National Safety Council

CAMPBELL INSTITUTE

Campbell Institute Research Outlook

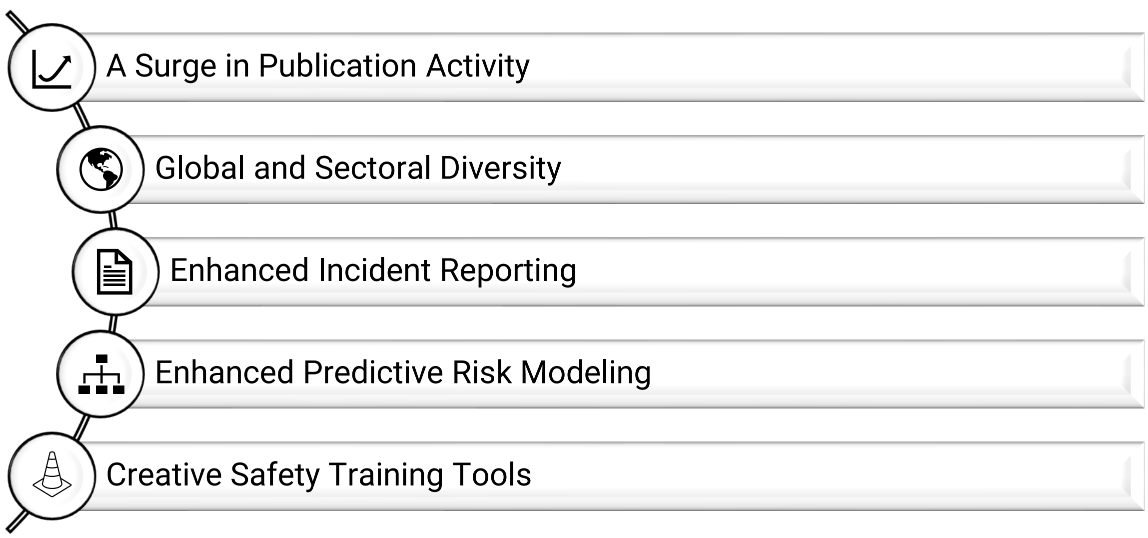
Carolyn Evemy, MA, Research Associate

National Safety Council

Exploring the Role of Generative AI in Occupational Environment, Health and Safety

Executive Summary

This report presents the results of an exploratory literature review investigating the evolving landscape of Generative Artificial Intelligence (GenAI) in occupational environment, health and safety domains. Drawing from a curated pool of 27 peer-reviewed articles, industry reports and grey literature sources published globally between 2018 and 2024, this review aims to capture a snapshot of how GenAI is beginning to make its mark across diverse EHS functions. Several noteworthy trends surfaced from this review:



While there are several innovative use cases and emerging interest in GenAI applications and value within EHS, this review identifies and addresses several challenges in adoption. Skepticism, skill gaps, and ethical and regulatory questions are addressed in this review. Limitations in GenAI are especially important to consider when GenAI tools are making or supporting any safety-critical decisions.

This review is intended as a snapshot rather than a comprehensive, systematic analysis. It draws from both academic and professional sources, but given the fast-moving nature of AI development, the findings may quickly evolve. By offering a high-level, interpretive glimpse into the state of play, it hopes to inspire further reflection, experimentation and strategic foresight as GenAI technologies continue to evolve.

Background

GenAI is increasingly seen as a transformative tool, especially within EHS and safety management systems, where many processes are labor-intensive and burdened by administrative tasks and the challenges of real-time compliance. By automating information processing and providing real-time insights, GenAI has the potential to reduce practitioners' manual workload while improving access to accurate, up-to-date information across all organizational levels (Howard, 2019).

It is important to clarify that **"AI" and "GenAI" are not interchangeable terms** in this context. Many existing EHS AI applications use machine learning for classification or prediction, but they do not create new safety knowledge or content. GenAI, by contrast, focuses on producing new content, such as text, images or reports. GenAI is unique in its use of deep learning models pre-trained on vast amounts of data, such as large language models (LLMs).

