

SafeVoice-AI-Release: An Artificial Intelligence-Based Early Intervention Ecosystem for Workplace Psychosocial Risk Management

Nuruaslizawati Ayob, Siti Hajar Abdul Rauf

Universiti Sultan Zainal Abidin

Corresponding Author Email: hajarrauf@unisza.edu.my

DOI Link: <http://dx.doi.org/10.6007/IJARBSS/v16-i9/28926>

Published Date: 27 September 2026

Abstract

Workplace psychosocial risks, including work stress, burnout, workplace bullying, organizational conflict, and psychological distress, have become increasingly prevalent across contemporary organizations. Despite growing awareness of these issues, many employees remain reluctant to report psychosocial concerns due to fear of retaliation, stigmatization, or disclosure of their identities. Consequently, organizations often fail to identify psychosocial risks at an early stage, resulting in adverse outcomes for employee well-being and organizational performance. This article introduces SafeVoice-AI-Release, an artificial intelligence-based early intervention ecosystem designed to facilitate anonymous reporting, psychosocial risk screening, automated risk analysis, and evidence-based intervention recommendations. The system integrates the SafeVoice Psychosocial Screening Inventory (SPSI), an AI Analytics Engine, a Predictive Risk Scoring Model, and organizational dashboards within a unified digital platform. The innovation was developed using a design and development methodology involving needs assessment, system blueprinting, user journey mapping, AI architecture development, system architecture modelling, and prototype design. SafeVoice-AI-Release is designed to enable employees to report concerns anonymously while allowing organizations to identify psychosocial trends, classify risk levels, and implement targeted interventions. The proposed system contributes to workplace well-being by shifting organizational responses from reactive problem management to proactive prevention and early intervention. The innovation demonstrates the potential of artificial intelligence to support sustainable workplace well-being and organizational resilience through data-driven decision-making.

Keywords: Artificial Intelligence, Psychosocial Risk, Workplace Well-Being, Anonymous Reporting, Early Intervention

Introduction

Contemporary workplaces are being reshaped by digitalisation, technological innovation, changing employment arrangements, intensified performance expectations, and increasingly

blurred boundaries between work and personal life. Although these transformations have created new opportunities for flexibility, productivity, and organizational efficiency, they have also altered the psychosocial conditions under which employees work. Work stress, burnout, workplace bullying, interpersonal conflict, and declining psychological well-being are no longer adequately understood as isolated individual difficulties; rather, they increasingly reflect the interaction between employees and the organizational, relational, and technological environments in which work takes place (Backhaus et al., 2024; Deady et al., 2024; Gimpel et al., 2024). The World Health Organization (WHO) (2022) similarly positions mental health at work as an organizational and public health concern, emphasising that working conditions can either protect or undermine employee well-being. Depression and anxiety alone are estimated to result in approximately 12 billion lost working days annually worldwide, demonstrating that workplace mental health has consequences not only for individual employees but also for organizational productivity and wider socioeconomic sustainability (World Health Organization & International Labour Organization, 2022).

This changing understanding of workplace mental health reflects a wider shift in contemporary social science from approaches that primarily locate stress and psychological difficulty within individuals towards perspectives that recognise the organizational and social production of psychosocial risk. Employees' well-being is shaped not only by their capacity to cope with demands, but also by workload, job control, organizational support, interpersonal relations, fairness, communication, workplace culture, and access to mechanisms through which concerns can be expressed and addressed (Amoadu et al., 2024; Deady et al., 2024). Accordingly, contemporary psychosocial risk management increasingly emphasises organizational prevention rather than relying predominantly on individual coping strategies after distress has already developed. This orientation is reflected in the WHO Guidelines on Mental Health at Work, which recommend organizational interventions that identify and modify psychosocial risk factors within working environments (WHO, 2022). In Malaysia, this policy direction has become particularly salient following the introduction of the *Guidelines on Psychosocial Risk Assessment and Management at the Workplace 2024*, which explicitly emphasise organizational screening, risk control, record keeping, and continuing monitoring of psychosocial conditions (Department of Occupational Safety and Health, 2024; Abd Shukor & Khalid, 2025).

Yet the presence of formal policies or workplace support mechanisms does not necessarily mean that psychosocial concerns become visible to organizations. One of the central problems lies in the relationship between employee voice and employee silence. Employees may recognise harmful conditions but nevertheless withhold concerns when speaking up is perceived as personally or professionally risky. Fear of retaliation, stigma, breaches of confidentiality, negative career consequences, loss of credibility, and distrust of organizational procedures can therefore prevent psychosocial problems from entering formal reporting systems (Bogaers et al., 2021; Dewa et al., 2021; McGrath et al., 2023). This issue is increasingly important within organizational scholarship because silence is not simply the absence of reporting; it may reflect workplace power relations, psychological safety, organizational culture, and employees' assessments of whether expressing concerns is likely to produce meaningful change. Recent evidence reinforces this connection. A systematic review and meta-analysis involving 84 studies and 34,975 participants found that greater employee silence and lower voice were significantly associated with higher burnout,

suggesting that organizational environments in which concerns remain unspoken may themselves be consequential for employee well-being (Lainidi et al., 2026).

This perspective has important implications for workplace psychosocial risk management. Conventional systems often depend on formal complaints, periodic employee surveys, counselling referrals, or human resource procedures that become activated only after an employee decides to disclose a concern. Such mechanisms remain necessary, but they may fail to capture emerging risks when employees are unwilling to identify themselves or when psychosocial difficulties have not yet developed into formally reportable incidents. Moreover, psychosocial screening, employee reporting, risk assessment, intervention planning, and organizational monitoring frequently operate through separate organizational processes (Ryan et al., 2023; Deady et al., 2024). The resulting problem is therefore twofold: psychosocial concerns may remain hidden because employees do not speak, while information that is available may remain fragmented because organizations lack mechanisms for integrating it. This creates a persistent gap between the existence of psychosocial information and the organizational capacity to transform that information into timely preventive action.

Artificial intelligence (AI) creates new possibilities within this gap because it can support the analysis of multiple forms of psychosocial information, including structured screening data and unstructured employee narratives. Natural Language Processing, Large Language Models, and predictive analytical approaches may assist organizations in identifying recurring themes, patterns, and risk signals that are difficult to detect through manual review alone (Dubosson et al., 2022; Guo et al., 2024). However, the increasing use of AI in workplaces also introduces an important contemporary social-science debate concerning surveillance, employee autonomy, privacy, transparency, and the power associated with data-driven decision-making. Recent research cautions against assuming that workplace AI is inherently beneficial. AI may enhance decision support and organizational responsiveness, but it may also increase perceptions of monitoring, reduce autonomy, generate privacy concerns, and create new psychosocial risks when implemented without adequate governance (Ferrarini & Leonelli, 2026; Karimova, 2026). The challenge is therefore not merely to introduce AI into workplace well-being systems, but to design AI-supported mechanisms that preserve confidentiality, augment rather than replace professional judgement, and translate psychosocial information into proportionate human and organizational responses.

Against this background, an important gap remains at the intersection of psychosocial risk management, employee voice, and AI-enabled organizational decision support. Existing scholarship has generated substantial knowledge about individual psychosocial hazards such as stress, burnout, workplace bullying, conflict, and psychological well-being. Other studies have examined disclosure barriers, organizational interventions, psychosocial safety, or applications of AI to workplace and mental-health-related data. However, these streams of research remain insufficiently connected. In particular, there is limited conceptual integration of (a) multidimensional psychosocial screening, (b) protected or anonymous employee reporting, (c) AI-assisted interpretation of structured and narrative information, (d) risk prioritisation, and (e) intervention and organizational monitoring within a continuous early-intervention process. The literature reviewed in this study therefore suggests that the critical gap is not simply the absence of another reporting application or another psychosocial

assessment instrument, but the absence of an integrated socio-technical framework capable of connecting employee voice, psychosocial assessment, analytical interpretation, and organizational action.

To address this gap, this paper proposes SafeVoice-AI-Release, an Artificial Intelligence-Based Early Intervention Ecosystem for Workplace Psychosocial Risk Management. Guided by the WHO Healthy Workplace Framework and contemporary approaches to psychosocial risk management, the ecosystem integrates the SafeVoice Psychosocial Screening Inventory (SPSI), anonymous reporting, AI-assisted analytics, predictive risk assessment, intervention recommendation, and organizational monitoring within a unified framework. Its conceptual contribution lies in connecting three areas that are frequently examined separately: psychosocial risk, employee voice and silence, and AI-supported organizational decision-making. Rather than positioning technology as a substitute for professional or organizational responsibility, SafeVoice-AI-Release conceptualises AI as a decision-support mechanism within a broader human-centred process of prevention, assessment, intervention, and monitoring. Accordingly, this study aims to develop and conceptually validate an integrated AI-supported ecosystem for the early identification and management of workplace psychosocial risks, with the broader objective of strengthening preventive organizational responses to employee well-being.

