Protex AI

From Chaos to Clarity

Mastering Unstructured Data in Health and Safety Management

Whitepaper

INTRODUCTION

The digital revolution has led to an exponential growth of data sources – from social media data to wearables data. Because of this, the health and safety sector is bursting with data that is unstructured and siloed, making it challenging for health and safety teams to draw actionable insights.

Each day, more and more organizations move to implement and deploy technology with the intention of refining operational health and safety, changing behaviors, and improving health and safety metrics – achieving these outcomes is not a given, linear, or correlated. This whitepaper aims to shed light on the importance of data integration and analysis, offering a path to transform unstructured data into valuable insights for proactive health and safety management.

Speaking of structure, let's bring some to this narrative by defining what is structured and unstructured data. Not all data is the same.

Structured data is organized and decipherable data, typically known as quantitative data. This data is typically managed through structured query language (SQL), users can quickly input, search and manipulate data – for example this can be details from a site OHS induction – name, address, company, etc. Structured data is easily used by machine learning algorithms, it does not require an in-depth understanding of the data and it can be easily interpreted, as there are many tools available to analyze the data. Typically, structured data has limited usage and can really only be used for its intended purpose.

Structured Data	Unstructured Data
Structured data stands for information that is highly organized, factual, and to-the-point	Unstructured data doesn't have any predefined structure to it and comes in all its diversity of forms
Quantitative	Qualitative
Data warehouses relational databases	Data lakes non-relational databases
Several predetermined formats	A huge array of formats

Unstructured data cannot be processed and analyzed using conventional data tools and methods. This data is typically known as qualitative data. This type of data cannot be in an SQL tool, it is best managed through data lakes, so that the data can be preserved in its raw form. Examples of unstructured data include text from safety observation submissions, Internet of Things (IoT) data from worker locations, working patterns, sensor data from wearables, vehicles and artificial intelligence such as safety chat bots. Most data that exists is unstructured data, by 2025 80% of global data will be unstructured, with 95% of businesses prioritizing unstructured data management.

Moving from compliance to commitment and measuring safety in a contemporary way involves a holistic approach that considers both leading and lagging indicators, this is critical when understanding and using unstructured data in the health and safety arena. To effectively impact safety performance and culture, we must move from lagging indicators and structured data to leading indicators and unstructured data. This is uncomfortable, uncharted and uneasy for many, but it can be done, where there is a plan in place for utilizing unstructured data.

Exploring how unstructured data is present in the workplace and the challenges to health and safety.

Advanced technologies such as computer vision, IoT and virtual reality are being leveraged by organizations to improve workplace health and safety. When deployed, implemented, and embedded, technology solutions can create efficient and effective operations, reducing operational and safety risk, and improving operational health and safety. When deploying a technology solution, some workplaces focus more on what the solution is going to resolve, and not necessarily the metrics and data that it may generate, and how that data can be used. There are many risks with deploying advanced technologies, especially those that generate unstructured data. Unstructured data often lacks standardization and needs significant pre-processing and feature extraction efforts – this, if not managed correctly, can pose and create many operational and cultural health and safety challenges.

SCENARIO

Let's take computer vision for example, when used in a health and safety application, a computer vision camera is installed in work area, health and safety rules are established and defined. For example, CCTV capture workers who are not wearing the required personal protective equipment (PPE) in a specific zone, the computer vision technology is capturing the rule and the rule breaches through images, audio and video and then notifying operational leaders of rule breaches, there are some challenges that come with this.

Work integration:
If a rule is being breached, it's essential to establish system processes to ensure that data is utilized in real-time and integrated into standard operating procedures. Managers should be promptly notified to ensure that the safety breach is addressed and resolved effectively.

Overload of data and an absence of action:
Computer vision (CV) provides eyes in the back of everyone's head, until it starts identifying that there is a systemic behavioral issue in certain work zones, and managers receive far too many notifications of safety breaches that they too stop responding to them. The challenge then shifts from detecting to effectively managing and prioritizing these alerts. A proactive approach is necessary, focusing on prioritizing alerts based on severity and potential impact, enabling a strategic response to safety breaches. By refining CV systems to filter and highlight critical data, we can transition from a reactive to a proactive stance in safety management, ensuring that every alert is both noticed and actionable.

