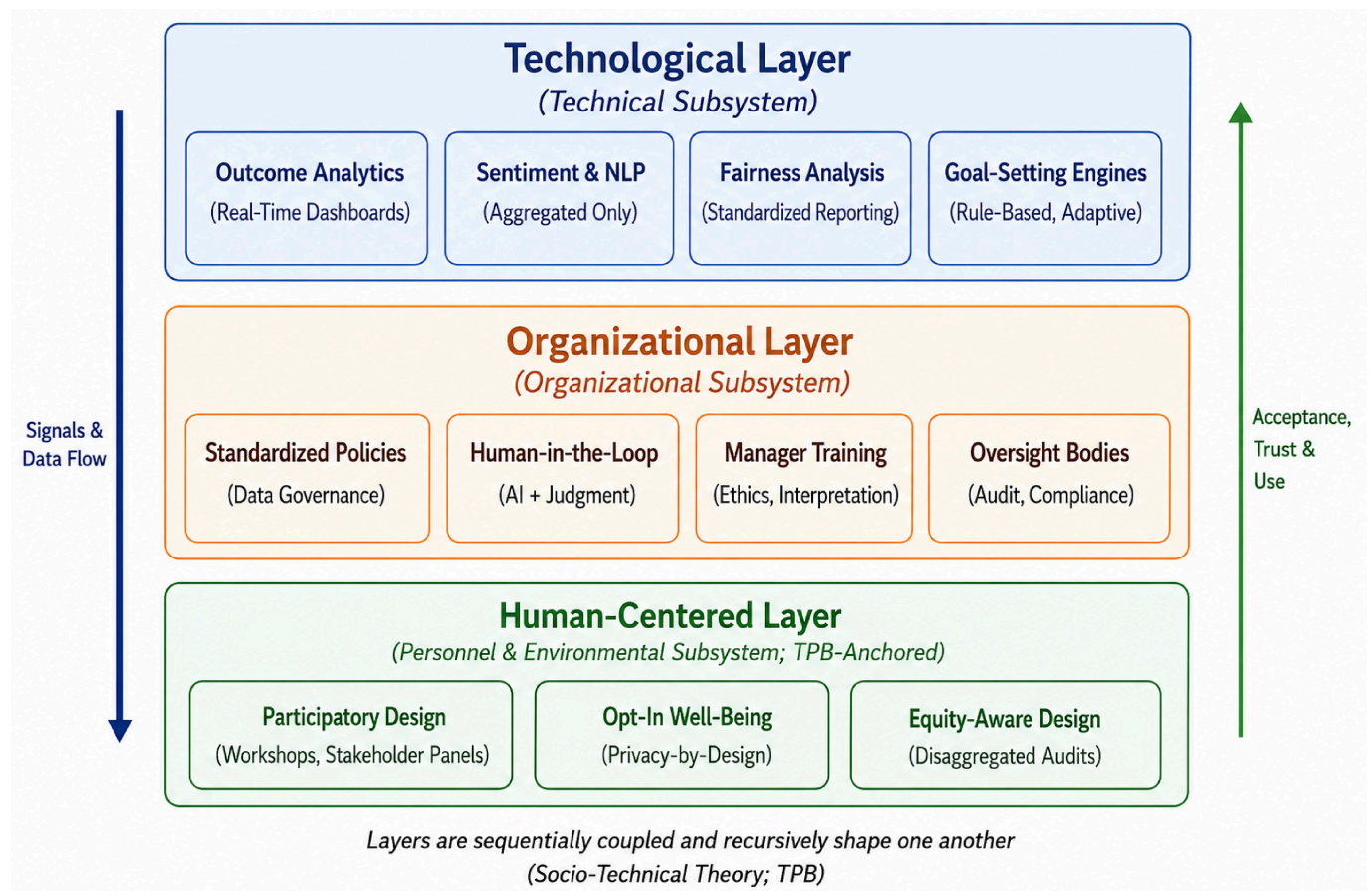


Figure 1. AI-Driven Socio-Technical Performance Management Framework for Teleworking Environments.

The proposed framework operates through an integrated process in which empirically identified teleworking challenges are translated into AI-enabled performance management capabilities informed by organizational and behavioral theories, including socio-technical theory, the Theory of Planned Behavior (TPB), and supporting perspectives from Self-Determination Theory (SDT) and the Job Demands–Resources (JD-R) model. As illustrated in Figure 1, the framework connects empirical evidence from the Canadian public service with these theoretical foundations to guide the design of AI-enabled mechanisms such as adaptive feedback, collaboration analytics, well-being monitoring, and fairness auditing. These capabilities are embedded within governance and human oversight structures intended to ensure transparency, ethical accountability, privacy protection, and trust-centered management. Rather than positioning AI as a standalone replacement for managerial judgment, the framework conceptualizes AI as an enabling layer that augments human decision-making and supports organizational outcomes related to employee well-being, productivity, fairness, engagement, and non-intrusive performance visibility in teleworking environments.

While the proposed framework has not yet undergone full organizational implementation or longitudinal testing, several forms of analytical and theoretical verification were incorporated into the study to strengthen its rigor and alignment with the research question. First, the framework was empirically grounded in a sequential mixed-methods analysis of teleworking experiences within the Canadian public service, including large-scale social media analysis of over 205,000 tweets, document analysis, survey findings, and semi-structured interviews. This ensured that the framework's design requirements were derived from documented workplace challenges rather than hypothetical assumptions. Second, the framework was theoretically verified through alignment with socio-technical theory, the Theory of Planned Behavior (TPB), Self-Determination Theory (SDT), and the Job Demands–Resources (JD-R) model, which collectively explain the organizational, behavioral, motivational, and well-being dynamics associated with teleworking and AI-enabled management systems. Third, the proposed AI-enabled mechanisms were systematically mapped to specific teleworking challenges identified in the literature and empirical findings. For example, adaptive feedback mechanisms address communication and alignment challenges; collaboration analytics support coordination and performance visibility; well-being monitoring responds to stress and isolation concerns; and fairness auditing mechanisms address risks related to bias and inconsistent evaluation. Finally, governance and human oversight structures were incorporated to mitigate risks associated with surveillance, opacity, and excessive algorithmic control. Accordingly, the framework should be understood as a theoretically and empirically informed socio-technical design artifact that demonstrates conceptual and analytical coherence, while recognizing that future pilot implementation, comparative evaluation, and longitudinal testing remain necessary to establish practical effectiveness across diverse organizational contexts.

5.1. Framework Architecture

The framework integrates three interdependent layers, mapped onto the socio-technical subsystems: technological, organizational, and human-centered governance. Importantly, the proposed architecture is conceptualized as a decision-support and augmentation system rather than an autonomous AI decision-making regime, relying on established and currently deployable technologies such as dashboard analytics, natural language processing (NLP), and rule-based decision-support systems. This design choice reflects both the empirical findings of this study and the practical constraints of public-sector implementation, where fully autonomous or experimental AI applications remain limited. Importantly, these layers are not independent components but function as a coupled system in which technological outputs, organizational rules, and human behavioral responses continuously shape one another.

The relationship between layers is therefore sequential and recursive: technological systems generate performance and engagement signals, organizational structures interpret and regulate these signals, and human actors (influenced by attitudes, subjective norms, and perceived behavioral control as defined in the Theory of Planned Behavior) determine the degree of acceptance, trust, and use of the system.