Literature Review

Work Stress

Work stress is widely recognised as a key psychosocial risk factor because it often precedes a range of adverse individual and organizational outcomes. Excessive workload, time pressure, role ambiguity, and work–life imbalance have been associated with emotional exhaustion, lower job satisfaction, absenteeism, and reduced organizational commitment (Antón, 2009; Wang et al., 2020; Jim et al., 2024; Julianto et al., 2024). Beyond these immediate effects, prolonged exposure to work stress may increase vulnerability to burnout and psychological distress (Backhaus et al., 2024; Loretto et al., 2010; Ryan & Cooney, 2024). This makes work stress an important indicator for the early identification of psychosocial risks and highlights the need for mechanisms capable of detecting emerging concerns before they affect employee well-being and organizational functioning (Oz, 2019; Ryan et al., 2023).

Burnout

Burnout reflects the cumulative impact of prolonged occupational strain and is frequently associated with reduced engagement, declining performance, absenteeism, and turnover intentions (Antón, 2009; Wang et al., 2020). Rather than representing an isolated individual problem, burnout often signals broader workplace issues, including excessive demands, inadequate support, and persistent organizational pressures (Backhaus et al., 2024; Ryan et al., 2023). Its significance lies in its role as an early warning indicator of deteriorating workplace conditions, making it an important dimension for psychosocial risk assessment and intervention planning (Ryan & Cooney, 2024).

Workplace Bullying

Workplace bullying constitutes a significant psychosocial risk because of its persistent effects on employee well-being and workplace relationships. Experiences of intimidation, exclusion, humiliation, and psychological mistreatment have been linked to anxiety, depression,

reduced self-esteem, job dissatisfaction, and lower organizational commitment (Boudrias et al., 2021; Bunce et al., 2024; Chatziioannidis et al., 2018). However, bullying frequently remains underreported due to concerns regarding stigma, retaliation, and confidentiality (Chatziioannidis et al., 2018). This reporting gap limits organizational awareness of emerging psychosocial risks and reinforces the importance of confidential reporting mechanisms within workplace well-being systems.

Organizational Conflict

Although conflict is an inevitable aspect of organizational life, unresolved conflict can become a significant source of psychosocial strain. Persistent disagreements, communication breakdowns, and interpersonal tensions may undermine collaboration, workplace morale, and employee well-being (Antón, 2009; Julianto et al., 2024). Research suggests that unresolved conflict not only affects interpersonal relationships but may also reduce team effectiveness and organizational performance (Antón, 2009; Dong et al., 2024). Consequently, the timely identification of conflict is important for preventing the escalation of workplace tensions into more serious psychosocial problems (Ryan et al., 2023).

Psychological Well-Being

Psychological well-being represents a protective dimension of workplace functioning and provides an important counterbalance to psychosocial risk indicators. Employees who demonstrate greater resilience, emotional stability, and workplace satisfaction are generally better equipped to manage occupational demands and adapt to workplace challenges (Dong et al., 2024). Contemporary workplace well-being research increasingly recognises that organizational health depends not only on reducing psychosocial hazards but also on strengthening positive psychological functioning (Ryan & Cooney, 2024; Dong et al., 2024). Psychological well-being is therefore an important indicator of both employee resilience and organizational health (Ryan et al., 2023; Dong et al., 2024).

Limitations of Existing Workplace Reporting Systems

Despite increasing awareness of workplace psychosocial risks, many organizations continue to rely on reporting systems that are primarily complaint-driven and reactive in nature (Ryan et al., 2023; Ryan & Cooney, 2024). Employees may be reluctant to disclose psychosocial concerns because of fears related to confidentiality, stigma, retaliation, or negative career consequences (Chatziioannidis et al., 2018; Ryan & Cooney, 2024). As a result, organizations often become aware of workplace problems only after substantial individual or organizational harm has occurred.

A further limitation is that workplace reporting systems are frequently designed to document incidents rather than identify psychosocial risks at an earlier stage. Psychosocial screening, risk assessment, intervention planning, and organizational monitoring often operate as separate processes, resulting in fragmented information and limited coordination (Ryan et al., 2023). This fragmentation reduces organizations' capacity to recognise emerging risks, implement preventive measures, and respond proactively to workplace well-being concerns (Oz, 2019; Gimpel et al., 2024; Ryan & Cooney, 2024). These limitations highlight the need for integrated approaches that support continuous monitoring, early detection, and evidence-informed intervention (Oz, 2019; Ryan et al., 2023).

Synthesis of Workplace Psychosocial Risk Dimensions

The literature indicates that workplace psychosocial risks cannot be adequately explained through a single construct. Instead, psychosocial conditions emerge through the interaction of work demands, interpersonal relationships, organizational processes, and individual psychological functioning (Deady et al., 2024; Leka & Kortum, 2008). Although these dimensions are often examined separately, their effects are closely interconnected (Galanakis & Tsitouri, 2022; Zeshan et al., 2024).

Excessive work demands may contribute to burnout, while workplace bullying and unresolved organizational conflict can intensify psychological strain and undermine employee well-being (Backhaus et al., 2024; Hayat & Afshari, 2021). At the same time, psychological well-being may influence how employees respond to workplace pressures and organisational challenges (Amoadu et al., 2024; Zeshan et al., 2024). These dimensions therefore operate as an interconnected system rather than as independent factors, shaping both employee experiences and organizational outcomes (Schaufeli, 2017; Boudrias et al., 2021).

Based on this synthesis, five domains were identified as central to workplace psychosocial risk assessment: Work Stress, Burnout, Workplace Bullying, Organizational Conflict, and Psychological Well-Being. Together, these dimensions capture both psychosocial risk factors and protective aspects of employee functioning, providing a more comprehensive understanding of workplace well-being (Amoadu et al., 2024; Schaufeli, 2017; Galanakis & Tsitouri, 2022). These domains subsequently informed the development of the SafeVoice Psychosocial Screening Inventory (SPSI) and the conceptual design of the SafeVoice-AI-Release ecosystem.

Conceptual Foundation of SafeVoice-AI-Release

SafeVoice-AI-Release was developed in response to two persistent challenges identified in workplace psychosocial risk research: the underreporting of psychosocial concerns and the limited capacity of conventional reporting systems to support prevention and early intervention (Deady et al., 2024). In many organizations, psychosocial issues are recognised only after substantial individual or organizational consequences have occurred (Schaufeli, 2017). Guided by the WHO Healthy Workplace Framework, the proposed framework views workplace well-being as the outcome of interactions between psychosocial conditions, organizational practices, and employee functioning (Leka & Kortum, 2008). This perspective highlights the need to move beyond incident reporting towards the continuous identification and management of psychosocial risks (Bergh et al., 2017; Deady et al., 2024; Schaufeli, 2017).

The framework is grounded in the assumption that psychosocial risks emerge progressively through workplace experiences and organizational processes rather than as isolated incidents (Backhaus et al., 2024). Changes in workload, interpersonal relationships, organizational climate, and psychological functioning often precede more visible psychosocial problems (Backhaus et al., 2024; Amoadu et al., 2024). Consequently, psychosocial risks are unlikely to be effectively managed through approaches that depend solely on formal complaints after problems have become apparent (Deady et al., 2024). Understanding psychosocial risk as a dynamic process therefore shifts attention towards the identification of early indicators that may signal future harm and require preventive intervention (Schaufeli, 2017).

Central to the framework is the SafeVoice Psychosocial Screening Inventory (SPSI), which incorporates five domains: Work Stress, Burnout, Workplace Bullying, Organizational Conflict, and Psychological Well-Being. These domains were selected because workplace psychosocial conditions are multidimensional and cannot be adequately explained through a single construct (Deady et al., 2024; Leka & Kortum, 2008). Together, they capture both sources of psychosocial risk and indicators of employee resilience, providing a broader understanding of workplace well-being and reflecting the interconnected nature of psychosocial experiences within organizational settings (Amoadu et al., 2024; Galanakis & Tsitouri, 2022; Zeshan et al., 2024).

A central contribution of SafeVoice-AI-Release lies in its integration of psychosocial screening, anonymous reporting, risk assessment, intervention planning, and organizational monitoring within a single framework. This integration is important because psychosocial concerns frequently emerge through the interaction of multiple factors that remain fragmented across separate organizational processes (Deady et al., 2024; Sivris & Leka, 2015). Reliance on a single source of information may obscure emerging risks, whereas the integration of assessment and reporting data provides a more comprehensive understanding of workplace psychosocial conditions (Bergh et al., 2017). Integrating these functions within a continuous process strengthens organizational capacity to identify concerns earlier and coordinate appropriate responses (Schaufeli, 2017).