Credibility and quality of data:
Typically, unstructured data comes in large volumes, and when processing such high volumes of data there is a risk of data errors, which can beg the question of the credibility and quality of the data.

Digital literacy of the operational and safety teams:
With such a large volume of data being generated during a shift, teams may not know how to interpret the data and/or understand what to do with the data in an operational sense.

Scaling data from tactical to strategic:
Unstructured data is very large and complex, which makes it challenging to scale from a site to an enterprise, particularly when working on systemic issues or enterprise wide strategies.

The adoption of computer vision technologies, IoT, sensors, and wearables represents a comprehensive win and a progressive approach to health and safety. However, these technologies also pose challenges, such as the accumulation of potentially irrelevant data, underutilization of data, lack of data quality control, or failure to capture the right data to address specific health and safety issues. Overcoming these challenges necessitates a tailored strategy and capable resources with the expertise to identify, develop, and implement solutions for advancing operational health and safety in the future.

Strategies for Data Integration

Deploying technology is often believed to be the solution to reducing risk, eliminating safety incidents and improving workplace health and safety behaviors. When technologies are simply implemented in a business because they are new shiny tools, or they are not linked to the health and safety strategy, or a safety business problem, it can cause many issues – particularly in relation to unstructured data.
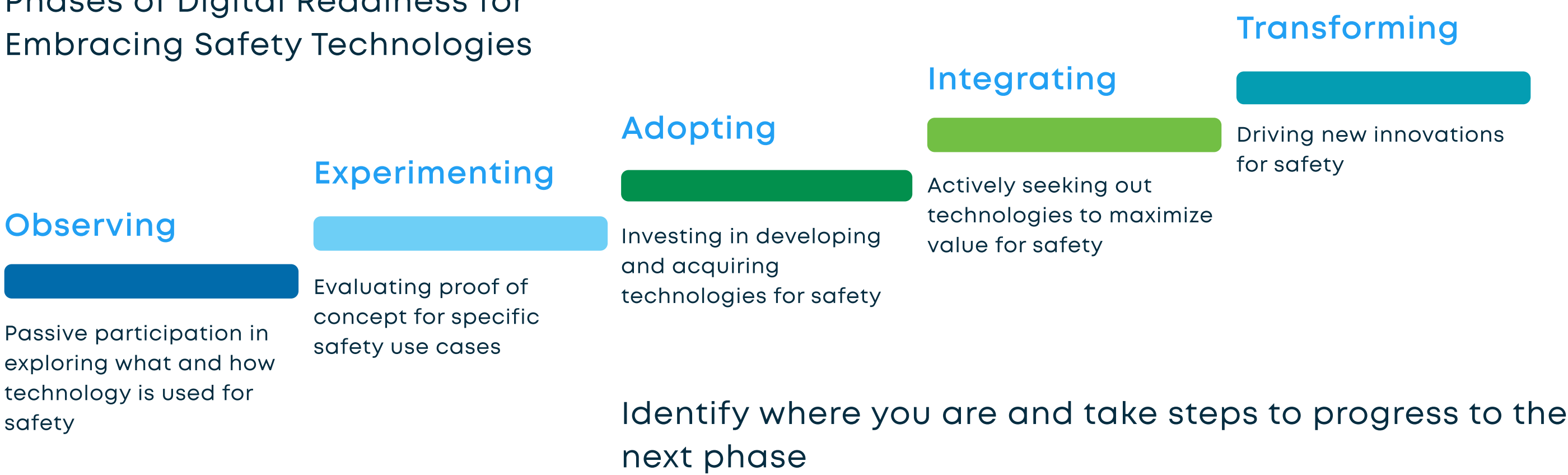
Health and safety leaders must be intentional in crafting their digital strategies and roadmaps, establishing robust system architectures, and determining which metrics and data are crucial both presently and in the future. It's imperative for these leaders to define the role of technology in this proces. To achieve a successful, data-driven Occupational Health and Safety (OHS) program, organizations must adopt a holistic and organized approach to their data strategies, empowering a data-driven culture in health and safety initiatives.

As part of developing a digital strategy, health and safety leaders should assess their own, their teams', and the business's current digital literacy levels and determine the desired proficiency. This understanding is crucial for aligning with planned technology projects and leveraging the data generated by these solutions to measure safety performance, inform the design of health and safety management systems, and shape future strategies.