GenAI is also different from other safety technologies like computer vision systems or virtual/augmented reality (VR/AR), which rely on pattern recognition or pre-scripted logic rather than generating novel outputs. For instance, a visual AI system that detects personal protective equipment (PPE) violations or an AR overlay for hazard alerts can enhance safety compliance, but these tools do not generate new content or insights on their own. Based on the definition of GenAI provided by the Network of the National Library of Medicine, for the purposes of this review, GenAI is defined as:

> *AI systems that produce new content such as text, speech, images, video or structured outputs by learning from existing data and generating novel outputs.*

Enthusiasm around GenAI is high, but there is a growing consensus that we need to critically examine not just GenAI's technological capabilities but also the ethical and operational risks of its adoption in high-stakes environments.

This review analyzes recent literature on GenAI in EHS to better understand how these opportunities and challenges are being addressed in the present. It aims to identify key themes, gaps and emerging directions that can inform future research and responsible implementation.

Key Findings

Publication Trends: The sample of literature collected for this review indicated that the recent interest in GenAI applications for EHS has accelerated noticeably in the past two years. Of the 27 total academic and grey literature papers identified that addressed specific applications of GenAI within EHS contexts, only a handful of articles (See Figure 1) were published between 2018 and 2022. In 2023, that number rose to 6, and 2024 saw 62% of total publications identified (N = 17). It also hints at the novelty of GenAI tools as a research priority, suggesting that their full potential in the EHS domain is only beginning to be explored and understood.

Identified Academic and Grey Publications by Year

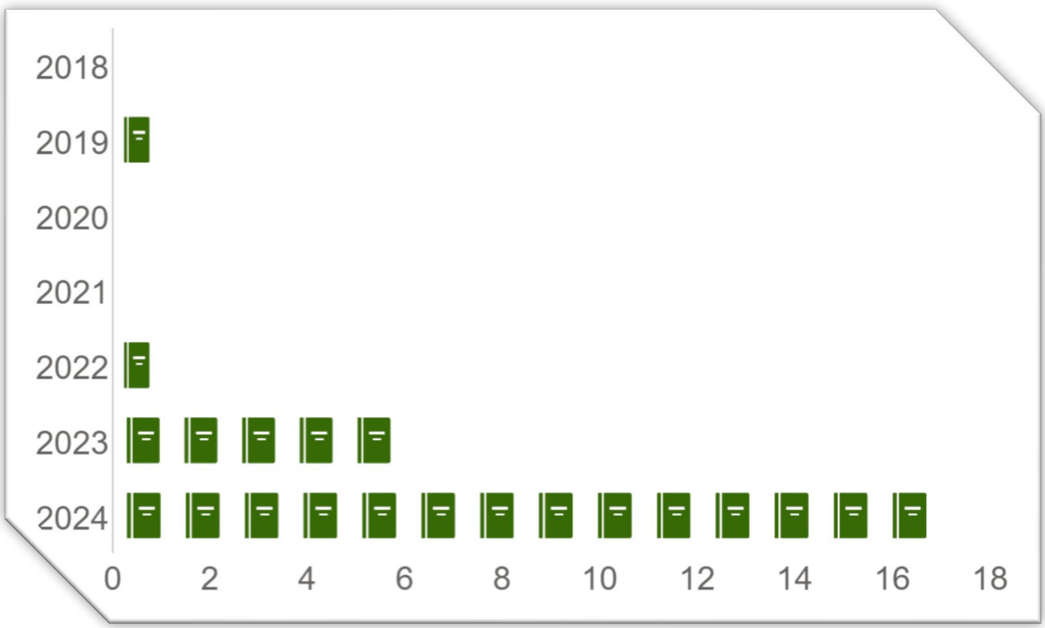


Figure 1. A bar graph where the x-axis represents the year when each identified publication was published, and the y-axis represents the total count of identified publications per year.

Global and Sectoral Diversity: The literature and case studies on GenAI in EHS span diverse regions and industries. Contributions come from North America, Europe, Asia and include insights from international agencies and organizations (e.g., European Agency for Safety and Health at Work (EU-OSHA), International Labour Organization (ILO)) as well as country-specific ones. Likewise, many industry sectors are exploring GenAI in EHS, from construction and manufacturing to agriculture and health care. This broad engagement suggests that interest in GenAI's potential for improving workplace safety and health is global and cross-industry, rather than isolated to a single sector.