Within this integrated framework, AI-assisted analytics functions as a mechanism for identifying patterns and relationships across multiple sources of psychosocial information. Its contribution lies not merely in processing large volumes of data but in supporting the recognition of emerging psychosocial risks that may not be readily observable through conventional reporting mechanisms. By transforming fragmented psychosocial information into actionable insights, AI-assisted analytics supports earlier organizational responses and more informed decision-making.

By linking assessment, reporting, analysis, and intervention within a continuous cycle, the framework conceptualises workplace psychosocial risk management as an ongoing organizational responsibility rather than a reactive response to isolated incidents (Schaufeli, 2017; Sivris & Leka, 2015; Deady et al., 2024). In doing so, it advances a preventive conceptualisation of workplace psychosocial risk management in which assessment, reporting, monitoring, and intervention are understood as interconnected components of organizational well-being (Leka & Kortum, 2008; Bergh et al., 2017). Such an approach provides a foundation for earlier, more coordinated, and evidence-informed responses to workplace psychosocial concerns (Amoadu et al., 2024; Deady et al., 2024).

System Development Methodology

The development of SafeVoice-AI-Release adopted a Design and Development Research (DDR) approach to guide the systematic design of the proposed workplace psychosocial risk management ecosystem (Aris, 2024; Padzil et al., 2021; Richey & Klein, 2013; Easterday et al., 2018). DDR was selected because it provides a structured pathway for translating theoretical and empirical knowledge into practical innovation (Aris, 2024; Richey & Klein, 2013). Within this study, the approach enabled the integration of psychosocial risk theory, organizational

reporting practices, and AI-assisted decision support into a coherent framework aligned with workplace well-being objectives (Leicht-Deobald et al., 2019; Kawakami et al., 2023).

The development process began with a needs assessment informed by literature on workplace psychosocial risks, employee well-being, and organizational reporting systems (Padzil et al., 2021; Aris, 2024; Easterday et al., 2018). Particular attention was given to work stress, burnout, workplace bullying, organizational conflict, psychological well-being, and barriers to reporting workplace concerns (Hayat & Afshari, 2021; Amoadu et al., 2024). This stage established the conceptual foundation of the framework by identifying a recurring limitation in existing workplace practices: psychosocial concerns are frequently addressed after harm has occurred rather than through mechanisms that support prevention and ongoing monitoring.

Subsequently, a system blueprint was developed to define system objectives, stakeholder roles, functional requirements, and data flows (Aris, 2024; Padzil et al., 2021; Easterday et al., 2018). Establishing these relationships was important because workplace psychosocial risks rarely exist as isolated events (Amoadu et al., 2024). Effective management requires the integration of assessment, reporting, analysis, and intervention processes that are often dispersed across different organizational functions (Kawakami et al., 2023).

User journey mapping was then undertaken to examine interactions among employees, counsellors, human resource personnel, and system administrators. Because psychosocial risk management involves multiple stakeholders with distinct responsibilities, mapping these interactions helped ensure that reporting pathways, screening procedures, and intervention mechanisms reflected practical organizational processes and decision-making needs (Leicht-Deobald et al., 2019; Kawakami et al., 2023).

The development of the system and AI architectures focused on creating a framework capable of transforming psychosocial information into actionable organizational insights. While the system architecture supported data collection, storage, processing, and reporting, the AI architecture was designed to facilitate complaint classification, psychosocial risk identification, predictive risk assessment, and intervention recommendation. This integration was particularly important because psychosocial concerns often emerge through patterns and relationships that may remain obscured when information is fragmented across separate reporting channels (Kawakami et al., 2023). Within this framework, AI-assisted analytics functions as a mechanism for recognising emerging psychosocial patterns and supporting more informed organizational responses.

The final stage involved prototype development based on the conceptual framework and system architecture (Aris, 2024; Padzil et al., 2021; Easterday et al., 2018). Rather than serving solely as a technical demonstration, the prototype was designed to illustrate how psychosocial screening, anonymous reporting, analytics, intervention recommendation, and organizational monitoring could operate within a unified ecosystem. This stage provided a practical representation of how multiple sources of psychosocial information could be integrated to support organizational decision-making and workplace well-being management (Kawakami et al., 2023).

Overall, the DDR approach provided a systematic pathway for translating workplace psychosocial risk literature into a functional ecosystem design (Richey & Klein, 2013; Aris, 2024; Padzil et al., 2021). Rather than treating technology as an end in itself, the development process ensured that technological innovation remained grounded in psychosocial risk management principles, thereby aligning system design with the broader objective of strengthening workplace well-being and organizational responsiveness (Leicht-Deobald et al., 2019; Kawakami et al., 2023).

Expert Validation of the SafeVoice-AI-Release Framework

To enhance the content validity, practical relevance, and feasibility of the proposed SafeVoice-AI-Release ecosystem, an expert validation process was conducted involving ten subject matter experts from multidisciplinary fields relevant to workplace psychosocial risk management. The expert panel consisted of three industrial and organizational psychologists, three human resource management specialists, two artificial intelligence and data analytics experts, and two organizational counselling practitioners. The selection of experts was based on their professional experience, academic qualifications, and demonstrated expertise in workplace well-being, psychosocial risk assessment, organizational intervention, or artificial intelligence applications (Hasson et al., 2000; Lynn, 1986).

The validation process aimed to evaluate the conceptual adequacy and applicability of the SafeVoice-AI-Release framework prior to full prototype implementation and empirical testing. Experts were provided with the conceptual framework, system architecture, component descriptions, and operational workflow of the proposed ecosystem. They were subsequently asked to assess the framework using a structured evaluation form comprising five dimensions: relevance, clarity, practicality, feasibility, and innovation (Polit & Beck, 2006). Relevance referred to the extent to which the framework addressed contemporary workplace psychosocial risk management challenges. Clarity examined whether the framework components, processes, and relationships were clearly articulated and understandable. Practicality assessed the potential usefulness of the framework in real organizational settings. Feasibility evaluated the likelihood of implementation within existing organizational structures and technological environments. Innovation focused on the originality and added value of integrating psychosocial screening, anonymous reporting, artificial intelligence analytics, predictive risk assessment, and intervention planning within a unified ecosystem (Nambisan et al., 2017).

Content validity was assessed using the Content Validity Index (CVI). Experts rated each evaluation item within the five dimensions on a four-point scale ranging from 1 (not relevant) to 4 (highly relevant). Item-level Content Validity Index (I-CVI) values were calculated by dividing the number of experts assigning ratings of 3 or 4 by the total number of experts. Dimension-level CVI values were subsequently calculated by averaging the I-CVI values of the items within each dimension, while the overall Scale-Level Content Validity Index using the averaging method (S-CVI/Ave) was calculated across all items. In addition, percentage agreement was calculated to provide an indication of expert consensus regarding the suitability of the proposed framework (Lynn, 1986; Polit & Beck, 2006).

The validation findings demonstrated strong support for the SafeVoice-AI-Release framework. The overall S-CVI/Ave was 0.91, exceeding the commonly recommended

benchmark of 0.90 for excellent scale-level content validity (Polit & Beck, 2006). Percentage agreement across the five evaluation dimensions ranged from 85% to 95%, reflecting substantial consensus among experts regarding the framework's conceptual relevance and practical potential (Polit & Beck, 2006). Experts particularly highlighted the value of integrating anonymous reporting mechanisms with psychosocial screening and AI-assisted analytics, noting that such integration could strengthen organizational capacity for early psychosocial risk detection and intervention. This emphasis is broadly consistent with contemporary guidance advocating preventive and organizational approaches to workplace mental health (World Health Organization & International Labour Organization, 2022).

Several recommendations were provided by the expert panel to further strengthen the framework. These included the development of detailed psychosocial assessment indicators within the SafeVoice Psychosocial Screening Inventory (SPSI), refinement of the proposed predictive risk scoring mechanism, incorporation of ethical governance principles for artificial intelligence implementation, and enhancement of data privacy and confidentiality safeguards. These recommendations were incorporated into the revised conceptual framework and will inform future prototype development and empirical validation phases. The emphasis on ethical governance and data safeguards is consistent with UNESCO's (2021) Recommendation on the Ethics of Artificial Intelligence.

Overall, the expert validation process provided preliminary evidence supporting the conceptual soundness, relevance, and practical applicability of the SafeVoice-AI-Release ecosystem. The findings suggest that the framework has potential as a workplace psychosocial risk management innovation and warrants further development through prototype implementation, usability testing, and organizational pilot studies (World Health Organization & International Labour Organization, 2022).