Technological Layer (Technical Subsystem). This layer encompasses the AI-driven tools that form the analytical foundation of the framework, with a clear emphasis on augmenting—rather than replacing—human managerial judgment. Outcome-focused AI analytics respond to the performance visibility problem documented in Section 4.2.1, providing managers with real-time dashboards that aggregate task completion, workflow patterns, and collaboration metrics using data generated through routine digital work systems—without keystroke logging or webcam monitoring.

NLP-based sentiment and engagement monitoring is applied in a limited and aggregated manner, responding to the isolation and well-being evidence from Sections 4.2.2 and 4.2.4, detecting team-level disengagement trends (e.g., tone, frequency) rather than inferring individual psychological states. Fairness-related analysis is operationalized through standardized reporting and comparison mechanisms, respond to the equity gaps documented in Section 4.2.3, applying standardized evaluation criteria across departments and geographies without relying on fully automated decision-making.

Adaptive goal-setting engines is implemented through rule-based or semi-automated systems, respond to the work–life balance complexity from Section 4.2.4, recalibrating individual goals based on workload, team capacity, and contextual circumstances, rather than fully autonomous optimization.

These technological outputs do not directly determine performance decisions; instead, they function as inputs into organizational governance processes where interpretation and contextualization occur.

Organizational Layer (Organizational Subsystem). This layer provides the governance and process infrastructure that determines whether the technological tools produce equitable outcomes. Standardized, transparent performance management policies across departments respond to the fairness and consistency findings from Section 4.2.3, where inter-departmental competition and policy inconsistency created perceived inequity. These policies are further extended to include formal data governance protocols, specifying what types of employee data may be collected, how such data can be used, and clear restrictions prohibiting secondary or unauthorized uses (e.g., disciplinary actions based on well-being or engagement indicators).

Feedback loops integrating AI insights with human judgment (rather than automated decision-making) respond to the nuanced performance evidence from Section 4.2.1, where managers who combined outcome data with relational engagement were most effective. Importantly, these processes are governed by explicit “human-in-the-loop” requirements, ensuring that all consequential decisions—particularly those related to performance evaluation, promotion, or workload allocation—remain subject to managerial review and contextual interpretation.

Manager training and support infrastructure responds to the documented need for new leadership approaches, recognizing that the survey’s strongest predictor was organizational support (odds ratio ≈ 6.8), not monitoring capability. This includes training managers in the interpretation, limitation, and appropriate use of AI-supported dashboards and analytics, ensuring that technological outputs are used as decision-support tools rather than as deterministic evaluation instruments. Training programs also incorporate data ethics and privacy awareness components, equipping managers to understand data boundaries, respect employee consent conditions, and avoid inappropriate inference or overreliance on algorithmically generated insights.

In addition, organizational governance structures include designated oversight mechanisms (e.g., data governance committees or AI oversight boards) responsible for monitoring compliance with privacy standards, reviewing system outputs, and coordinating with audit processes. These structures ensure that data use remains aligned with organizational policy, legal requirements, and employee trust expectations.

Human-Centered Layer (Personnel and Environmental Subsystems). This layer directly reflects Theory of Planned Behavior constructs (attitudes, subjective norms, perceived behavioral control) and ensures that the framework’s design and governance reflect the documented primacy of attitudes, norms, and culture in determining telework outcomes. Participatory design processes involving employees in AI system selection and governance respond to the attitudes and norms findings from Section 4.2.5, where these constructs

overwhelmed perceived behavioral control. This participatory approach is operationalized through structured consultation mechanisms (e.g., workshops, surveys, and stakeholder review panels) rather than algorithmic co-design systems.

Opt-in well-being support with clear privacy safeguards responds to the “always-on” problem from Section 4.2.4, where employees described complex and context-dependent well-being dynamics. These safeguards are grounded in a privacy-by-design approach, including data minimization (collection of only work-relevant indicators), aggregation at the team level, and strict limitations on data use. In particular, well-being analytics are explicitly separated from performance evaluation processes, and no content-level monitoring, behavioral profiling, or individual psychological inference is conducted. These mechanisms are therefore limited to voluntary participation and non-identifiable, aggregated indicators, ensuring that no individual-level behavioral surveillance is undertaken and that employee autonomy and informational boundaries are preserved.

Equity-aware design addressing documented demographic disparities responds to the gender and marital status findings from Section 4.2.3, ensuring that the framework does not reproduce existing inequities. This is operationalized through periodic equity audits and disaggregated reporting of outcomes rather than automated fairness enforcement algorithms, thereby maintaining transparency while avoiding reliance on opaque or fully automated decision-making systems.

5.2. Speculative Possibilities: How the Framework Could Operate

For each AI capability area, this section describes how it could function within the proposed framework, grounding each speculation in the empirical evidence. Importantly, the term “speculative” is used here in a bounded and implementation-oriented sense, referring to feasible extensions of existing decision-support and analytics systems rather than autonomous AI functionality.

Outcome-Focused AI Analytics in Practice. AI systems could rely on existing enterprise reporting and dashboard technologies to aggregate task completion, workflow patterns, and collaboration metrics to give managers real-time performance dashboards—without keystroke logging or webcam monitoring. The empirical evidence supports this: the interviews showed that managers who shifted to outcome-based evaluation were more effective, and the survey found that organizational support (not monitoring) was the strongest predictor of better telework arrangements. Ethical guardrails would include transparency about what data is collected, employee access to their own analytics, and prohibition of punitive use of well-being data.

Proactive Well-being Monitoring. NLP analysis of communication patterns (not content) could detect team-level disengagement trends; opt-in chatbot-based check-ins could support individual well-being assessment. Well-being support functions would be limited to aggregated, anonymised, and opt-in indicators derived from communication metadata patterns (e.g., frequency and responsiveness trends), rather than content-level or psychological inference-based analysis. This ensures alignment with both technical feasibility and privacy constraints. The empirical evidence supports this: isolation was a documented risk factor (odds ratio = 0.871), socialization was protective (odds ratio = 3.973), and interviewees described mental health tradeoffs in nuanced terms that a crude survey would miss. Guardrails would include opt-in participation only, aggregate reporting at the team level rather than individual surveillance, and clear separation between well-being support and performance evaluation.

Fairness-Enhancing AI. Standardized evaluation criteria could be applied consistently across departments and geographies through standardized reporting dashboards and comparative analytics across units; algorithmic audits could flag disparities in recogni-

tion, promotion, or workload distribution. These tools would support the identification of disparities in outcomes such as recognition, workload distribution, and evaluation consistency. The empirical evidence supports this: inter-departmental competition via differential telework policies was documented, gender-based disparities were identified, and the disconnect between government rhetoric (97–99% positive) and employee experience highlights a fairness gap that standardized tools could help close. Guardrails would include regular bias audits with diverse oversight, human review of all consequential decisions, and representation of affected groups in system governance.

Adaptive Goal-Setting. Rule-based performance management AI systems embedded within existing HR and workflow platforms could recalibrate individual goals based on workload, team capacity, and personal circumstances—recognizing that teleworking parents, single employees, and employees in different time zones face different constraints. The survey showed that education, marital status, and work–life balance all influenced teleworking quality; the interviews documented how personal circumstances shaped the experience in ways that one-size-fits-all goal-setting cannot accommodate. Guardrails would include employee agency in goal negotiation, safeguards against algorithmic bias in workload distribution, and transparency in how personal circumstances inform adaptations. Importantly, adaptations would remain managerially mediated rather than algorithmically determined, ensuring that contextual information informs decision-making without introducing automated bias into performance evaluation.

5.3. What This Framework Does Not—And Cannot—Do

An honest assessment of limitations, informed by the empirical evidence, is essential to appropriately situate the scope and applicability of the proposed framework.