Beyond these overarching trends, the literature reveals some major thematic areas where GenAI is being applied or considered in EHS settings:

Incident Reporting and Real-Time Risk Detection: GenAI techniques are being used to improve incident reporting and hazard detection by analyzing unstructured safety reports and logs with natural language processing (NLP) (Weichelt et al., 2023). Academic and grey literature was identified that explored the use of GenAI-enhanced tools to identify hazards from maintenance records.

Some specific examples of incident reporting and real-time risk detection were identified in this review. They include tools like a Virtual Risk Assistant (VRA) tested in Estonian workplaces that demonstrated how rule-based safety systems, augmented by an LLM, can support real-time ergonomic risk assessment. The results of this study also identified that the VRA system generally tended to assign higher hazard levels than their human counterparts (Koppel et al., 2024). While still emerging, the use of GenAI tools for EHS under this capacity allows for streamlining frontline workflows and transforming how safety teams manage risk.

Genny AI, a proprietary GenAI system developed by Benchmark Gensuite, supports frontline teams and program leaders by improving data collection, accelerating summarization and surfacing operational issues faster and more accurately than traditional approaches. Featured in OSHA’s “New Products” brief in 2024, these capabilities enable proactive risk management while enhancing efficiency and performance.

Similarly featured in “New Products” was Protex AI, which launched Protex Copilot, a GenAI safety assistant that empowers EHS and operations teams to shift from reactive to proactive safety management. Protex Copilot’s GenAI-powered features include conversational data analysis, auto-generated toolbox talk presentations, assisted site walks and smart suggestions for corrective actions designed to enhance situational awareness and support faster, data-driven safety decisions.

In parallel, the Human Digital Twin system described by Davila-Gonzalez & Martin (2024) demonstrates how generative AI can be incorporated into real-time worker risk detection by combining physical, emotional, and cognitive data. Within this framework, speech is transcribed and analyzed using NLP and machine learning to extract emotional cues, while OpenAI's GPT-3 (Davinci) model engages workers through conversational assessments to generate real-time evaluations of stress, anxiety, and overall wellbeing. These AI-powered evaluations enable early detection of emerging risks and support timely interventions to improve safety outcomes.

Enhanced Predictive Risk Modeling: Literature on GenAI is beginning to explore enhanced predictive risk modeling by enabling more dynamic simulations, filling gaps in incomplete datasets, and generating synthetic data to train or augment safety models.

In industries such as insurance and heavy manufacturing, these capabilities are helping to improve operational readiness by making risk models more robust, flexible and scenario-rich (McNamara, 2024; Giangiulio, 2023). For example, GenAI might be useful to simulate rare or hypothetical hazard scenarios that traditional data-driven models may not capture, supporting better preparation for low-frequency, high-impact events.

Lange et al. (2024) identify significant challenges related to data availability, sample diversity, and model life cycle completion in workplace AI research. These limitations, which currently constrain traditional ML and DL models, are similarly critical for training robust generative AI models for risk prediction and narrative generation.

Safety Training Materials: GenAI is playing an increasingly important role in the development of customized safety training tools by producing dynamic, context-specific content tailored to diverse workforce needs. These systems can automatically generate job-specific safety scenarios, quizzes and multilingual training manuals, making it easier to deliver accessible and culturally relevant education across global worksites. By leveraging large language models, GenAI enables rapid content creation that reflects site-specific risks, real-time procedural changes or worker demographics.

For instance, Hussain et al. (2024) introduced a GenAI-enhanced virtual reality (AI-VR) training system designed for migrant construction workers. The system generated tailored training content addressing linguistic and contextual barriers, improving engagement and comprehension. This reflects a broader trend toward immersive and adaptive training environments that are increasingly supported by generative technologies.

Although platforms like ABB Electrification Service Assist apps are primarily focused on real-time troubleshooting (Mazzoleni, 2024), they also support field technicians by combining AR-based visual guidance and real-time task instructions. This allows workers to learn procedures on-site while reducing reliance on prior experience or extensive pre-training. GenAI, including LLMs like ChatGPT, is being integrated to provide interactive explanations, answer technical questions during maintenance tasks, and automatically generate documentation that reinforces learning and standardizes procedures (Mazzoleni, 2024).