Table 1

Expert Panel Composition

Expert Category	Number of Experts
Industrial and Organizational Psychologists	3
Human Resource Specialists	3
Artificial Intelligence Experts	2
Organizational Counsellors	2
Total	10

Table 2

Expert Evaluation Criteria

Evaluation Dimension	Description
Relevance	Suitability of the framework for workplace psychosocial risk management
Clarity	Clarity of framework structure and component relationships
Practicality	Applicability in organizational settings
Feasibility	Potential for implementation and operationalization

Evaluation Dimension	Description
Innovation	Originality and contribution of the framework

Table 3

Expert Validation Results

Dimension	CVI	Percentage Agreement (%)
Relevance	0.95	95
Clarity	0.90	90
Practicality	0.90	90
Feasibility	0.85	85
Innovation	0.95	95
Overall (S-CVI/Ave)	0.91	91

The conceptual framework of SafeVoice-AI-Release

The conceptual framework of SafeVoice-AI-Release was developed through the integration of occupational health theory, workplace psychosocial risk literature, and contemporary artificial intelligence technologies (Aust et al., 2023; Deady et al., 2024; Dubosson et al., 2022; Guo et al., 2024). It is informed by the World Health Organization (WHO) Healthy Workplace Framework (2010), which recognises psychosocial working conditions as a critical determinant of employee well-being and organizational effectiveness (Deady et al., 2024; Miguel et al., 2023). Consistent with this perspective, the framework views psychosocial risk management as a preventive organizational responsibility that requires the identification and management of emerging concerns before they develop into more serious individual or organizational consequences (Loh et al., 2024; Deady et al., 2024; Berglund et al., 2024). This orientation is particularly relevant given the tendency of workplace psychosocial issues to remain unrecognised until they have already affected employee functioning, workplace relationships, or organizational performance (Dubosson et al., 2022).

The framework is organised around five psychosocial constructs: Work Stress, Burnout, Workplace Bullying, Organizational Conflict, and Psychological Well-Being (Aust et al., 2023; Amoadu et al., 2024; Galanis et al., 2024). These constructs were selected because workplace psychosocial conditions are shaped by the interaction of occupational demands, interpersonal relationships, organizational processes, and individual psychological functioning (Dubosson et al., 2022; Deady et al., 2024). Work Stress and Burnout reflect the psychological consequences of workplace demands (Aust et al., 2023), while Workplace Bullying and Organizational Conflict represent interpersonal and organizational sources of strain (Amoadu et al., 2024; Galanis et al., 2024; Bernotaitė & Malinauskiene, 2017). In contrast, Psychological Well-Being captures a protective dimension associated with employees' capacity to adapt to workplace challenges (Amoadu et al., 2024). Considered collectively, these constructs provide a broader understanding of workplace psychosocial conditions than any single indicator alone and form the basis of the SafeVoice Psychosocial Screening Inventory (SPSI) (Deady et al., 2024; Miguel et al., 2023). Their inclusion reflects the premise that psychosocial risk arises through the interaction of multiple workplace conditions rather than from isolated experiences or events (Dubosson et al., 2022).

A central assumption of the framework is that psychosocial risks may remain obscured when relevant information is dispersed across separate organizational processes (Daniels et al., 2021; Deady et al., 2024; Dubosson et al., 2022). Conventional reporting systems frequently rely on individual complaints or isolated incidents, which may provide only a partial representation of workplace psychosocial conditions (Dubosson et al., 2022; Stratton et al., 2022). To address this limitation, the framework incorporates a technological layer comprising Artificial Intelligence (AI) Analytics, Natural Language Processing (NLP), Large Language Models (LLMs), and Predictive Risk Scoring mechanisms. The purpose of this layer is not merely analytical but integrative. By identifying patterns across psychosocial screening outcomes and anonymous employee reports, these technologies facilitate the transformation of fragmented information into structured insights that can support organizational assessment, decision-making, and risk management (Dubosson et al., 2022). This capability is particularly important because psychosocial concerns often emerge gradually and may not be readily identifiable through a single source of information (Deady et al., 2024).

The framework also incorporates an intervention layer that connects psychosocial assessment with organizational action (Aust et al., 2023; Daniels et al., 2021; Loh et al., 2024). This component recognises that the value of psychosocial assessment lies not solely in identifying risks but in its capacity to inform timely and coordinated responses (Berglund et al., 2024; Miguel et al., 2023). Through the integration of early intervention strategies, counselling support, human resource responses, and organizational monitoring mechanisms, the framework seeks to ensure that identified concerns are linked to appropriate support pathways (Deady et al., 2024; Stratton et al., 2022; Berglund et al., 2024). This connection between assessment and intervention is central to the framework because the effectiveness of psychosocial risk management depends on an organization's ability to translate information into meaningful action that supports employee well-being and organizational resilience (Aust et al., 2023; Daniels et al., 2021).

Taken together, SafeVoice-AI-Release functions as an Artificial Intelligence Early Intervention Ecosystem in which psychosocial assessment, anonymous reporting, AI-assisted analysis, risk classification, intervention planning, and organizational monitoring operate as interconnected processes (Dubosson et al., 2022). Unlike conventional workplace reporting systems that primarily document incidents after they occur, the framework emphasises prevention, continuous monitoring, and organizational responsiveness (Loh et al., 2024; Berglund et al., 2024; Deady et al., 2024). In doing so, it advances a conceptualisation of workplace psychosocial risk management in which assessment, reporting, analysis, and intervention are understood as mutually reinforcing components of organizational well-being (Daniels et al., 2021; Aust et al., 2023; Miguel et al., 2023). By integrating psychosocial theory, workplace reporting practices, and AI-enabled analytics within a unified framework, SafeVoice-AI-Release provides a conceptual foundation for strengthening organizational capacity to understand, manage, and respond to workplace psychosocial risks in a more systematic and preventive manner (Guo et al., 2024).

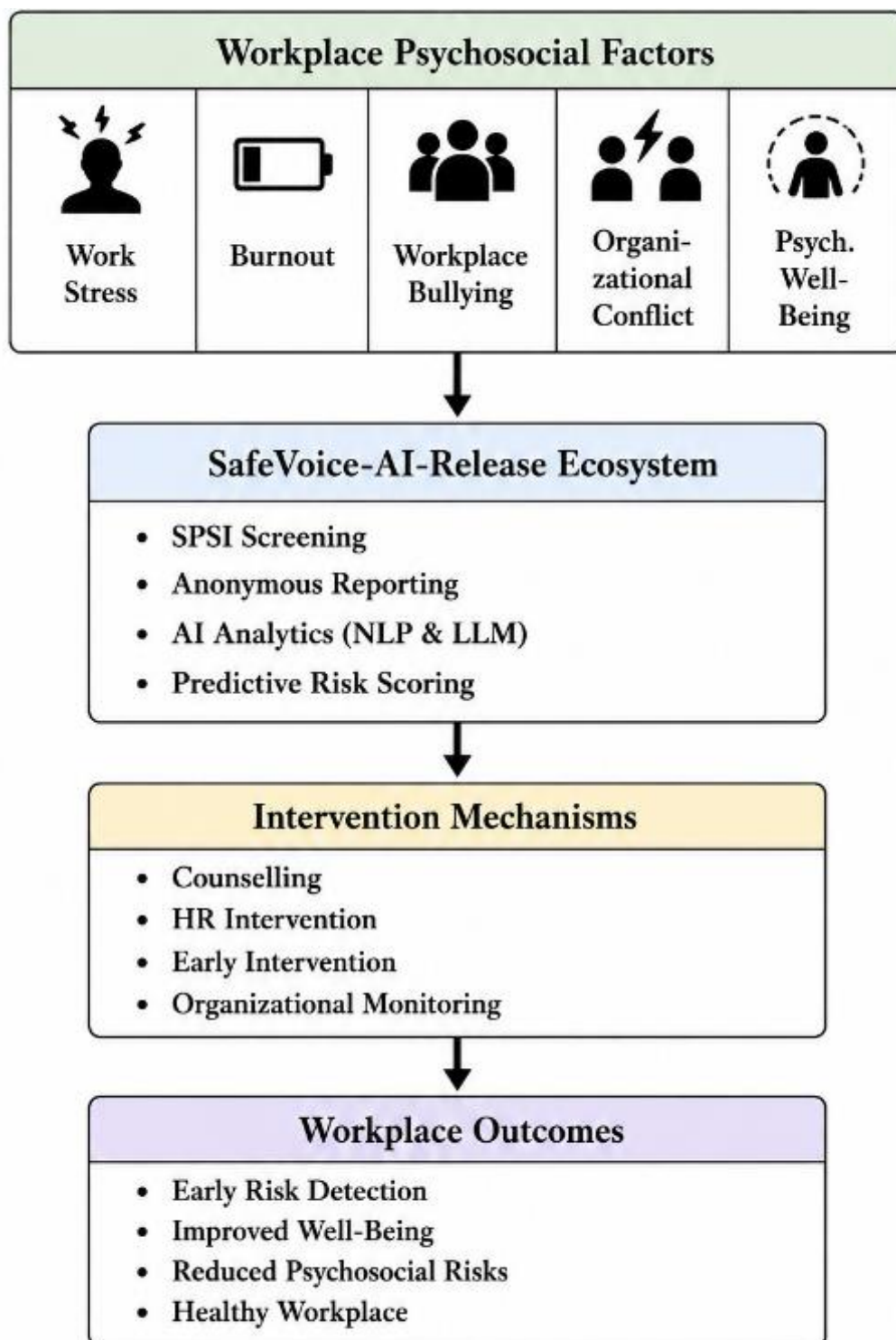


Figure 1 illustrates the conceptual framework underpinning SafeVoice-AI-Release, highlighting the relationships between the theoretical foundation, psychosocial constructs, technological components, and intervention mechanisms that collectively support workplace psychosocial risk management.