First, the survey finding that digital skills showed no independent effect on teleworking quality (despite bivariate correlations) suggests that technology alone does not determine outcomes—organizational and cultural factors dominate. No AI framework can substitute for genuine organizational commitment to supporting teleworkers.

Second, the interviews revealed that political pressure—downtown business impacts, citizen perceptions, union dynamics—drove return-to-office mandates more than performance evidence. No AI framework can override political decision-making, however robust its data.

Third, the documented disconnect between government policy rhetoric (overwhelmingly positive) and employee experience (mixed and complex) suggests that framework adoption requires genuine organizational commitment, not performative endorsement. AI tools deployed in an organizationally hostile environment will not produce the outcomes described here.

Fourth, AI tools cannot substitute for the hallway conversations, spontaneous encounters, and serendipitous idea exchange that some interviewees valued. The framework augments human connection—it does not replace it.

Fifth, and critically, the framework should entail privacy and data governance considerations associated with the use of aggregated workforce analytics. Although the system is explicitly designed to avoid individual-level surveillance, the collection of workplace interaction metadata (e.g., communication frequency, workflow patterns, and task completion signals) introduces legitimate concerns regarding data protection, consent, and the boundaries of acceptable organizational data use. Consistent with the ethical AI literature in public administration, these concerns are not treated as technical side constraints but as socio-technical and governance challenges that emerge through the interaction of institutional practices, organizational infrastructures, and system design choices (Morley et al., 2021, 2020; Mergel et al., 2019; Floridi et al., 2018).

In this respect, the framework does not claim to empirically evaluate or validate privacy outcomes; rather, it articulates normative and design-oriented guidance informed by established responsible AI and algorithmic governance scholarship (e.g., [Raji et al., 2020](#); [Jobin et al., 2019](#)). To address these concerns, the framework adopts privacy-by-design principles, including data minimization, aggregation at team level, strict separation between well-being analytics and performance evaluation, and prohibition of content-level monitoring or behavioral profiling. These safeguards are conceptualized as governance principles that must be enacted through organizational routines, rather than as automatically enforceable technical guarantees.

However, consistent with prior research on the implementation gap in ethical AI, the effectiveness of such safeguards is contingent upon organizational capacity, institutional enforcement, and regulatory context, rather than being guaranteed by system design alone ([Kroll, 2021](#); [Mergel et al., 2019](#)). This reflects the broader socio-technical understanding that ethical outcomes in AI systems are not embedded in technology itself, but are continuously produced and negotiated through practice, governance, and use.

5.4. Phased Implementation Roadmap, KPIs, and Evaluation Strategy

It is important to recognize that no single implementation roadmap can be universally applied across organizational contexts. Variations in institutional mandates, governance structures, levels of digital maturity, and regulatory constraints require that AI-enabled performance management systems be adapted to the specific conditions within which they are deployed. Consistent with socio-technical systems theory, effective implementation depends not only on technological infrastructure, but also on alignment with organizational processes, cultural norms, and accountability structures.

Accordingly, the roadmap proposed here is not intended as a prescriptive or one-size-fits-all model, but rather as a context-sensitive implementation framework grounded in the empirical findings of this study. In particular, the design reflects the demonstrated importance of attitudes, social norms, and organizational support in shaping telework outcomes, while also incorporating insights from emerging research on algorithmic accountability, which emphasizes the need for structured oversight, documentation, and continuous evaluation across the system lifecycle ([Raji et al., 2020](#)).

To ensure feasibility within the public-sector context examined in this research, a streamlined three-phase implementation approach, spanning approximately 12 to 18 months, is proposed, with each phase incorporating explicit governance, documentation, and evaluation mechanisms.

The first phase, pilot and socio-technical alignment (0–4 months), focuses on limited-scale deployment combined with early-stage governance design. Rather than prioritizing technical sophistication, this phase emphasizes trust-building, transparency, and participatory engagement. AI-supported performance dashboards are introduced within a small number of pilot teams, with a focus on validating outcome-based performance metrics as an alternative to activity-based monitoring. In line with audit-based accountability approaches, this phase also includes the development of baseline documentation practices, including clear articulation of system purpose, data inputs, and intended use cases. Co-design workshops involving employees and managers are used to align system functionality with organizational norms, while initial governance protocols establish responsibilities for data stewardship, system oversight, and acceptable use. Evaluation at this stage focuses on user acceptance, perceived usefulness, employee trust, perceived fairness, and the interpretability and reliability of system outputs.

The second phase, controlled expansion and governance integration (4–10 months), involves scaling the system across organizational units while formalizing accountability

structures and oversight processes. Consistent with the finding that organizational support is a key predictor of telework effectiveness, this phase prioritizes managerial capability and institutional alignment. Implementation is extended to departments with comparable work structures, accompanied by the standardization of outcome-based performance indicators. At the same time, governance mechanisms are strengthened through the formal assignment of roles and responsibilities for system monitoring, audit processes, and decision oversight. Human-in-the-loop decision-making is institutionalized to ensure that algorithmic outputs are interpreted within their organizational context, and documentation practices are expanded to include model performance tracking and decision rationales. Performance is evaluated through reductions in ambiguity in performance expectations, increased consistency in evaluations, improvements in perceived organizational support, and the degree of alignment between AI-generated insights and managerial decisions.

The third phase, institutionalization and continuous evaluation (10–18 months), focuses on embedding the framework within routine organizational practices and establishing ongoing audit and accountability processes. At this stage, the emphasis shifts from implementation to sustained monitoring, evaluation, and iterative refinement. AI-supported processes are integrated into formal performance management systems, and continuous monitoring mechanisms are used to detect performance drift, unintended consequences, or emerging inequities. In line with lifecycle governance models, periodic reviews are conducted to assess system impacts on fairness, workload distribution, and employee outcomes, supported by systematic documentation and reporting practices. Cross-unit benchmarking enables organizational learning, while governance bodies retain authority to modify or recalibrate system components as needed. Evaluation in this phase centers on broader organizational outcomes, including employee engagement and retention, stability of telework performance, reduction in equity gaps, and compliance with governance and accountability standards.

Taken together, this phased approach aligns implementation with both socio-technical principles and emerging best practices in algorithmic accountability, ensuring that system deployment is accompanied by structured oversight, transparent documentation, and continuous evaluation rather than treated as a purely technical exercise.

5.5. Algorithmic Accountability and Audit Mechanism

To address concerns related to bias, transparency, accountability, and the governance of sensitive employee data, the proposed framework incorporates a structured algorithmic auditing and governance mechanism grounded in the lifecycle accountability model advanced by [Raji et al. \(2020\)](#). In contrast to approaches that treat auditing as a one-time technical validation, this framework conceptualizes accountability as an ongoing, socio-technical process that is embedded across all stages of system design, deployment, and use, including continuous verification of compliance with data protection and privacy requirements.

Central to this approach is the institutionalization of auditability as an organizational capability, rather than a purely technical feature. This involves establishing clear lines of responsibility for AI system oversight, formalizing documentation practices, and integrating continuous evaluation mechanisms into routine organizational processes. It also includes the formalization of data governance standards specifying permissible data inputs, access controls, retention policies, and boundaries on the use of employee-related data. In line with [Raji et al. \(2020\)](#), the framework emphasizes that effective accountability requires not only technical testing, but also governance structures capable of interpreting, contesting, and acting upon algorithmic outputs, while ensuring that data use remains consistent with privacy-by-design principles and organizational policy constraints.