Ethical and Regulatory Considerations

The integration of GenAI into EHS workflows introduces complex ethical challenges that extend beyond traditional automation concerns. Because GenAI systems are designed to produce novel content such as safety guidance, risk assessments and procedural narratives, they present unique risks related to transparency, accuracy, data privacy and human autonomy. Outputs from LLMs may appear authoritative yet be factually incorrect or contextually misleading, particularly when deployed without adequate oversight.

EU-OSHA (2021) stresses the importance of explainability and worker-centered design when using AI in occupational safety contexts. These principles are especially critical for generative tools, in which the reasoning behind outputs is often opaque. Studies have shown that open-ended GenAI systems like ChatGPT can generate unsafe or misleading safety advice when prompted on complex or ambiguous topics, underlining the necessity of human review, especially in high-stakes environments (Oviedo-Trespalacios et al., 2023; Koppel et al., 2024).

The potential for GenAI tools to operate with false confidence or generate outputs that exceed their programmed scope adds urgency to calls for regulatory frameworks and validation protocols specific to generative systems. Niehaus et al., (2022) argue that keeping a “human-in-control” architecture is vital.

As GenAI continues to be explored in EHS settings, there is a growing consensus that ethical integration must involve technical safeguards, organizational policies, clear roles for human oversight, and accountability structures capable of addressing generative systems' unique risks. While some frameworks have been explored, limitations in GenAI's data-driven approximations and statistical assurances make them more prone to rare but potentially severe failures that are difficult to predict and prevent (Bajcsy & Fisac, 2024).

Organizational Adoption Challenges

While GenAI holds significant promise for transforming EHS operations, its adoption remains uneven due to a range of organizational, technical and cultural barriers. Unlike more deterministic automation tools, GenAI introduces new forms of uncertainty, such as concerns about the accuracy, appropriateness or interpretability of machine-generated content. Adopting these concerns in safety-critical environments could have serious consequences for the workforce if not handled with extreme caution and care.

Common challenges include limited AI literacy among EHS personnel, skepticism toward AI-generated narratives or recommendations, and fears that generative systems could displace human judgment or undermine professional expertise.

Small and medium-sized businesses (SMBs), in particular, face added difficulties in deploying GenAI solutions due to resource constraints, the cost of implementation and unclear compliance pathways amid evolving GenAI regulations (ILO, 2024; El-Helaly, 2024).

Infrastructure also plays a major role: GenAI tools typically require integrated digital workflows, clean and structured data, and robust governance or safety management systems to function effectively. Without these foundations, organizations may struggle to validate GenAI outputs or use them consistently across safety programs.

Despite these hurdles, momentum is building. McCann (2024) highlights construction initiatives that are already using GenAI to support real-time safety decisions, while Rai (2024) points to startups specifically focused on workplace injury prevention through GenAI-powered tools.

These trends are echoed in high-risk domains such as process safety and crisis management, where GenAI is beginning to support risk analysis, response simulations and predictive planning (Jamieson, 2023; Karinshak, 2024).

As GenAI continues to mature, its successful adoption will depend not only on technological capability but also on organizational readiness, including workforce training, cross-functional collaboration, and regulatory awareness.

Emerging Trends and Limitations

The landscape of GenAI in EHS is evolving rapidly, with emerging trends pointing toward deeper integration with already emerging technologies in the safety world like IoT-connected wearables, vision-language models and immersive risk simulations (Park and Kang, 2024; McCann, 2024). These developments position GenAI as a text generator and a multimodal intelligence layer capable of interpreting sensory data, predicting hazards and generating real-time safety insights.

For example, Zhang et al. (2024) employed a vision-language model to assess driver behavior and identify potential risks, demonstrating how GenAI can bridge video input with natural language interpretation. Similarly, Chandra and Chakraborty (2024) applied an LLM to support emergency planning in radiological contexts, an area where the generation of realistic, scenario-specific response guidance is necessary.

These advances suggest a future in which GenAI can contribute to proactive safety management, enabling more adaptive simulations, dynamic training environments and real-time hazard communication tools. However, such a promise is tempered by significant limitations.

An emerging concern is the tendency of GenAI systems to “hallucinate,” producing outputs that are grammatically convincing but factually incorrect or lacking context (Sridi and Brigui, 2023). Even small inaccuracies can lead to operational confusion or decision-making errors in safety-critical settings.