SafeVoice Psychosocial Screening Inventory (SPSI)

The SafeVoice Psychosocial Screening Inventory (SPSI) serves as the primary assessment component within the SafeVoice-AI-Release ecosystem (Amoadu et al., 2024; Deady et al., 2024; Dubosson et al., 2022; Kawakami et al., 2023). Developed from the synthesis of workplace psychosocial risk literature, the instrument was designed to provide a structured assessment of psychosocial conditions that may not be readily observable through organizational reporting mechanisms alone (Dubosson et al., 2022; Kawakami et al., 2023). This function is particularly important because psychosocial concerns often emerge gradually and may remain undetected until they affect employee well-being and organizational functioning (Dubosson et al., 2022).

At its current conceptual stage, the provisional SPSI comprises twenty items distributed across five domains: Work Stress, Burnout, Workplace Bullying, Organizational Conflict, and Psychological Well-Being. These domains were selected because they represent both psychosocial risk factors and protective dimensions of employee functioning (Amoadu et al., 2024; Hayat & Afshari, 2021; Ribeiro et al., 2021). Work Stress and Burnout reflect the effects of occupational demands, whereas Workplace Bullying and Organizational Conflict capture interpersonal and organizational sources of strain (Hayat & Afshari, 2021; Ribeiro et al., 2021). Psychological Well-Being provides an indicator of employees' capacity to cope with workplace challenges (Amoadu et al., 2024; Hayat & Afshari, 2021; Kawakami et al., 2023). Considered collectively, these domains offer a multidimensional assessment of workplace psychosocial conditions and provide the foundation for subsequent analysis, risk assessment, and intervention planning (Dubosson et al., 2022). Rather than focusing on a single indicator, the SPSI recognises that workplace well-being is shaped by the interaction of multiple psychosocial factors operating simultaneously (Kawakami et al., 2023; Amoadu et al., 2024; Hayat & Afshari, 2021; Ribeiro et al., 2021). The proposed items require subsequent psychometric validation before the SPSI can be used as an empirically established workplace screening instrument.

Anonymous Reporting Module

Underreporting remains a persistent challenge in workplace psychosocial risk management, as employees may be reluctant to disclose concerns because of fears related to confidentiality, stigma, retaliation, or potential career consequences (Kyung et al., 2023; Klinefelter et al., 2021; Dewa et al., 2021; Dewa, 2014). As a result, organizational responses are often initiated only after concerns have become more visible or difficult to address (Kyung et al., 2023; Zadow et al., 2017).

The Anonymous Reporting Module was developed to reduce these barriers by providing employees with a confidential mechanism for reporting workplace concerns (Dahler-Larsen et al., 2020; Dewa, 2014). The module accommodates reports relating to workplace bullying, harassment, abuse of authority, organizational conflict, burnout-related concerns, and other psychosocial difficulties that may affect workplace functioning (Amoadu et al., 2024; Kiaos, 2025). Beyond its reporting function, this component contributes contextual information that may not be captured through structured psychosocial assessments. Anonymous reports therefore complement SPSI data by providing insights into employees' lived workplace experiences and enabling a more comprehensive understanding of organizational psychosocial conditions (Dahler-Larsen et al., 2020; Kiaos, 2025).

AI Analytics Engine

The AI Analytics Engine functions as the analytical core of the SafeVoice-AI-Release ecosystem. Its inclusion reflects the assumption that psychosocial concerns are often embedded within narrative accounts, workplace experiences, and patterns of interaction that may be difficult to identify through manual review alone (Uronen et al., 2022; Ricketts et al., 2023; Young et al., 2019; Byon et al., 2023; Becker et al., 2025).

Drawing on Natural Language Processing (NLP) and Large Language Model (LLM) capabilities, the engine analyses employee narratives to identify recurring psychosocial themes, classify reported concerns, and detect patterns associated with emerging risks (Uronen et al., 2022; Becker et al., 2025). This process enables qualitative workplace experiences to be transformed into structured information that can support organizational decision-making (Khairuddin et al., 2022; Ricketts et al., 2023; Zhang et al., 2019).

Importantly, the value of the AI Analytics Engine extends beyond complaint categorisation (Ricketts et al., 2023; Becker et al., 2025; Khairuddin et al., 2022). By identifying relationships across reports and psychosocial indicators, it facilitates a broader understanding of workplace conditions and supports the recognition of issues that may otherwise remain fragmented across departments, teams, or reporting channels. In this way, the component strengthens organizational capacity to interpret psychosocial information and respond more strategically to emerging concerns.

SafeVoice Predictive Risk Scoring Model

The SafeVoice Predictive Risk Scoring Model was developed to integrate psychosocial assessment outcomes with information derived from employee reports within a unified risk assessment framework. The model is based on the premise that psychosocial risks are more accurately understood when multiple sources of information are considered together rather than in isolation (WHO, 2010).

A composite risk score may be generated by combining SPSI results with indicators derived from anonymous employee reports. This integration enables the system to capture both self-reported psychosocial conditions and workplace experiences, thereby providing a broader representation of psychosocial risk within the organization (Leka et al., 2011).

For conceptual purposes, the proposed model organises psychosocial risk into four provisional categories: Low Risk, Moderate Risk, High Risk, and Critical Risk. These categories are intended to facilitate the prioritisation of intervention needs and resource allocation in accordance with broader psychosocial risk-management principles (European Agency for Safety and Health at Work [EU-OSHA], 2022; World Health Organization, 2010). The categories and their corresponding thresholds remain provisional and require empirical validation.

At the current stage of development, the model remains conceptual. Weighting parameters, classification thresholds, and predictive algorithms have yet to be empirically established. Future research is therefore needed to evaluate the reliability, predictive performance, and practical applicability of the model across different organizational contexts.

Intervention Recommendation Module

The Intervention Recommendation Module was developed to bridge the gap between psychosocial risk identification and organizational action. Within workplace settings, the value of psychosocial assessment lies not solely in detecting risk but in informing responses that are proportionate to the severity of identified concerns. Without a mechanism that translates assessment outcomes into action, risk identification alone may have limited practical value (World Health Organization, 2010).

To address this issue, the module links intervention strategies to risk classifications generated through the Predictive Risk Scoring Model. Recommendations range from self-help resources, psychoeducational programmes, and resilience-building activities for lower-risk cases to counselling support, conflict mediation, human resource involvement, and crisis management procedures for higher-risk cases. This tiered approach reflects the assumption that psychosocial concerns vary in severity and therefore require differentiated forms of organizational support (Leka et al., 2011; International Labour Organization, 2016).

More importantly, the module reinforces the preventive orientation of SafeVoice-AI-Release by embedding intervention planning within the broader risk management process. Rather than treating psychosocial concerns as isolated incidents requiring reactive responses, the framework positions intervention as an integral component of continuous workplace well-being management (World Health Organization, 2010; Leka et al., 2015).

Organizational Dashboard

The Organizational Dashboard serves as the primary decision-support interface within the SafeVoice-AI-Release ecosystem. Information derived from psychosocial screening, anonymous reporting, AI-assisted analytics, and risk assessment processes is consolidated into a single monitoring environment to support organizational oversight of psychosocial conditions (Leka & Jain, 2010; World Health Organization, 2010).

Its significance extends beyond data visualisation. Workplace psychosocial risks often emerge through recurring patterns that may be difficult to identify when information remains dispersed across different organizational systems. By integrating multiple sources of psychosocial information, the dashboard provides a broader view of risk distributions, complaint trends, intervention outcomes, and organizational well-being patterns.

This capability supports both operational and strategic decision-making. In addition to facilitating the monitoring of emerging concerns, the dashboard provides evidence that can inform organizational policies, employee support initiatives, and resource allocation decisions (Leka et al., 2011). Consequently, monitoring becomes part of an ongoing organizational process rather than a response triggered only after psychosocial problems become visible.

System Integration and Innovation

The principal contribution of SafeVoice-AI-Release lies in its integration of psychosocial assessment, employee reporting, AI-assisted analytics, risk assessment, intervention planning, and organizational monitoring within a single framework. Existing workplace systems frequently address these functions separately, resulting in fragmented information

and limited capacity to develop a comprehensive understanding of workplace psychosocial conditions (Leka et al., 2011; World Health Organization, 2010).