Operationally, the auditing mechanism is structured across three interrelated stages. First, pre-deployment evaluation focuses on validating data integrity, assessing potential sources of bias, and ensuring that system objectives are aligned with organizational policies and normative expectations. This stage explicitly includes privacy impact assessments, verification of data minimization practices, and confirmation that no sensitive or non-work-related data (e.g., personal communications content or behavioral surveillance data) are incorporated into system inputs. Documentation of model assumptions, data provenance, and intended use cases creates an auditable record of design decisions and data boundaries.

Second, continuous monitoring during deployment tracks system performance over time, including the detection of model drift, inconsistencies in outputs, and emergent unintended consequences. This monitoring also includes ongoing checks for inappropriate data use, scope creep in data collection, and any unintended linkage between well-being indicators and performance evaluation processes. This aligns with the notion of ongoing “post-deployment auditing” highlighted by [Raji et al. \(2020\)](#), recognizing that many risks—particularly those related to privacy and fairness—only become visible in real-world use.

Third, periodic impact assessments evaluate the broader organizational effects of the system, including implications for fairness, workload distribution, and the integrity of performance evaluation processes across different employee groups. These assessments also examine employee perceptions of data use, trust in the system, and potential privacy concerns, ensuring that the framework remains socially and ethically acceptable in practice.

Importantly, the framework embeds human oversight at all stages of the audit process, reflecting the socio-technical premise that accountability cannot be fully automated. Managers and designated oversight bodies are responsible for interpreting algorithmic outputs, particularly in high-stakes or context-dependent decisions, and for ensuring that system recommendations are not applied in a mechanistic or decontextualized manner. This “human-in-the-loop” approach is complemented by formal governance structures that define roles, escalation pathways, and decision rights in cases where algorithmic outputs are contested or produce adverse outcomes, including cases involving potential breaches of data privacy or inappropriate data interpretation.

To ensure institutional accountability, audit findings are systematically reviewed by a designated governance entity (e.g., an AI oversight or ethics committee) with the authority to intervene when necessary. Such interventions may include recalibrating models, restricting system use in specific contexts, or suspending components that fail to meet established fairness, transparency, or data protection thresholds. In addition, the framework supports the maintenance of audit trails and documentation repositories to enable traceability, reproducibility, and external scrutiny where required, including documentation of data handling practices and audit decisions related to privacy compliance.

By embedding auditing within a broader governance architecture, this approach moves beyond compliance-oriented models toward a more robust form of continuous, practice-based accountability, in which technical evaluation, organizational oversight, and ethical considerations are jointly integrated. This design directly addresses the “AI accountability gap” identified by [Raji et al. \(2020\)](#) by ensuring that responsibility for system behavior—and the handling of sensitive data—remains visible, distributed, and actionable throughout the lifecycle of the framework.

5.6. Framework Verification and Illustrative Validation

The proposed AI-driven socio-technical performance management framework was first verified through alignment between empirically identified teleworking challenges, AI-enabled performance management mechanisms, and supporting theoretical foundations. Empirical verification was established by mapping framework components to findings

derived from the Canadian public-service dataset, including machine learning and big data analysis, document analysis, survey findings, and semi-structured interviews. The framework was also theoretically verified through integration with socio-technical theory, the Theory of Planned Behavior (TPB), Self-Determination Theory (SDT), and the Job Demands–Resources (JD-R) model, ensuring conceptual consistency between organizational challenges, behavioral mechanisms, and technological design principles. In addition, internal logical coherence was established through explicit traceability between identified teleworking challenges, corresponding AI-enabled capabilities, governance mechanisms, and intended organizational outcomes related to employee well-being, productivity, fairness, engagement, and non-intrusive performance visibility.

However, verification alone does not establish whether the proposed framework can plausibly operate within realistic teleworking environments. To provide an initial form of artifact validation, the framework is therefore illustrated through the following hypothetical teleworking scenario. While this does not constitute full organizational implementation or longitudinal testing, it demonstrates the framework's intended operational logic and practical applicability within a realistic organizational setting.

Consider a hypothetical Canadian public-sector department—for illustrative purposes, a policy analysis branch of approximately 25 employees distributed across three regional offices, transitioning to a 60/40 hybrid arrangement after operating fully remotely during the pandemic—operating within a hybrid teleworking environment involving knowledge-intensive administrative and policy-related work. Managers report increasing difficulty maintaining performance visibility, ensuring fairness in evaluations, supporting employee well-being, and sustaining collaboration across geographically dispersed teams. Employees simultaneously report communication fragmentation, unclear performance expectations, reduced feedback quality, concerns about excessive monitoring, and growing feelings of isolation and burnout. These challenges reflect the broader tensions identified in both the literature and the empirical findings of this study.

Within the proposed framework, AI-enabled productivity and collaboration analytics are used to identify workflow bottlenecks, coordination gaps, and uneven workload distribution without relying on intrusive surveillance mechanisms such as keystroke logging or continuous webcam monitoring. Adaptive feedback systems continuously align employee goals with organizational objectives while providing personalized developmental recommendations based on workload patterns and task progress. Sentiment analysis and well-being monitoring mechanisms identify indicators of stress, disengagement, or burnout risk through aggregated communication patterns and self-reported well-being indicators, allowing managers to intervene proactively through supportive organizational measures rather than punitive control mechanisms. At the same time, fairness auditing mechanisms evaluate performance assessment patterns across teams to identify potential inconsistencies or biases in evaluation processes. In concrete terms, the productivity dashboard might reveal that one regional team consistently completes briefing notes more quickly than its peers, but with substantially higher subsequent revision rates—prompting a managerial conversation about workload calibration and review processes rather than an automatic comparative ranking. The well-being monitoring layer might flag a sustained six-week decline in aggregated communication tone and response latency within a particular sub-team, triggering an opt-in check-in and a review of workload distribution rather than a performance flag against individual employees. The fairness auditing layer might surface that high-visibility assignments have been disproportionately allocated to staff in the headquarters office relative to regional offices, prompting a review of assignment-allocation practices. In each case, AI-generated signals are routed through governance structures

and managerial interpretation, with action taken by the manager in light of contextual information rather than by the algorithm itself.

Importantly, these AI-enabled mechanisms operate within governance and human oversight structures informed by socio-technical theory and the Theory of Planned Behavior (TPB). Human managerial judgment remains central to interpretation and decision-making, while governance safeguards ensure transparency, privacy protection, explainability, and employee participation in system use. Supporting perspectives from Self-Determination Theory (SDT) and the Job Demands–Resources (JD-R) model further inform the framework's emphasis on balancing accountability with autonomy, organizational support, workload management, and employee well-being.

Through this illustrative application, the framework demonstrates how AI-enabled performance management could support organizational outcomes related to employee well-being, productivity, fairness, engagement, and non-intrusive performance visibility in teleworking environments. Rather than optimizing surveillance or maximizing algorithmic control, the framework prioritizes socio-technical alignment between technological capabilities, organizational governance, and human-centered management practices. In this sense, the scenario-based application demonstrates how the framework is designed to support employee well-being, productivity, and fairness through supportive and adaptive performance management practices rather than through surveillance-oriented activity monitoring alone. Accordingly, the illustrative scenario should be understood as an initial operational demonstration of framework applicability rather than definitive empirical proof of organizational effectiveness. While future pilot implementation, comparative evaluation, and longitudinal testing remain necessary to establish practical effectiveness across diverse organizational contexts, this scenario-based validation demonstrates the framework's operational logic and its potential capacity to address the core challenges identified in the research question.