Phases of Readiness

In considering safety innovation for Industry 4.0, a five-phase model can be used to help determine progress in organizational readiness.

Phases of Digital Readiness for Embracing Safety Technologies

Transforming

Driving new innovations for safety

Integrating

Actively seeking out technologies to maximize value for safety

Adopting

Investing in developing and acquiring technologies for safety

Experimenting

Evaluating proof of concept for specific safety use cases

Observing

Passive participation in exploring what and how technology is used for safety

Identify where you are and take steps to progress to the next phase

Once the digital literacy is understood and digital safety projects are mapped out, the team should develop specific change management governance and plans. The NSC has built a ready-made tool to assess digital literacy and readiness **https://www.nsc.org/ digitalreadiness**.

Knowing the level of digital literacy will help you and your teams understand how to use and analyze data. There are many ways to analyze data, and this must be built into the strategy, to ensure a continuous improvement cycle (more about how to analyze unstructured data later on). With the capturing of any data, it is important to communicate what the data is telling us, and it can also be used to influence senior leaders on safety outcomes needed within the business – a concept termed 'numbers to narrative' will help elevate health and safety and digital solutions within a business.

The Role of AI and Analytics

Generative AI is altering the way we approach unstructured data. With the arrival of large language models such as ChatGPT, unstructured data sets can now be analyzed and consolidated at rapid speeds in ways that were previously impossible. This has led to a significant shift in how businesses view unstructured data, positioning it as a critical enabler of innovation. A substantial advantage of generative AI is its ability to unlock previously inaccessible insights from unstructured data. Traditionally difficult to analyze due to its unstructured nature, this data, with the assistance of generative AI, now reveals patterns and trends that were once hidden. These insights lead businesses to clear actions and next steps, benefiting industries across the board. Particularly in health and safety, analyzing large-scale data such as safety observations overlaid with incidents enables the identification of systemic issues across workplaces or enterprises. With these insights, companies can make more informed decisions about risk reduction, safety communication, and focus areas, as well as gain deeper insights into workforce dynamics, safety experiences, training requirements, and workplace design.

With the assistance of generative AI, businesses can deliver critical insights to C-suite executives with unparalleled speed, enabling swift assessment and decision-making even in last-minute requests or emergency situations. For instance, in scenarios requiring rapid risk evaluation, generative AI systems empower executives to generate comprehensive reports with a simple click. Whether executives are identifying trends in safety non-compliance, assessing risk levels across different shifts and locations, or pinpointing systemic safety issues, this agility ensures prompt response capability. Even on short notice, executives have access to the insights needed to guide strategic decisions, reinforcing the organization's commitment to maintaining a safe and compliant work environment while enhancing operational efficiency

What was the latest event in unit 2A today?

The most recent event in Unit 2A was 'Person walking to close to Forklift."
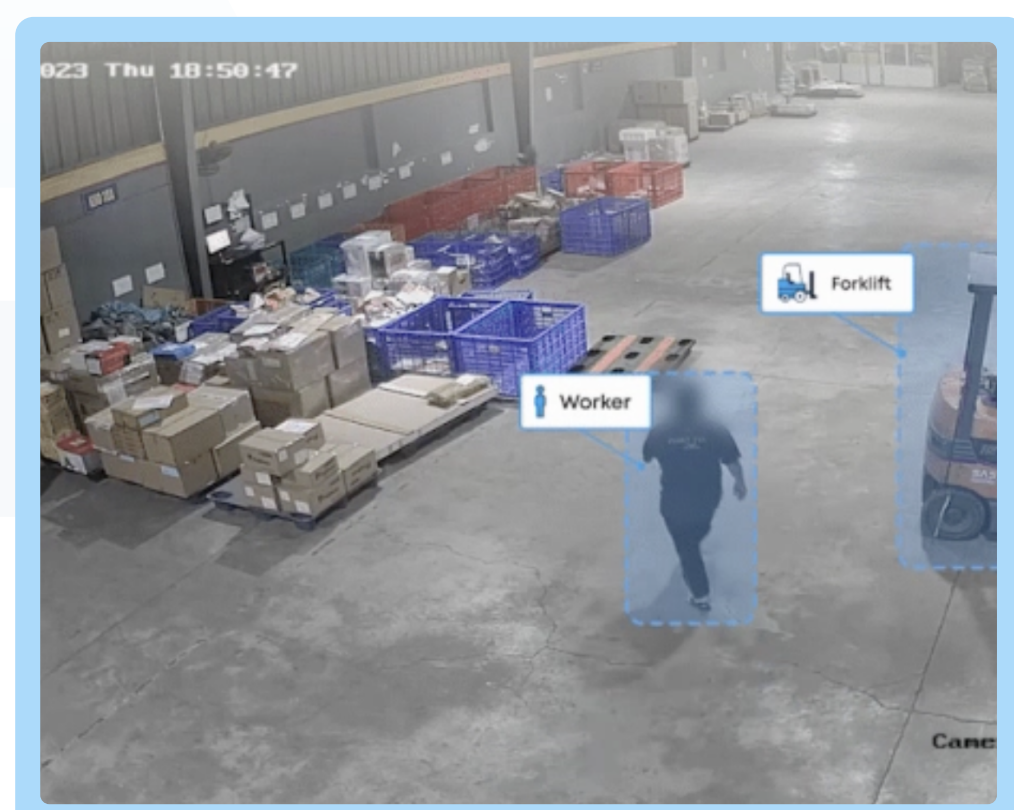
Actionable Insights