This fragmentation is particularly problematic because psychosocial risks rarely emerge through a single event or indicator. Instead, they develop through the interaction of individual experiences, workplace relationships, organizational processes, and broader psychosocial conditions (Leka & Jain, 2010; Cox et al., 2000). Addressing these risks therefore requires mechanisms capable of connecting information across multiple sources rather than relying on isolated complaints or assessments.

SafeVoice-AI-Release was developed in response to this challenge by conceptualising psychosocial risk management as a continuous cycle linking assessment, reporting, analysis, intervention, and monitoring. Through this integration, the framework shifts attention from incident-based management towards a preventive and evidence-informed approach that supports organizational responsiveness to emerging psychosocial concerns (European Agency for Safety and Health at Work [EU-OSHA], 2022; Leka et al., 2011).

Accordingly, the innovation of SafeVoice-AI-Release lies not in any single technological component but in the integration of previously disconnected processes within a unified workplace psychosocial risk management ecosystem.

Expected Outcomes and Impact

SafeVoice-AI-Release was developed in response to increasing recognition that workplace well-being is shaped by psychosocial conditions that may remain unaddressed until significant consequences have occurred (World Health Organization, 2010; Leka & Jain, 2010). By integrating psychosocial screening, anonymous reporting, AI-assisted analytics, and intervention planning within a single ecosystem, the framework seeks to strengthen organizational capacity for the earlier identification and management of psychosocial concerns.

For employees, the framework provides opportunities to disclose workplace concerns through confidential reporting mechanisms while supporting the identification of psychosocial vulnerabilities through structured assessment. Reducing barriers to disclosure is particularly important because underreporting remains a persistent obstacle to effective psychosocial risk management. Earlier recognition of concerns creates greater opportunities for timely support and intervention.

For organizations, the integration of assessment data, employee reports, and analytical processes enables a more comprehensive understanding of workplace psychosocial conditions than would be possible through isolated reporting systems alone. This information can support the identification of recurring psychosocial issues, inform preventive strategies, and strengthen organizational responsiveness to emerging risks (Leka et al., 2011; European Agency for Safety and Health at Work [EU-OSHA], 2022).

The framework also contributes to professional practice by providing counsellors, human resource practitioners, and organizational leaders with structured information to support assessment, intervention, and follow-up activities. Access to integrated psychosocial

information may improve coordination across support systems and facilitate more evidence-informed decision-making.

More broadly, SafeVoice-AI-Release reflects a preventive conceptualisation of workplace psychosocial risk management in which assessment, reporting, analysis, intervention, and monitoring operate as interconnected processes. By shifting attention from reactive case management towards the continuous management of psychosocial conditions, the framework provides a foundation for healthier, safer, and more sustainable workplace environments (World Health Organization, 2010; Leka et al., 2011).

Discussion

Workplace psychosocial risks are increasingly recognised as important determinants of employee well-being and organizational performance (Leka et al., 2011). Nevertheless, organizational responses remain largely reactive, relying on formal complaints after concerns have become visible. Such approaches are limited because psychosocial risks often develop gradually and may remain unreported due to stigma, fear of retaliation, confidentiality concerns, or perceived career consequences (Dewa et al., 2021; Klinefelter et al., 2021; McGrath et al., 2023). SafeVoice-AI-Release addresses this limitation by conceptualising psychosocial risk management as an ongoing organizational process that integrates assessment, reporting, analysis, and intervention.

A key contribution of the framework is its emphasis on earlier recognition of psychosocial concerns. Conventional workplace well-being initiatives frequently depend on surveys or formal reporting mechanisms, both of which require employees to disclose their experiences. Although valuable, these approaches may fail to capture concerns that have not yet reached the threshold of formal complaints. The integration of psychosocial screening and anonymous reporting broadens the evidence available to organizations and strengthens their capacity to recognise emerging psychosocial risks.

The inclusion of Work Stress, Burnout, Workplace Bullying, Organizational Conflict, and Psychological Well-Being reflects the multidimensional nature of workplace psychosocial conditions. Workplace well-being is shaped by the interaction of occupational demands, interpersonal relationships, organizational processes, and individual psychological functioning (Leka et al., 2011; WHO, 2010). By incorporating both risk and protective dimensions, the SPSI provides a more comprehensive representation of workplace psychosocial conditions than approaches focused solely on psychosocial hazards.

The framework also responds to the fragmentation that characterises many organizational approaches to psychosocial risk management. Assessment, reporting, intervention, and monitoring are often undertaken as separate activities, limiting the ability to develop a coherent understanding of workplace conditions (Leka et al., 2011). By linking these processes within a continuous cycle, SafeVoice-AI-Release enables psychosocial information to inform organizational responses more systematically.

The integration of AI Analytics, Natural Language Processing (NLP), and Large Language Model (LLM) capabilities further strengthens this process. Employee narratives often contain psychosocial information that may not be readily captured through structured assessments

or conventional reporting mechanisms. AI-assisted analytics supports the identification of recurring themes and patterns across multiple sources of information, providing additional evidence to inform professional judgement and organizational decision-making (Shatte et al., 2019).

The framework also has practical implications for workplace well-being management. Integrating psychosocial assessment and reporting data may assist organizations in identifying risk patterns, prioritising interventions, and evaluating workplace initiatives, while improving coordination among counsellors, human resource practitioners, and organizational leaders (Leka et al., 2011; WHO, 2022).

Several limitations should be acknowledged. SafeVoice-AI-Release remains a conceptual framework and has not yet been implemented or evaluated within organizational settings. The SPSI therefore requires empirical validation, while the proposed Predictive Risk Scoring Model requires further testing to establish its predictive performance. Future research should examine system usability, user acceptance, and the effectiveness of AI-assisted recommendations across different workplace contexts.

Overall, SafeVoice-AI-Release advances a preventive conceptualisation of workplace psychosocial risk management by integrating psychosocial assessment, anonymous reporting, AI-assisted analytics, risk assessment, intervention planning, and organizational monitoring within a single ecosystem. Its contribution lies in bringing together functions that are often treated separately, thereby supporting a more coordinated approach to understanding and managing workplace psychosocial conditions.

Conclusion

This article proposed SafeVoice-AI-Release as a conceptual framework for workplace psychosocial risk management developed in response to the limitations of reactive reporting systems that often identify concerns only after significant consequences have occurred. By integrating psychosocial assessment, anonymous reporting, AI-assisted analytics, risk assessment, intervention planning, and organizational monitoring, the framework advances a more preventive approach to managing workplace psychosocial conditions.

A central contribution of the framework is its conceptualisation of psychosocial risk management as an ongoing organizational process rather than a response to isolated incidents. This perspective recognises that psychosocial concerns are shaped by the interaction of individual experiences, workplace relationships, and organizational conditions, requiring coordinated mechanisms that connect assessment, reporting, analysis, intervention, and monitoring.

The framework also contributes to discussions on workplace well-being by illustrating how psychosocial screening and employee reporting data may be integrated with AI-assisted analytics to support organizational decision-making. Its significance lies not in the adoption of technology alone but in the integration of previously fragmented processes within a unified psychosocial risk management ecosystem.

As a conceptual framework, SafeVoice-AI-Release requires further empirical development and evaluation. Future research should focus on validating the SafeVoice Psychosocial Screening Inventory (SPSI), examining the predictive performance of the proposed risk assessment model, and evaluating implementation across different organizational contexts. Such work is necessary to determine the practical applicability of the framework and its potential contribution to workplace psychosocial well-being.