6. Discussion

6.1. Contributions to Theory and Practice

This paper makes three interrelated contributions. Methodologically, it demonstrates how empirical evidence from one domain—the lived experience of teleworking in the Canadian public service—can ground speculative framework design in another—AI-driven performance management. The approach bridges descriptive research (what is) and prescriptive design (what could be), providing a model for evidence-informed framework development in domains where full implementation data is not yet available.

Theoretically, the integration of socio-technical theory with the Theory of Planned Behavior provides a dual lens that captures both system design and human behavior. The empirical finding that attitudes and norms overwhelm perceived behavioral control has direct implications for implementation strategy: organizations investing in AI-driven performance management should allocate at least as much attention to cultural change, participatory design, and demonstrated value as they do to technical infrastructure.

Practically, the framework offers a structured approach for organizations transitioning from presence-based to outcome-based performance management, with each element tied to documented challenges and evidence-based design requirements rather than aspirational possibilities alone.

6.2. Implications for Policy

Empirical evidence has several implications for public sector teleworking policy. The documented inter-departmental inconsistency in telework policies argues for standardized AI-driven performance management frameworks as an equity measure—not merely an

efficiency tool. The political dynamics of return-to-office mandates suggest that AI-driven performance management tools could provide the evidence base that political decision-makers currently lack, demonstrating that teleworkers are productive without requiring physical presence. However, the evidence also cautions that data alone may not override political imperatives driven by downtown economic concerns or citizen perceptions. Importantly, this evidentiary role is conditional: AI systems do not resolve political conflict but instead reframe it by making competing claims about productivity observable and comparable across organizational units.

More concretely, the influence of AI-generated performance data operates through existing administrative and policy decision-making processes rather than through direct persuasion or causal influence on political actors. In practice, AI systems convert operational work data into structured indicators (e.g., productivity trends, task completion rates, and cross-unit comparisons), which are then integrated into formal organizational artifacts such as executive dashboards, departmental performance reports, and policy briefing documents. These artifacts are routinely used in return-to-office deliberations and related workforce policy decisions. As a result, decision-makers are exposed to standardized and comparable evidence on remote and in-office performance within a shared reporting environment.

By embedding performance information into these official governance channels, AI systems reduce reliance on selective interpretation, anecdotal evidence, or department-specific narratives when assessing telework effectiveness. Instead, they establish a shared evidentiary baseline that constrains how performance claims are constructed and justified in formal discussions. Importantly, this does not eliminate political considerations or override decision-maker preferences; rather, it ensures that such decisions are made within a consistent, system-generated evidence environment that structures how competing arguments are evaluated and compared.

At the same time, this evidence-based visibility may partially counter political pressure by shifting debates from perceptions of telework to comparable, system-generated performance indicators across in-office and remote contexts. However, it is important to emphasize that such data does not eliminate political economy considerations (e.g., downtown economic impacts or public sentiment) but instead reshapes the evidentiary basis upon which these competing pressures are negotiated.

The gendered impacts of teleworking require that any AI-driven performance management framework include equity auditing as a core, not peripheral, feature. The finding that women bore disproportionate domestic burdens during telework means that performance evaluation systems must account for the unequal conditions under which work is performed, rather than assuming uniform productivity environments across employees. Accordingly, the framework must operationalize equity through mechanisms such as disaggregated performance reporting, contextual interpretation of productivity indicators, and the integration of equity-focused audit processes that systematically assess disparities across demographic groups. In practice, this means that AI-supported evaluation systems should avoid uniform performance benchmarks and instead enable managers to interpret outputs in light of differing work contexts (e.g., caregiving responsibilities, flexible schedules, or constrained working hours). This shifts the role of AI from enforcing standardized evaluation to supporting more context-sensitive and equitable decision-making. By embedding these considerations into both the analytical layer (through disaggregated data and reporting) and the governance layer (through periodic equity audits and oversight), the framework ensures that gender-related disparities are not only identified but actively addressed within performance management processes.

Importantly, the strong association between perceived organizational support and improved teleworking outcomes (odds ratio ≈ 6.8) should be interpreted cautiously. While

this suggests a strong statistical relationship, it does not establish causality. One plausible alternative explanation is selection and perception bias, whereby employees who are already more motivated, higher performing, or better integrated into organizational routines are also more likely to report higher levels of perceived support. This limitation is inherent in cross-sectional survey designs. This implies that the observed relationship may reflect underlying organizational and behavioral alignment rather than the causal effect of support alone. Nevertheless, even under this alternative interpretation, the finding remains theoretically meaningful: it reinforces that perceived organizational conditions whether causally driven or perceptually reinforced are more strongly associated with telework outcomes than surveillance intensity or technological monitoring capacity. This supports the broader socio-technical argument of the paper that organizational context mediates the effectiveness of AI-enabled performance systems.

6.3. Limitations and Future Research

Several limitations should be acknowledged. The empirical evidence is drawn from the Canadian public service, and generalizability to private sector organizations, other countries, or other levels of government requires further study. The AI capabilities described in Section 5 are speculative extrapolations from current technology—real-world implementation may reveal unexpected challenges in areas such as algorithmic bias, user acceptance, and organizational resistance. The survey response rate (6.4%) and interview sample ($n = 6$) are modest; larger-scale studies are needed to validate the design requirements derived here, though the triangulation across four methods partially mitigates this limitation. The Twitter data predates the platform's transformation under new ownership, and the applicability of social media sentiment analysis may have shifted since data collection. Importantly, the observed association between perceived organizational support and teleworking effectiveness may reflect endogeneity or reverse causality, whereby pre-existing employee motivation, engagement, or performance levels shape perceptions of support rather than the reverse. This limitation is inherent in cross-sectional designs and reinforces the need to interpret the strong observed association between perceived organizational support and teleworking outcomes as correlational rather than causal.

These limitations are not peripheral but directly inform the interpretation and application of the proposed framework. In particular, they reinforce the paper's central argument that AI-enabled performance management systems operate within broader organizational and socio-technical contexts, and that their effectiveness cannot be inferred from technological capability alone.

Future research should pilot specific elements of the proposed framework in organizational settings to test feasibility and impact. Longitudinal studies tracking the introduction of AI-driven performance management tools in telework environments would provide crucial evidence about adoption dynamics, unintended consequences, and the sustainability of initial benefits. Comparative studies across public and private sectors, and across national contexts, would strengthen the generalizability of the framework's design principles. In addition, replication on larger samples using a priori-specified models and penalized-likelihood estimators (such as Firth's bias-reduced logistic regression) would address small-sample limitations in the present logistic regression analysis and provide more stable inferential estimates of the relationships observed here.

7. Conclusions

This paper has argued that AI-driven performance management for teleworkers must be grounded in evidence about what teleworkers actually experience and need—not merely in what technology makes possible. By integrating a comprehensive literature review on AI

capabilities with empirical evidence from a mixed-methods study of the Canadian public service, and by combining socio-technical theory with the Theory of Planned Behavior, the paper has proposed a framework in which each design element responds to documented challenges: the performance visibility problem, isolation and communication barriers, fairness and equity gaps, well-being complexity, and the primacy of attitudes and norms. Rather than claiming universal applicability, the framework is intended as a contextually grounded and theoretically informed design proposition that requires further empirical validation.