For businesses, generative AI delivers unthinkable opportunities. Here are a few ways unstructured data can be used to collect actionable insights which drive behaviors, reduce operational risk, and create operational efficiencies:

Data Mining for Leaders and Operators Behavior:
Utilize text mining algorithms on safety observations submitted by workers, close out rates, incidents and injuries to uncover hidden patterns, trends, and preferences. This can inform/identify work areas requiring physical site inspections and operational leaders who require behavioral coaching and support.

Automated Toolbox / Daily Briefing Generation:
Employ generative AI models to create work zone/site specific safety briefings based on observations, near misses, audit results, plant inductions and risk register. This will create bespoke and extremely relevant safety talking points to drive engagement and relevant conversations.

Predictive Analytics for Market Trends:
Analyze unstructured data from incident and injury reports, and unactioned observations/hazard reports to predict incidents.

Enhanced Risk Management:
Implement risk management analysis on unstructured data like toolbox talks, hazard reports, safety committee meetings, risk registers, SWMS reviews and training compliance and completion to help with proactive risk assessment.

Employee Safety Sentiment and Experience:
Use AI to analyze how the workforce feels about the level of safety and the safety culture at the workplace through induction feedback, training feedback, toolbox talks, employee surveys and hazard observations.

Operational Efficiency:
Apply machine learning models to analyze unstructured operational data (like logs, sensor outputs) for predictive maintenance, reducing downtime, using unsafe equipment and improving efficiency.

Procedural Breaches:
Apply machine learning models to analyze unstructured operational data (like logs, sensor outputs) for insights on work areas with procedural breaches, near misses and incidents.