References

- Abd Shukor, S., & Khalid, M. S. (2025). Psychosocial risk assessment and management: Critical appraisal and advancement of Malaysia's PRISMA 2024 guidelines. *International Journal of Education, Psychology and Counseling*, 10(61), 100–129. <https://doi.org/10.35631/IJEPC.1061008>
- Amodu, M., Ansah, E. W., & Sarfo, J. (2024). Preventing workplace mistreatment and improving workers' mental health: a scoping review of the impact of psychosocial safety climate. *BMC Psychology*, 12. <https://doi.org/10.1186/s40359-024-01675-z>
- Antón, C. (2009). The impact of role stress on workers' behaviour through job satisfaction and organizational commitment.. *International journal of psychology : Journal international de psychologie*, 44 3, 187-94 . <https://doi.org/10.1080/00207590701700511>
- Aris, N. M. (2024). Design and Development Research (DDR) Approach in Designing Design Thinking Chemistry Module to Empower Students' Innovation Competencies. *Journal of Advanced Research in Applied Sciences and Engineering Technology*. <https://doi.org/10.37934/araset.44.1.5568>
- Aust, B., Møller, J. L., Nordentoft, M., Frydendall, K. B., Bengtsen, E., Jensen, A., Garde, A., Kompier, M., Semmer, N., Rugulies, R., & Jaspers, S. (2023). How effective are organizational-level interventions in improving the psychosocial work environment, health, and retention of workers? A systematic overview of systematic reviews. *Scandinavian Journal of Work, Environment & Health*, 49, 315 - 329. <https://doi.org/10.5271/sjweh.4097>
- Backhaus, I., Lohmann-Haislah, A., Burr, H., Nielsen, K. M., Di Tecco, C., & Dragano, N. (2024). Organizational change: challenges for workplace psychosocial risks and employee mental health. *BMC Public Health*, 24. <https://doi.org/10.1186/s12889-024-19815-w>
- Becker, M., Hwang, S., Schriver, E., Douma, C., Duffy, C., Atkins, J., McShane, C., Lubken, J., Hanish, A., McGreevey, J., Regli, S. H., & Mowery, D. L. (2025). Automatically identifying event reports of workplace violence and communication failures using large language models. *AMIA Joint Summits on Translational Science Proceedings*.
- Bergh, L., Leka, S., & Zwetsloot, G. (2017). Tailoring Psychosocial Risk Assessment in the Oil and Gas Industry by Exploring Specific and Common Psychosocial Risks. *Safety and Health at Work*, 9, 63 - 70. <https://doi.org/10.1016/j.shaw.2017.05.001>
- Berglund, R. T., Kombeiz, O., & Dollard, M. F. (2024). Manager-driven intervention for improved psychosocial safety climate and psychosocial work environment. *Safety Science*. <https://doi.org/10.1016/j.ssci.2024.106552>
- Bernotaite, L., & Malinauskiene, V. (2017). Workplace bullying and mental health among teachers in relation to psychosocial job characteristics and burnout. *International journal of occupational medicine and environmental health*, 30(4), 629–640. <https://doi.org/10.13075/ijomeh.1896.00943>
- Bogaers, R., Geuze, E., Van Weeghel, J., Leijten, F., Rüsche, N., Van De Mheen, D., Varis, P., Rozema, A., & Brouwers, E. (2021). Decision (not) to disclose mental health conditions

- or substance abuse in the work environment: a multiperspective focus group study within the military. *BMJ Open*, 11. <https://doi.org/10.1136/bmjopen-2021-049370>
- Boudrias, V., Trépanier, S.-G., & Salin, D. (2021). A systematic review of research on the longitudinal consequences of workplace bullying and the mechanisms involved. *Aggression and Violent Behavior*, 56, 101508. <https://doi.org/10.1016/j.avb.2020.101508>
- Bunce, A., Hashemi, L., Clark, C., Stansfeld, S., Myers, C.-A., & McManus, S. (2024). Prevalence and nature of workplace bullying and harassment and associations with mental health conditions in England: a cross-sectional probability sample survey. *BMC Public Health*, 24. <https://doi.org/10.1186/s12889-024-18614-7>
- Byon, H. D., Harris, C., Crandall, M., Song, J., & Topaz, M. (2023). Identifying Type II workplace violence from clinical notes using natural language processing. *Workplace health & safety*, 71(10), 484–490. <https://doi.org/10.1177/21650799231176078>
- Chatziioannidis, I., Bascialla, F. G., Chatzivalsama, P., Vouzas, F., & Mitsiakos, G. (2018). Prevalence, causes and mental health impact of workplace bullying in the Neonatal Intensive Care Unit environment. *BMJ Open*, 8. <https://doi.org/10.1136/bmjopen-2017-018766>
- Cox, T., Griffiths, A., & Rial González, E. (2000). *Research on work-related stress*. European Agency for Safety and Health at Work.
- Dahler-Larsen, P., Sundby, A., & Boodhoo, A. (2020). Can occupational health and safety management systems address psychosocial risk factors? An empirical study. *Safety Science*, 130. <https://doi.org/10.1016/j.ssci.2020.104878>
- Daniels, K., Watson, D., Nayani, R., Tregaskis, O., Hogg, M., Etuknwa, A., & Semkina, A. (2021). Implementing practices focused on workplace health and psychological wellbeing: A systematic review.. *Social science & medicine*, 277, 113888. <https://doi.org/10.1016/j.socscimed.2021.113888>
- Deady, M., Sanatkar, S., Tan, L., Glozier, N., Gayed, A., Petrie, K., Dalgaard, V. L., Stratton, E., LaMontagne, A. D., & Harvey, S. B. (2024). A mentally healthy framework to guide employers and policy makers. *Frontiers in Public Health*, 12. <https://doi.org/10.3389/fpubh.2024.1430540>
- Department of Occupational Safety and Health. (2024). *Guidelines on psychosocial risk assessment and management at the workplace 2024*. Ministry of Human Resources Malaysia.
- Dewa C. S. (2014). Worker attitudes towards mental health problems and disclosure. *The international journal of occupational and environmental medicine*, 5(4), 175–186.
- Dewa, C. S., van Weeghel, J., Joosen, M. C. W., Gronholm, P. C., & Brouwers, E. P. M. (2021). Workers' Decisions to Disclose a Mental Health Issue to Managers and the Consequences. *Frontiers in psychiatry*, 12, 631032. <https://doi.org/10.3389/fpsy.2021.631032>
- Dong, R. K., Li, X., & Hernan, B. R. (2024). Psychological safety and psychosocial safety climate in workplace: A bibliometric analysis and systematic review towards a research agenda. *Journal of Safety Research*, 91, 1–19. <https://doi.org/10.1016/j.jsr.2024.08.001>
- Dubosson, M., Fragnière, E., & Meier, S. (2022). Early detection of human-related risks in an increasingly digitized work environment. *Digital Transformation and Society*. <https://doi.org/10.1108/dts-05-2022-0017>
- Easterday, M. W., Lewis, D. R. G., & Gerber, E. (2018). The logic of design research. *Learning: Research and Practice*, 4, 131 - 160. <https://doi.org/10.1080/23735082.2017.1286367>

- European Agency for Safety and Health at Work. (2022). *Managing psychosocial risks in European workplaces*. Publications Office of the European Union.
- Ferrarini, F., & Leonelli, S. (2026). Risks and benefits posed by artificial intelligence to workers' mental health: A systematic literature review. *Industrial Management & Data Systems*, 126(7), 2179–2204. <https://doi.org/10.1108/IMDS-05-2025-0661>
- Galanakis, M., & Tsitouri, E. (2022). Positive psychology in the working environment. Job demands-resources theory, work engagement and burnout: A systematic literature review. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.1022102>
- Galanis, P., Moisoglou, I., Katsiroumpa, A., & Mastrogianni, M. (2024). Association between Workplace Bullying, Job Stress, and Professional Quality of Life in Nurses: A Systematic Review and Meta-Analysis. *Healthcare*, 12. <https://doi.org/10.3390/healthcare12060623>
- Gimpel, H., Lanzl, J., Regal, C., Urbach, N., Becker, J. S., & Tegtmeier, P. (2024). Stress from Digital Work: Toward a Unified View of Digital Hindrance Stressors. *Inf. Syst. Res.*, 36, 896-915. <https://doi.org/10.1287/isre.2022.0691>
- Guo, Z., Lai, A., Thygesen, J., Farrington, J., Keen, T., & Li, K. (2024). Large Language Models for Mental Health Applications: Systematic Review. *JMIR Mental Health*, 11. <https://doi.org/10.2196/57400>
- Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, 32(4), 1008–1015. <https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x>
- Hayat, A., & Afshari, L. (2021). Supportive organizational climate: a moderated mediation model of workplace bullying and employee well-being. *Personnel Review*. <https://doi.org/10.1108/pr-06-2020-0407>
- International Labour Organization. (2016). *Workplace stress: A collective challenge*. International Labour Office.
- Jim, E., Pio, R. J., Asaloei, S., Leba, S. M. R., Angelianawati, D., & Werang, B. R. (2024). Work-Related Stress, Emotional Exhaustion, Job Satisfaction, and Organizational Commitment of Indonesian Healthcare Workers. *International Journal of Religion*. <https://doi.org/10.61707/6fzykj38>
- Julianto, F., Imelda, S., & Tamba, M. (2024). The Influence of Work Stress, Role Conflict and Emotional Exhaustion on Organizational Commitment at PT. Vinsia Bingkai Medan. *Journal Boas: Business, Economics, Accounting And Management*. <https://doi.org/10.54209/boas.v3i01.296>
- Karimova, T. 2026. AI Systems at Work: A Changing Psychosocial Work Environment, ILO Working Paper 170 (Geneva, ILO). <https://doi.org/10.54394/00034098>
- Kawakami, A., Chowdhary, S., Iqbal, S. T., Liao, Q., Olteanu, A., Suh, J., & Saha, K. (2023). Sensing Wellbeing in the Workplace, Why and For Whom? Envisioning Impacts with Organizational Stakeholders. *Proceedings of the ACM on Human-Computer Interaction*, 7, 1 - 33. <https://doi.org/10.1145/3610207>
- Khairuddin, M. Z. F., Hasikin, K., Abd Razak, N. A., Lai, K. W., Osman, M. Z., Aslan, M. F., Sabanci, K., Azizan, M. M., Satapathy, S. C., & Wu, X. (2022). Predicting occupational injury causal factors using text-based analytics: A systematic review. *Frontiers in public health*, 10, 984099. <https://doi.org/10.3389/fpubh.2022.984099>
- Kiaos, T. (2025). Cultural blind spots: Identifying hidden psychosocial hazards in the workplace. *Health Promotion Journal of Australia*.