The empirical evidence offers a clear message. Organizational support matters more than monitoring (odds ratio ≈ 6.8). Socialization protects while isolation harms (odds ratios = 3.973 and 0.871, respectively). Attitudes and social norms overwhelm technical skills in predicting telework outcomes (TPB equation: $\beta = 6.293$ for attitudes, $\beta = 56.008$ for norms, $\beta = 0$ for perceived behavioral control). However, these relationships should be interpreted cautiously: the association between perceived organizational support and performance outcomes may reflect reverse causality or self-selection effects, whereby more motivated or higher-performing employees are also more likely to perceive higher levels of support. Government rhetoric runs far ahead of employee experience (97–99% positive policy sentiment versus mixed lived reality). And political dynamics constrain what any framework can achieve.

The proposed framework takes these findings as contextually bounded design inputs rather than universal prescriptions. It envisions AI not as a surveillance mechanism but as a support system—one that provides performance visibility without intrusion, detects well-being risks without violating privacy, standardizes evaluation without erasing context, and adapts to individual circumstances without reinforcing inequity. In particular, the framework integrates equity as a core design requirement, recognizing that telework experiences and performance outcomes are shaped by structurally unequal conditions, including the disproportionate domestic burdens observed among women. Drawing on intersectional perspectives (Crenshaw, 1991) and AI fairness scholarship, including empirical evidence showing that algorithmic systems can reproduce bias in real-world applications (Buolamwini & Gebru, 2018) as well as broader critiques of fairness as shaped by social and institutional power relations (e.g., AI Now reports on bias, opacity, and accountability gaps), the key implication is that fairness cannot be treated as a purely technical problem that can be solved by adjusting algorithms alone. Instead, fairness must be understood as something that emerges from both technology and organizational context. This means that performance management systems must go beyond technical fixes and include disaggregated analysis of outcomes across groups, context-sensitive evaluation of performance, and continuous equity auditing over time to identify and address unequal impacts as they arise during system use. In practical terms, this shifts fairness from a one-time technical validation of the model to an ongoing governance process, where fairness is continuously monitored, interpreted, and corrected throughout system design, deployment, and organizational use.

At the same time, the framework's scope is deliberately constrained. The findings do not establish causal relationships, nor do they generalize beyond the institutional and cultural conditions of the Canadian public sector. More broadly, the analysis reinforces that technology alone cannot solve problems that are fundamentally organizational, cultural, and political. These limitations are not peripheral but central to the interpretation of the findings, underscoring that the framework should be understood as a theoretically informed and empirically grounded starting point rather than a validated solution.

The fundamental question underlying this paper, asking “how are your employees doing?” rather than “what are your employees doing?”, is not merely rhetorical. It captures a fundamental reorientation in performance management philosophy, from control to

care, from surveillance to support, from compliance to development. However, the extent to which AI-enabled systems can meaningfully contribute to this shift remains an open empirical question. Future research should therefore prioritize (a) longitudinal validation of the framework in organizational settings, (b) cross-sectoral comparisons between public and private organizations, and (c) comparative international studies to assess how institutional context shapes both adoption and outcomes. AI-driven tools, thoughtfully designed within a socio-technical framework and grounded in empirical evidence, can help organizations make this transition. But this can only occur if the humans who design, deploy, and govern these tools commit to the same reorientation. Future research should focus on pilot implementation, user acceptance testing, and longitudinal evaluation of the proposed framework across organizational contexts to assess its practical effectiveness, unintended consequences, and scalability.

Author Contributions: Conceptualization, Y.W. and J.L.; methodology, Y.W.; software, Y.W.; validation, Y.W.; formal analysis, Y.W.; investigation, Y.W.; resources, Y.W.; data curation, Y.W.; writing—original draft preparation, Y.W.; writing—review and editing, J.L.; supervision, J.L.; project administration, Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Behavioural Research Ethics Board, University of Saskatchewan (protocol code 3869 approved 29 March 2023).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data available on request from the authors after publication.

Acknowledgments: The authors would like to express their gratitude to colleagues (Ken Coates, Carey Doberstein, Carl Gutwin, and Tarun Katapally) who offered valuable comments on an earlier work that preceded this manuscript. The insights provided by colleagues and their constructive criticism contributed to this present paper. During the preparation of this manuscript, the authors used Claude (Anthropic, claude.ai, [model version 1.1.4498], accessed March 2026) for the purposes of final review editorial suggestions on a completed draft of this manuscript. The authors have reviewed and edited the output and take full responsibility for the content of this publication. The comments of two anonymous reviewers were instrumental in shaping the final version of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
COVID-19	Coronavirus Disease 2019
EPM	Electronic Performance Monitoring
HR	Human Resources
IT	Information Technology
NLP	Natural Language Processing
ONA	Organizational Network Analysis

References

- Aguinis, H. (2009). *Performance management* (2nd ed.). Prentice Hall/Pearson Education.
- Aguinis, H., & Burgi-Tian, J. (2021). Talent management challenges during COVID-19 and beyond: Performance management to the rescue. *BRQ Business Research Quarterly*, 24(3), 233–240. [[CrossRef](#)]
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [[CrossRef](#)]
- Alvesson, M. (2004). *Knowledge work and knowledge-intensive firms*. Oxford University Press.

- Asana. (2025a). *How to streamline performance reviews with Asana AI: Tips from Asana's head of digital*. Available online: <https://asana.com/resources/performance-reviews-ai> (accessed on 1 January 2026).
- Asana. (2025b). *Is micromanagement holding you back?* Available online: <https://asana.com/resources/micromanagement> (accessed on 1 January 2026).
- Asana. (2025c). *Workflows and automation feature*. Available online: <https://asana.com/features/workflow-automation> (accessed on 1 January 2026).
- Ashdown, L. (2018). *Performance management: A practical introduction* (Vol. 16). Kogan Page Publishers.
- Bailey, D. E., & Kurland, N. B. (2002). A review of telework research: Findings, new directions, and lessons for the study of modern work. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 23(4), 383–400. [CrossRef]
- Baki, N. U., Rasdi, R. M., Krauss, S. E., & Omar, M. K. (2023). Integrating artificial intelligence in human resource functions: Challenges and opportunities. *International Journal of Academic Research in Business and Social Sciences*, 13(8), 1262–1277.
- Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of Managerial Psychology*, 22(3), 309–328. [CrossRef]
- Ball, K. (2010). Workplace surveillance: An overview. *Labour History*, 51(1), 87–106. [CrossRef]
- BambooHR. (2025). *AI summarized feedback*. Available online: <https://www.bamboohr.com/product-updates/ai-summarized-feedback> (accessed on 1 November 2025).
- Bernstein, E., Blunden, H., Brodsky, A., Sohn, W., & Waber, B. (2020). The implications of working without an office. Special issue on the new reality of WFH. In *Harvard business review: The big idea*. Harvard Business School Publishing.
- BetterUp. (2025a). *Running on empty: The workforce performance crisis and how to fix it*. Available online: <https://www.betterup.com/blog/workforce-performance-crisis> (accessed on 1 June 2026).
- BetterUp. (2025b). *The most comprehensive coaching platform* | BetterUp. Available online: <https://www.betterup.com> (accessed on 14 June 2025).
- Bélanger, F., Watson-Manheim, M. B., & Swan, B. R. (2013). A multi-level socio-technical systems telecommuting framework. *Behaviour & Information Technology*, 32(12), 1257–1279.
- Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J., & Shadbolt, N. (2018). 'It's reducing a human being to a percentage' perceptions of justice in algorithmic decisions. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (pp. 1–14). ACM.
- Bloom, N., Liang, J., Roberts, J., & Ying, Z. J. (2015). Does working from home work? Evidence from a Chinese experiment. *Quarterly Journal of Economics*, 130(1), 165–218. [CrossRef]
- Bostrom, R. P., & Heinen, J. S. (1977). MIS problems and failures: A socio-technical perspective. Part I: The causes. *MIS Quarterly*, 17–32. [CrossRef] [PubMed]
- Boyne, G. A. (2002). Public and private management: What's the difference? *Journal of Management Studies*, 39(1), 97–122. [CrossRef]
- Braesemann, F., Stephany, F., Teutloff, O., Kässi, O., Graham, M., & Lehdonvirta, V. (2022). The global polarisation of remote work. *PLoS ONE*, 17(10), e0274630. [CrossRef]
- Brynjolfsson, E., Horton, J. J., Ozimek, A., Rock, D., Sharma, G., & TuYe, H. Y. (2020). *COVID-19 and remote work: An early look at US data* (No. w27344). National Bureau of Economic Research.
- Buckingham, M., & Goodall, A. (2015). Reinventing performance management. *Harvard Business Review*, 93(4), 40–50.
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency* (pp. 77–91). PMLR.
- Callahan, C., Kimber, J., Hu, E., Tanner, L., & Kunkle, S. (2024). The real-world impact of app-based mindfulness on headspace members with moderate and severe perceived stress: Observational study. *JMIR mHealth and uHealth*, 12(1), e52968. [CrossRef]
- Clockwise. (2023). *Smart calendar assistant for teams*. Available online: <https://www.getclockwise.com> (accessed on 1 January 2024).
- Cosa, M., & Torelli, R. (2024). Digital transformation and flexible performance management: A systematic literature review of the evolution of performance measurement systems. *Global Journal of Flexible Systems Management*, 25(3), 445–466. [CrossRef]
- Cowgill, B. (2018). *Bias and productivity in humans and algorithms: Theory and evidence from resume screening* (Vol. 29, pp. 679–681). Columbia Business School, Columbia University.
- Crenshaw, K. (1991). Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review*, 43(6), 1241–1299. [CrossRef]
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.
- Culture Amp. (2025). *AI in HR: Powering the future of HR with AI* | Culture Amp. Culture Amp. Available online: <https://www.cultureamp.com/blog/ai-in-hr> (accessed on 1 January 2026).
- Dastin, J. (2018). *Amazon scraps secret AI recruiting tool that showed bias against women*. Reuters. Available online: <https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/> (accessed on 1 November 2025).