CONCLUSION

The pursuit of a safer, more technology-focused workplace has led to the emergance of data driven safety, which is shaping the redesign of workplace health and safety management systems. With more technology, comes more unstructured data, something which is becoming the norm and plays an important role in taking action on safety. To maximize the benefits of unstructured data, business and health and safety teams must have a detailed digital strategy which encompasses the collection, and use of unstructured data. Businesses must be aware of their digital literacy and readiness and be open to improving this, whilst continuing to deploy advanced technologies to reduce risk and enhance operations.
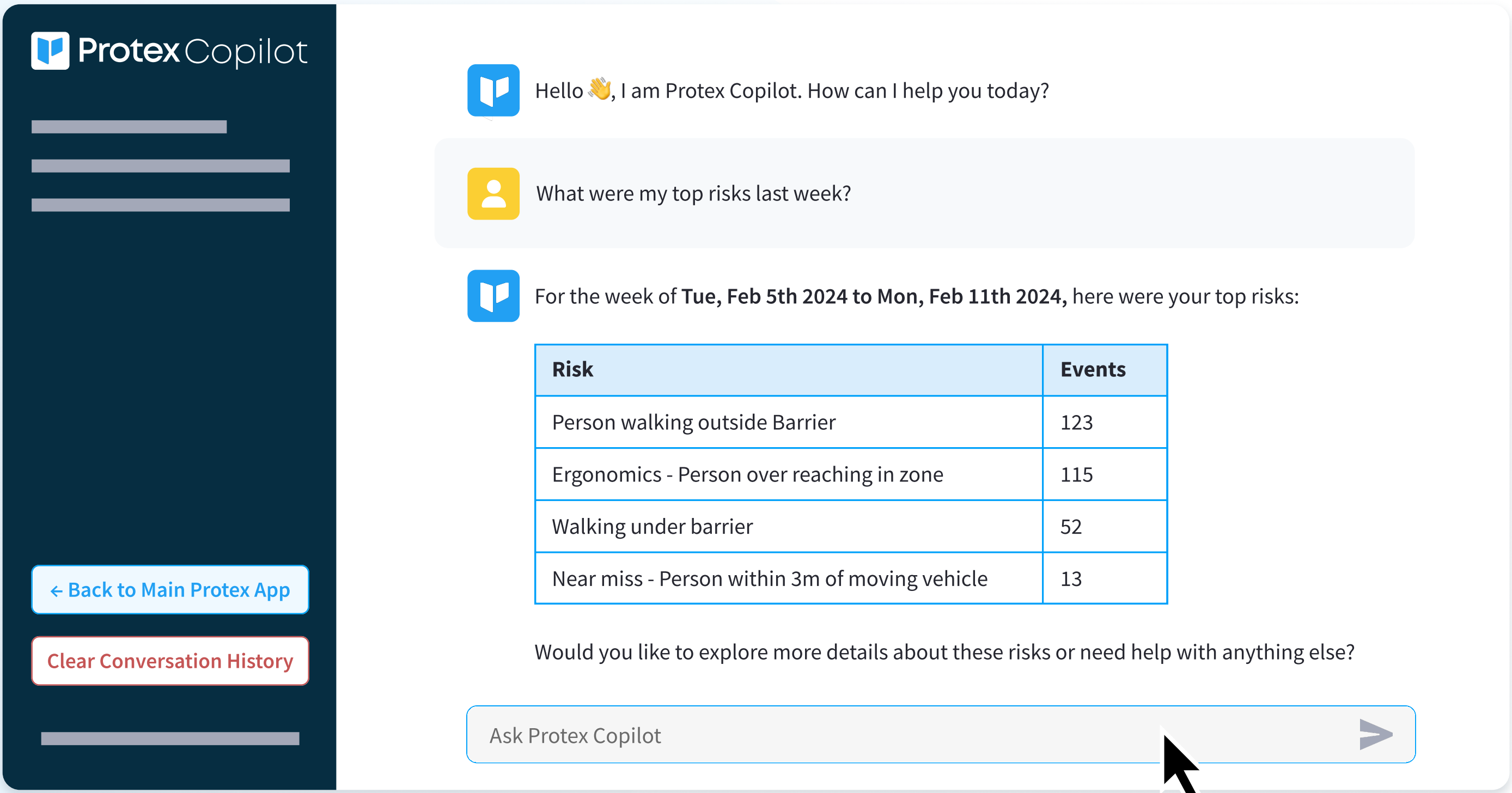
REDEFINING SAFETY MANAGEMENT:

Protex Copilot is the world's first generative AI assistant for health and safety, transforming how EHS teams manage workplace safety. Protex Copilot simplifies the handling of complex data, offering tailored insights that streamline daily operations and enhance safety measures.

With Protex Copilot, generating reports and diving deep into incident analysis becomes intuitive and efficient, aiding in the prevention and identification of safety issues. It elevates on-site evaluations with real-time incident replays, equipping teams with the data and visuals needed for precise improvements. Moreover, its smart, data-driven recommendations guide EHS leaders in prioritizing and tackling safety concerns with confidence.

But this is just the beginning. Protex Copilot's generative AI capabilities mean that the possibilities for enhancing workplace safety are virtually limitless. Need insights or have a safety query? Just ask Protex Copilot. It's designed to adapt to your needs, offering solutions to even the most complex safety challenges. Experience the future of health and safety management with Protex Copilot, where asking questions is the key to unlocking a safer workplace.

Ready to explore the possibilities of integrating AI into your safety strategy?

Contact our product experts