Additionally, data privacy and management were regularly mentioned throughout the literature. The ability of GenAI tools to generate synthetic data or extrapolated scenarios introduces privacy risks and potential regulatory complications, especially when outputs are derived from sensitive worker or operational information (Oviedo-Trespalacios et al., 2023).

Furthermore, some studies highlight that GenAI lacks situational awareness. It cannot fully understand the organizational, emotional or ethical nuance that often surrounds safety incidents. For these reasons, researchers consistently emphasize the importance of human-in-the-loop oversight, strong validation protocols and ongoing monitoring when deploying GenAI in any high-risk environment (Niehaus et al., 2022; Kremer, 2023).

As these tools become more advanced and accessible, the challenge will be technological refinement and the design of governance structures that ensure generative systems are accurate, equitable and safe for practical use in the EHS domain.

Discussion

The application of GenAI in occupational safety and health is emerging but largely experimental. The bulk of implementations reported in the literature are still in conceptual stages or pilot tests, with very few examples of fully validated use in high-risk, real-world settings (especially in U.S. workplaces). In general, data dependence is high: most GenAI-driven safety systems currently rely on a robust digital infrastructure, whether that is networks of IoT sensors, large volumes of quality text data (e.g., incident reports) or real-time video feeds. This underscores that successful GenAI adoption goes hand-in-hand with digital modernization of EHS data collection and management.

A common theme is that interpretability matters for GenAI in EHS. The lack of explainability in many GenAI model outputs presents risks in regulated safety environments where transparency is crucial. If a generative model flags a hazard or recommends an action, safety professionals need to understand why in order to trust and act on that information. Thus, human oversight remains critical. Across studies, AI is consistently framed as a decision-support tool, not a replacement for expert judgment. The role of certified safety professionals in interpreting and validating AI-generated insights is repeatedly emphasized as a safeguard to ensure technology augments, rather than overrides, human expertise.

Across the literature, GenAI systems are developed with clear intentions to solve real problems such as improving training effectiveness, reducing risk exposure, supporting maintenance technicians or enhancing crisis preparedness. However, many solutions rely on design choices that inherently shape how the safety problem is understood and addressed. For instance:

- Safety training may be framed as primarily a content-delivery or language-processing challenge
- Risk management might be modeled through data-driven logic
- Emergency response could be approached as a simulation task

While these framings are valid, they can abstract away important human elements of safety work, such as ambiguity in human communication, the emotional labor of incident response, or power dynamics in decision-making. A more expansive view of GenAI's role might consider not just how to automate or augment existing tasks, but also which aspects of safety work should remain fundamentally human-centered.

Another cross-cutting challenge is that, despite frequent references to issues like algorithmic bias, transparency or data governance, relatively few papers delve deeply into the ethical and regulatory implications of GenAI in EHS. Discussions of ethics in many studies are often brief caveats rather than detailed analyses. Only a small number of publications attempt to directly address challenges such as establishing fail-safes, accountability mechanisms or formal validation protocols for AI in safety-critical applications.

This gap raises concerns about the readiness of GenAI tools for integration into strictly regulated and high-stakes environments. In most cases, the absence of thorough validation or oversight frameworks means that proposed GenAI solutions are not yet fully aligned with the safety assurance standards set by regulating authorities such as OSHA.

An additional insight from the review is the narrow definition of "success" used in many GenAI-EHS studies. Identified metrics of success included processing speed, classification accuracy or the relevance of generated outputs, rather than practical safety outcomes or user acceptance. While technical feasibility is important, focusing solely on efficiency metrics overlooks whether the GenAI actually improves safety on the ground. In environments where trust, compliance and communication are essential to safety performance, a GenAI tool could have excellent algorithmic metrics yet still fail if users do not trust or understand its recommendations.

Finally, this review points to several opportunities for future research and development. One need is for longitudinal studies that observe AI integration in live EHS environments over time, to see how these tools perform beyond the pilot phase and what organizational factors affect success. Developing validation frameworks for safety AI is another priority, such as adapting existing safety standards or certification processes to accommodate AI-driven systems.

There is also room to design more multimodal AI systems that combine text generation with visual or sensor data analysis, reflecting the reality that EHS management often requires integrating multiple data types (reports, images, sensor alerts, etc.). Cross-cultural research should examine GenAI use in different countries or languages, given that safety practices and languages vary widely and most current studies are limited in geographic focus.