- Davenport, M. K., & Beier, M. E. (2020). Who has the option to age successfully at work? Considering nonwork factors. *Industrial and Organizational Psychology*, 13(3), 366–368. [CrossRef]
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. [CrossRef]
- Deloitte. (2023). *Workplace monitoring and the lack of trust in the workplace*. Deloitte Insights. Available online: <https://www.deloitte.com/us/en/insights/topics/leadership/workplace-monitoring-and-the-lack-of-trust-in-the-workplace.html> (accessed on 1 January 2025).
- De Vries, H., Tummers, L., & Bekkers, V. (2019). The benefits of teleworking in the public sector: Reality or rhetoric? *Review of Public Personnel Administration*, 39(4), 570–593. [CrossRef]
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147–160. [CrossRef]
- Dinh, V. H. (2026). Leveraging AI-enabled knowledge capabilities for new product performance: Evidence from manufacturing SMEs in emerging economies. *Industrial Marketing Management*, 133, 84–103. [CrossRef]
- Dutta, D., & Mishra, S. K. (2025). Artificial intelligence-based virtual assistant and employee engagement: An empirical investigation. *Personnel Review*, 54(3), 913–934. [CrossRef]
- Eurofound and the International Labour Office. (2017). *Working anytime, anywhere: The effects on the world of work*. Publications Office of the European Union, Luxembourg, and the International Labour Office.
- European Commission. (2019). *Ethics guidelines for trustworthy AI*. European Commission Publications Office.
- Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health*, 4(2), e7785. [CrossRef]
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. [CrossRef]
- Gajendran, R. S., & Harrison, D. A. (2007). The good, the bad, and the unknown about telecommuting: Meta-analysis of psychological mediators and individual consequences. *Journal of Applied Psychology*, 92(6), 1524–1541. [CrossRef]
- Gibbs, M., Mengel, F., & Siemroth, C. (2023). Work from home and productivity: Evidence from personnel and analytics data on information technology professionals. *Journal of Political Economy Microeconomics*, 1(1), 7–41. [CrossRef]
- GitLab. (2021). *The remote work report 2021 | The GitLab handbook*. Available online: <https://handbook.gitlab.com/handbook/company/culture/all-remote/remote-work-report/> (accessed on 9 July 2025).
- Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627–660. [CrossRef]
- Goodings, L., Ellis, D., & Tucker, I. (2024). *Understanding mental health apps: An applied psychosocial perspective*. Springer Nature.
- Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? *Field Methods*, 18, 59–82. [CrossRef]
- Haddara, M., Hsieh, J., Fagerström, A., Eriksson, N., & Sigurðsson, V. (2020). Exploring customer online reviews for new product development: The case of identifying reinforcers in the cosmetic industry. *Managerial and Decision Economics*, 41(2), 250–273. [CrossRef]
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83; discussion 83–135. [CrossRef] [PubMed]
- Hossain, S., Fernando, M., & Akter, S. (2025). Digital leadership: Towards a dynamic managerial capability perspective of artificial intelligence-driven leader capabilities. *Journal of Leadership & Organizational Studies*, 32(2), 189–208. [CrossRef]
- Hu, D., Dang, Y., & Yue, X. (2023). The Effect of Trust-based Management Strategy on Performance of Human-Machine Collaborative Team: A Dynamic Computational Model. *Procedia Computer Science*, 221, 710–717. [CrossRef]
- Humanyze. (2025). *AI-powered workforce analytics for enhanced organizational health*. Humanyze. Available online: <https://humanyze.com/> (accessed on 1 January 2026).
- Inkster, B., Sarda, S., & Subramanian, V. (2018). An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: Real-world data evaluation mixed-methods study. *JMIR mHealth and uHealth*, 6(11), e12106. [CrossRef] [PubMed]
- Jangbahadur, U., Ahlawat, S., Rozera, P., & Gupta, N. (2025). The effect of AI-enabled HRM dimensions on employee engagement and sustainable organisational performance: Fusion skills as a moderator. In *Evidence-based HRM: A global forum for empirical scholarship* (Vol. 13, No. 1, pp. 85–107). Emerald Publishing Limited.
- Jarrahi, M. H., Newlands, G., Lee, M. K., Wolf, C. T., Kinder, E., & Sutherland, W. (2021). Algorithmic management in a work context. *Big Data & Society*, 8(2), 20539517211020332. [CrossRef]
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1, 389–399. [CrossRef]
- John, B., Alsamarra'i, Z., & Panteli, N. (2022). Enhancing employee experience in the era of hybrid work: The case of Microsoft viva. *IEEE Software*, 40(2), 70–79. [CrossRef]