- Klinefelter, Z., Sinclair, R., Britt, T., Sawhney, G., Black, K., & Munc, A. (2021). Psychosocial safety climate and stigma: Reporting stress-related concerns at work. *Stress and Health*.
- Kyung, M., Lee, S. J., Dancu, C., & Hong, O. (2023). Underreporting of workers' injuries or illnesses and contributing factors: a systematic review. *BMC public health*, 23(1), 558. <https://doi.org/10.1186/s12889-023-15487-0>
- Lainidi, O., Johnson, J., Griffin, B., Koutsimani, P., Mouratidis, C., Keyworth, C., & O'Connor, D. B. (2026). Associations between burnout, employee silence and voice: a systematic review and meta-analysis. *Psychology & Health*, 41(8), 1245–1265. <https://doi.org/10.1080/08870446.2025.2509074>
- 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. <https://doi.org/10.1007/s10551-019-04204-w>
- Leka, S., & Jain, A. (2010). *Health impact of psychosocial hazards at work: An overview*. World Health Organization. Geneva, Switzerland.
- Leka, S., & Kortum, E. (2008). A European Framework to Address Psychosocial Hazards. *Journal of Occupational Health*, 50. <https://doi.org/10.1539/joh.m6004>
- Leka, S., Jain, A., Iavicoli, S., Vartia, M., & Ertel, M. (2011). The role of policy for the management of psychosocial risks at the workplace in the European Union. *Safety Science*, 49(4), 558–564. <https://doi.org/10.1016/j.ssci.2010.02.002>
- Leka, S., Van Wassenhove, W., & Jain, A. (2015). Is psychosocial risk prevention possible? Deconstructing common presumptions. *Safety Science*, 71, 61–67. <https://doi.org/10.1016/j.ssci.2014.03.014>
- Loh, M., Dollard, M. F., McLinton, S., & Brough, P. (2024). Translating psychosocial safety climate (PSC) into real-world practice: two PSC intervention case studies. *Journal of Occupational Health*, 66. <https://doi.org/10.1093/joccuh/uiae051>
- Loretto, W., Platt, S., & Popham, F. (2010). Workplace Change and Employee Mental Health: Results from a Longitudinal Study. *British Journal of Management*, 21, 526-540. <https://doi.org/10.1111/j.1467-8551.2009.00658.x>
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382–385. <https://doi.org/10.1097/00006199-198611000-00017>
- McGrath, M., Krysińska, K., Reavley, N., Andriessen, K., & Pirkis, J. (2023). Disclosure of Mental Health Problems or Suicidality at Work: A Systematic Review. *International Journal of Environmental Research and Public Health*, 20. <https://doi.org/10.3390/ijerph20085548>
- Miguel, C., Amarnath, A., Akhtar, A., Malik, A., Baranyi, G., Barbui, C., Karyotaki, E., & Cuijpers, P. (2023). Universal, selective and indicated interventions for supporting mental health at the workplace: an umbrella review of meta-analyses. *Occupational and Environmental Medicine*, 80, 225 - 236. <https://doi.org/10.1136/oemed-2022-108698>
- Nambisan, S., Lyytinen, K., Majchrzak, A., & Song, M. (2017). Digital innovation management: Reinventing innovation management research in a digital world. *MIS Quarterly*, 41(1), 223–238. <https://doi.org/10.25300/MISQ/2017/41:1.03>
- Oz, T. (2019). Digital Trails of Work Stressors. <https://doi.org/10.31219/osf.io/wxcqp>
- Padzil, M. R., Karim, A. A., & Husnin, H. (2021). Employing DDR to Design and Develop a Flipped Classroom and Project based Learning Module to Applying Design Thinking in Design and Technology. *International Journal of Advanced Computer Science and Applications*. <https://doi.org/10.14569/ijacsa.2021.0120988>

- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Ribeiro, N., Semedo, A., Gomes, D., Bernardino, R., & Singh, S. (2021). The effect of workplace bullying on burnout: the mediating role of affective well-being. *Management Research Review*. <https://doi.org/10.1108/mrr-07-2021-0514>
- Richey, R. C., & Klein, J. D. (2013). Design and development research. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 141–150). Springer. https://doi.org/10.1007/978-1-4614-3185-5_12
- Ricketts, J., Barry, D., Guo, W., & Pelham, J. (2023). A Scoping Literature Review of Natural Language Processing Application to Safety Occurrence Reports. *Safety*, 9(2), 22. <https://doi.org/10.3390/safety9020022>
- Ryan, L., & Cooney, F. (2024). Evaluation of a public health workplace staff psychosocial risk assessment and wellbeing initiative. *The European Journal of Public Health*, 34. <https://doi.org/10.1093/eurpub/ckae144.599>
- Ryan, L., Edwards, D., Lee, J., Connor, D. O., Kane, G., Smith, L., & Cooney, F. (2023). Regional public health staff workplace psychosocial risk assessment and health promotion initiative. *The European Journal of Public Health*, 33. <https://doi.org/10.1093/eurpub/ckad160.131>
- Schaufeli, W. (2017). Applying the Job Demands-Resources model: A 'how to' guide to measuring and tackling work engagement and burnout. *Organizational Dynamics*, 46, 120-132. <https://doi.org/10.1016/j.orgdyn.2017.04.008>
- Shatte, A. B. R., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: A scoping review of methods and applications. *Psychological Medicine*, 49(9), 1426–1448. <https://doi.org/10.1017/S0033291719000151>
- Sivris, K. C., & Leka, S. (2015). Examples of holistic good practices in promoting and protecting mental health in the workplace: current and future challenges. *Safety and Health at Work*, 6(4), 295–304. <https://doi.org/10.1016/j.shaw.2015.07.002>
- Stratton, E., Lampit, A., Choi, I., Gavelin, H. M., Aji, M., Taylor, J., Calvo, R., Harvey, S., & Glozier, N. (2022). Trends in Effectiveness of Organizational eHealth Interventions in Addressing Employee Mental Health: Systematic Review and Meta-analysis. *Journal of Medical Internet Research*, 24. <https://doi.org/10.2196/37776>
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. UNESCO Publishing.
- Uronen, L., Salanterä, S., Hakala, K., Hartiala, J., & Moen, H. (2022). Combining supervised and unsupervised named entity recognition to detect psychosocial risk factors in occupational health checks. *International Journal of Medical Informatics*, 160. <https://doi.org/10.1016/j.ijmedinf.2022.104695>
- Wang, P., Chu, P., Wang, J., Pan, R., Sun, Y., Yan, M., Jiao, L., Zhan, X., & Zhang, D. (2020). Association Between Job Stress and Organizational Commitment in Three Types of Chinese University Teachers: Mediating Effects of Job Burnout and Job Satisfaction. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.576768>
- World Health Organization. (2022). *WHO guidelines on mental health at work*. World Health Organization. <https://www.who.int/publications/i/item/9789240053052>
- World Health Organization. (2010). *Healthy workplaces: A model for action: For employers, workers, policymakers and practitioners*. World Health Organization. <https://www.who.int/publications/i/item/9789241599313>

- World Health Organization, & International Labour Organization. (2022). *Mental health at work: Policy brief*. World Health Organization. <https://www.who.int/publications/i/item/9789240057944>
- Young, I. J. B., Luz, S., & Lone, N. (2019). A systematic review of natural language processing for classification tasks in the field of incident reporting and adverse event analysis. *International journal of medical informatics*, 132, 103971. <https://doi.org/10.1016/j.ijmedinf.2019.103971>
- Zadow, A. J., Dollard, M. F., McInton, S. S., Lawrence, P., & Tuckey, M. R. (2017). Psychosocial safety climate, emotional exhaustion, and work injuries in healthcare workplaces. *Stress and health : journal of the International Society for the Investigation of Stress*, 33(5), 558–569. <https://doi.org/10.1002/smi.2740>
- Zeshan, M., Morelli, M., Rasool, S., Centobelli, P., & Cerchione, R. (2024). Empowering sustainable workplaces: A perspective on employee well-being in the light of the job demand resource model. *Sustainable Development*. <https://doi.org/10.1002/sd.3179>
- Zhang, F., Fleyeh, H., Wang, X., & Lu, M. (2019). Construction site accident analysis using text mining and natural language processing techniques. *Automation in Construction*, 99, 238-248. <https://doi.org/10.1016/j.autcon.2018.12.016>