Perhaps most importantly, a participatory design approach would be beneficial: involving EHS practitioners in the development of GenAI tools from the outset. Such collaboration can help ensure that the technology addresses real frontline needs and accounts for the complex social and operational realities of safety work, rather than being driven solely by technological push.

Engaging end-users in design and testing could improve trust and relevance, ultimately making GenAI a more effective partner in promoting workplace health and safety. This review centers on a subset of GenAI use cases most relevant to EHS practice today. Future work might expand to explore image and video generation, code synthesis or design-focused tools, which are also emerging within the broader GenAI landscape.

References

Bajcsy, A., & Fisac, J. F. (2024). Human–AI safety: A descendant of generative AI and control systems safety. *arXiv preprint arXiv:2405.09794*, 2024. https://doi.org/10.48550/arXiv.2405.09794 .

Chandra, A., & Chakraborty, A. (2024). Exploring the role of large language models in radiation emergency response. *Journal of radiological protection : official journal of the Society for Radiological Protection*, *44*(1), 10.1088/1361-6498/ad270c. https://doi.org/10.1088/1361-6498/ad270c

Davila-Gonzalez, S., & Martin, S. (2024). Human Digital Twin in Industry 5.0: A Holistic Approach to Worker Safety and Well-Being through Advanced AI and Emotional Analytics. *Sensors (Basel, Switzerland)*, *24*(2), 655. https://doi.org/10.3390/s24020655

El-Helaly M. (2024). Artificial Intelligence and Occupational Health and Safety, Benefits and Drawbacks. *La Medicina del lavoro*, *115*(2), e2024014. https://doi.org/10.23749/mdl.v115i2.15835

European Agency for Safety and Health at Work (EU-OSHA). (2021). Impact of Artificial Intelligence on Occupational Safety and Health. Retrieved from https://osha.europa.eu/en/publications/impact-artificial-intelligence-occupational-safety-and-health

Giangiulio, D. (2023). *4 Ways Generative AI Will Impact the Supply Chain*. Retrieved from https://www.mhlnews.com/technology-automation/article/21267817/4-ways-generative-ai-will-impact-the-supply-chain

Howard J. (2019). Artificial intelligence: Implications for the future of work. *American journal of industrial medicine*, *62*(11), 917–926. https://doi.org/10.1002/ajim.23037

Hussain, R., Sabir, A., Lee, D. Y., Zaidi, S. F. A., Pedro, A., Abbas, M. S., & Park, C. (2024). Conversational AI-based VR system to improve construction safety training of migrant workers. *Automation in Construction, 160*, 105315. https://doi.org/10.1016/j.autcon.2024.105315

International Labour Organization, & United Nations. Office of the Secretary-General's Envoy on Technology. (2024). Mind the AI divide: shaping a global perspective on the future of work (p. 22 p.). ILO/UN. https://doi.org/10.18356/9789211066524

Jamieson, D. (2023). *AI's Journey to Becoming the Best Process Safety Engineer in the Room*. The Chemical Engineer. https://www.thechemicalengineer.com/features/ai-s-journey-to-becoming-the-best-process-safety-engineer-in-the-room/

Karinshak, E. (2023). Simulations in human-AI crisis management systems. *Journal of Contingencies and Crisis Management*, vol 32, e12525. https://doi.org/10.1111/1468-5973.12525

Park, J., & Kang, D. (2024). Artificial Intelligence and Smart Technologies in Safety Management: A Comprehensive Analysis Across Multiple Industries. *Applied Sciences*, *14*(24), 11934. https://doi.org/10.3390/app142411934

Koppel, T., Tšernikova, O., & Kalvet, T. (2024). *A rule-based risk assessment tool based on large language model* [Presentation of Virtual Risk Assistant (VRA) tool at ISPIM 2024 conference].

Kremer, A., Luget, A., Mikkelsen, D., Soller, H., Strandell-Jansson, M., & Zingg, S. (2023). *As gen AI advances, regulators—And risk functions—Rush to keep pace*. Retrieved from https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/as-gen-ai-advances-regulators-and-risk-functions-rush-to-keep-pace

Lange, M., Löwe, A., Kayser, I., & Schaller, A. (2024). Approaches for the Use of AI in Workplace Health Promotion and Prevention: Systematic Scoping Review. *JMIR AI*, *3*, e53506. https://doi.org/10.2196/53506

Mazzoleni, A. (2024). Immersive Technology Boosts Service. Manufacturing Management, 23–24.