- Kalischko, T., & Riedl, R. (2021). Electronic performance monitoring in the digital workplace: Conceptualization, review of effects and moderators, and future research opportunities. *Frontiers in Psychology*, 12, 633031. [CrossRef] [PubMed]
- Kellogg, K. C., Valentine, M., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *The Academy of Management Annals*, 14(1), 366–410. [CrossRef]
- Kniffin, K. M., Narayanan, J., Anseel, F., Antonakis, J., Ashford, S. P., Bakker, A. B., Bamberger, P., Bapuji, H., Bhawe, D. P., Choi, V. K., Creary, S. J., Demerouti, E., Flynn, F. J., Gelfand, M. J., Greer, L. L., Johns, G., Kesebir, S., Klein, P. G., Lee, S. Y., . . . Vugt, M. V. (2021). COVID-19 and the workplace: Implications, issues, and insights for future research and action. *American Psychologist*, 76(1), 63–77. [CrossRef] [PubMed]
- Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121(4), 1133–1165.
- Kroll, J. A. (2021). Outlining traceability: A principle for operationalizing accountability in computing systems. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 758–771). Association for Computing Machinery.
- Kulkarni, K., Arati, B., Mitra, H., Veeresh, K., Asha, A., Shweta, S., & Shiba, S. (2024). Enhancing employee engagement and productivity with AI. *International Journal of Business and Management Invention*, 13(9), 188–193. [CrossRef]
- Langer, M., Oster, D., Speith, T., Hermanns, H., Kästner, L., Schmidt, E., Sesing, A., & Baum, K. (2021). What do we want from Explainable Artificial Intelligence (XAI)?—A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research. *Artificial Intelligence*, 296, 103473. [CrossRef]
- Lee, J. L., Billovičs, C., Chen, S. Y., Wickham, R. E., Kocher, B., Chen, C. E., & Lungu, A. (2025). Using machine learning to match clients and therapy providers: Evaluating clinical quality and cost of care. *Value in Health*, 28(9), 1327–1334. [CrossRef] [PubMed]
- Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Schafheitle, S., Wildhaber, I., & Kasper, G. (2019). The challenges of algorithm-based HR decision-making for personal integrity. *Journal of Business Ethics*, 160, 377–392. [CrossRef]
- Mabaso, C. M., & Manuel, N. (2024). Performance management practices in remote and hybrid work environments: An exploratory study. *SA Journal of Industrial Psychology*, 50(1), 13. [CrossRef]
- Madanchian, M., Taherdoost, H., Vincenti, M., & Mohamed, N. (2024). Transforming leadership practices through artificial intelligence. *Procedia Computer Science*, 235, 2101–2111. [CrossRef]
- Mancl, D., & Fraser, S. D. (2023). The future of work: Agile in a hybrid world. In *International conference on agile software development* (pp. 63–70). Springer Nature.
- Mäntylä, M. V., Graziotin, D., & Kuuttila, M. (2018). The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Computer Science Review*, 27, 16–32. [CrossRef]
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, 54(6), 1–35. [CrossRef]
- Mergel, I., Edelmann, N., & Haug, N. (2019). Defining digital transformation: Results from expert interviews. *Government Information Quarterly*, 36(4), 101385. [CrossRef]
- Microsoft. (2021). *Wellbeing and productivity tools | Microsoft viva insights*. Available online: <https://www.microsoft.com/en-us/microsoft-viva/insights> (accessed on 9 July 2025).
- Microsoft. (2025a). *Bot messages with AI-generated content—teams*. Available online: <https://learn.microsoft.com/en-us/microsoftteams/platform/bots/how-to/bot-messages-ai-generated-content> (accessed on 9 July 2025).
- Microsoft. (2025b). *Fostering employee wellbeing and improving productivity at Microsoft with Microsoft viva insights*. Available online: <https://www.microsoft.com/insidetrack/blog/fostering-employee-wellbeing-and-improving-productivity-at-microsoft-with-microsoft-viva-insights/> (accessed on 1 January 2026).
- Microsoft. (2025c). *What's new in Microsoft edge*. Available online: <https://www.microsoft.com/en-us/edge/update/137> (accessed on 14 June 2025).
- Mkhize, S., & Lourens, E. (2025). Managing the workforce in the era of digital transformation and remote work. In *Advances in strategic management and leadership*. IntechOpen.
- Morley, J., Elhalal, A., Garcia, F., Kinsey, L., Mökander, J., & Floridi, L. (2021). Ethics as a service: A pragmatic operationalisation of AI ethics. *Minds and Machines*, 31(2), 239–256. [CrossRef] [PubMed]
- Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From what to how: An initial review of publicly available AI ethics tools. *AI and Ethics*, 1, 185–197.
- Nayak, P. B., & Jagadeeswari, I. U. (2025). Organizational performance optimization: An AI-enabled framework for employee compensation and benefits management. *Research in Globalization*, 12, 100328. [CrossRef]
- Ouchi, W. G. (1979). A conceptual framework for the design of organizational control mechanisms. In C. Emmanuel, D. Otley, & K. Merchant (Eds.), *Readings in accounting for management control*. Springer. [CrossRef]
- Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM*. Open University Press.
- Pulakos, E. D., Hanson, R. M., Arad, S., & Moye, N. (2015). Performance management can be fixed: An on-the-job experiential learning approach for complex behavior change. *Industrial and Organizational Psychology*, 8(1), 51–76. [CrossRef]

- Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 469–481). Association for Computing Machinery.
- Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, 46(1), 192–210. [CrossRef]
- Raji, I., Smart, A., White, R., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 33–44). Association for Computing Machinery. [CrossRef]
- Reclaim.ai. (2025). Reclaim—AI calendar for work & life. Available online: https://reclaim.ai/?utm_agid=173478810386&creative=769685790448&device=c&placement=&gad_source=1&gad_campaignid=22153645674 (accessed on 1 January 2026).
- Rockmann, K. W., & Pratt, M. G. (2015). Contagious offsite work and the lonely office: The unintended consequences of distributed work. *Academy of Management Discoveries*, 1(2), 150–164. [CrossRef]
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. [CrossRef]
- Salganik, M. J. (2018). *Bit by bit: Social research in the digital age*. Princeton University Press.
- Shrestha, Y. R., Ben-Menahem, S. M., & Von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review*, 61(4), 66–83. [CrossRef]
- Statistics Canada. (2021). *Working from home: Productivity and preferences*. Available online: <https://www150.statcan.gc.ca/> (accessed on 1 January 2024).
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24–54. [CrossRef]
- Treasury Board of Canada Secretariat. (2019). *Directive on automated decision-making*. Government of Canada.
- Trist, E. (1993). Introduction to volume II. In E. Trist, H. Murray, & B. Trist (Eds.), *The social engagement of social science, Volume 2: A Tavistock anthology—The socio-technical perspective* (pp. 36–60). University of Pennsylvania Press. Available online: <http://www.jstor.org/stable/j.ctt1bj4q98.5> (accessed on 29 May 2023).
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. [CrossRef]
- Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(1), 157–178. [CrossRef]
- Wafa, Y. (2024). *Dual perspectives on better teleworking in the public sector: A mixed methods approach* [Doctoral dissertation, College of Graduate and Postdoctoral Studies, Johnson Shoyama Graduate School of Public Policy, University of Saskatchewan]. Available online: <https://harvest.usask.ca/items/076a4781-b369-49f4-af6e-0ce6406f1b7f> (accessed on 1 January 2025).
- Wang, B., Liu, Y., Qian, J., & Parker, S. K. (2021). Achieving effective remote working during the COVID-19 pandemic: A work design perspective. *Applied Psychology*, 70(1), 16–59. [CrossRef] [PubMed]
- Wang, J., & Panesar, A. (2022). Machine learning based lattice generation method derived from topology optimisation. *Additive Manufacturing*, 60, 103238. [CrossRef]
- Yanamala, K. K. R. (2023). Transparency, privacy, and accountability in AI-enhanced HR processes. *Journal of Advanced Computing Systems*, 3(3), 10–18. [CrossRef]
- Zuboff, S. (2023). The age of surveillance capitalism. In *Social theory re-wired* (pp. 203–213). Routledge.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.