McCann, T. (2024). Realizing a Construction Site Transformation. *ENR: Engineering News-Record*, *292*(7), 36. Retrieved from https://research.ebsco.com/linkprocessor/plink?id=7ce2c9ff-332e-3f0b-b5f8-efe88d07b869.

McNamara, P. (2024). AI and the aggregation of risk. *Asia Insurance Review*, 50–51. Retrieved from https://research.ebsco.com/linkprocessor/plink?id=8fe293ee-5e9b-3199-959e-c20ed7bed562

New Products. (2024). *Occupational Health & Safety*, *93*(3), 94–95. Retrieved from https://research.ebsco.com/linkprocessor/plink?id=6c7e311f-d5fe-3d3b-81f1-6a204a0cd191

Oviedo-Trespalacios, O., Peden, A. E., Cole-Hunter, T., Costantini, A., Haghani, M., Rod, J. E., Kelly, S., Torkamaan, H., Tariq, A., David Albert Newton, J., Gallagher, T., Steinert, S., Filtness, A. J., & Reniers, G. (2023). The risks of using ChatGPT to obtain common safety-related information and advice. *Safety Science*, *167*, N.PAG. https://doi.org/10.1016/j.ssci.2023.106244

Rai, S. (2024). *AI Startup Gets Funds to Help Prevent Widespread Work Accidents*. Bloomberg. Retrieved from https://www.bloomberg.com/news/articles/2024-02-27/ai-startup-gets-funds-to-help-prevent-widespread-work-accidents?embedded-checkout=true

Niehaus, S., Hartwig, M., Rosen, P. H., & Wischniewski, S. (2022). An Occupational Safety and Health Perspective on Human in Control and AI. *Frontiers in artificial intelligence*, *5*, 868382. https://doi.org/10.3389/frai.2022.868382

Sridi, C., & Brigui, S. (2023). The use of ChatGPT in occupational medicine: opportunities and threats. *Annals of occupational and environmental medicine*, *35*, e42. https://doi.org/10.35371/aoem.2023.35.e42

Weichelt, B. P., Pilz, M., Burke, R., Puthoff, D., & Namkoong, K. (2024). The Potential of AI and ChatGPT in Improving Agricultural Injury and Illness Surveillance Programming and Dissemination. *Journal of agromedicine*, *29*(2), 150–154. https://doi.org/10.1080/1059924X.2023.2284959

Zhang, K., Wang, S., Jia, N., Zhao, L., Han, C., & Li, L. (2024). Integrating visual large language model and reasoning chain for driver behavior analysis and risk assessment. *Accident; analysis and prevention*, *198*, 107497. https://doi.org/10.1016/j.aap.2024.107497

Appendix 1. Methods

This literature review was conducted to assess the current state of research on the application of GenAI within EHS domains. The search strategy was developed using a combination of peer-reviewed and grey literature sources. Literature was retrieved via the EBSCOhost subscription-based service, utilizing the following databases:

- Academic Search Complete – Scholarly articles across many subjects
- APA PsycINFO – Research in psychology and related fields
- MEDLINE Complete – Medical and health science journals
- Regional Business News – Local business news sources
- Business Source Ultimate – Business and economics articles

Open-access sources captured through stakeholder engagement and were also included to capture recent industry insights and emerging trends. Rayyan software was used to support data collection, abstract screening and literature management.

A total of 384 records were screened following the removal of 13 duplicates. Four researchers independently reviewed the titles and abstracts to determine relevance according to predefined inclusion criteria, focusing on the intersection of GenAI and EHS. One researcher conducted full-text reviews of 79 reports, from which 27 studies were ultimately included in the final analysis. Exclusion criteria included articles not addressing EHS or AI and studies lacking a substantive discussion of generative AI tools. This review is interpretive rather than exhaustive, offering a thematic synthesis of current findings to support strategic awareness and decision-making among EHS professionals.

Contact Us:
Carolyn.Evemy@nsc.org

Learn More:
www.thecampbellinstitute.org

National Safety Council