EXECUTIVE BRIEFING

Using AI in Construction to Reduce Injuries and Save Lives

DISCOVERING SAFETY

bam nuttall

mind FOUNDRY

Oxford Semantic Technologies

DISCOVERING SAFETY

bam nuttall

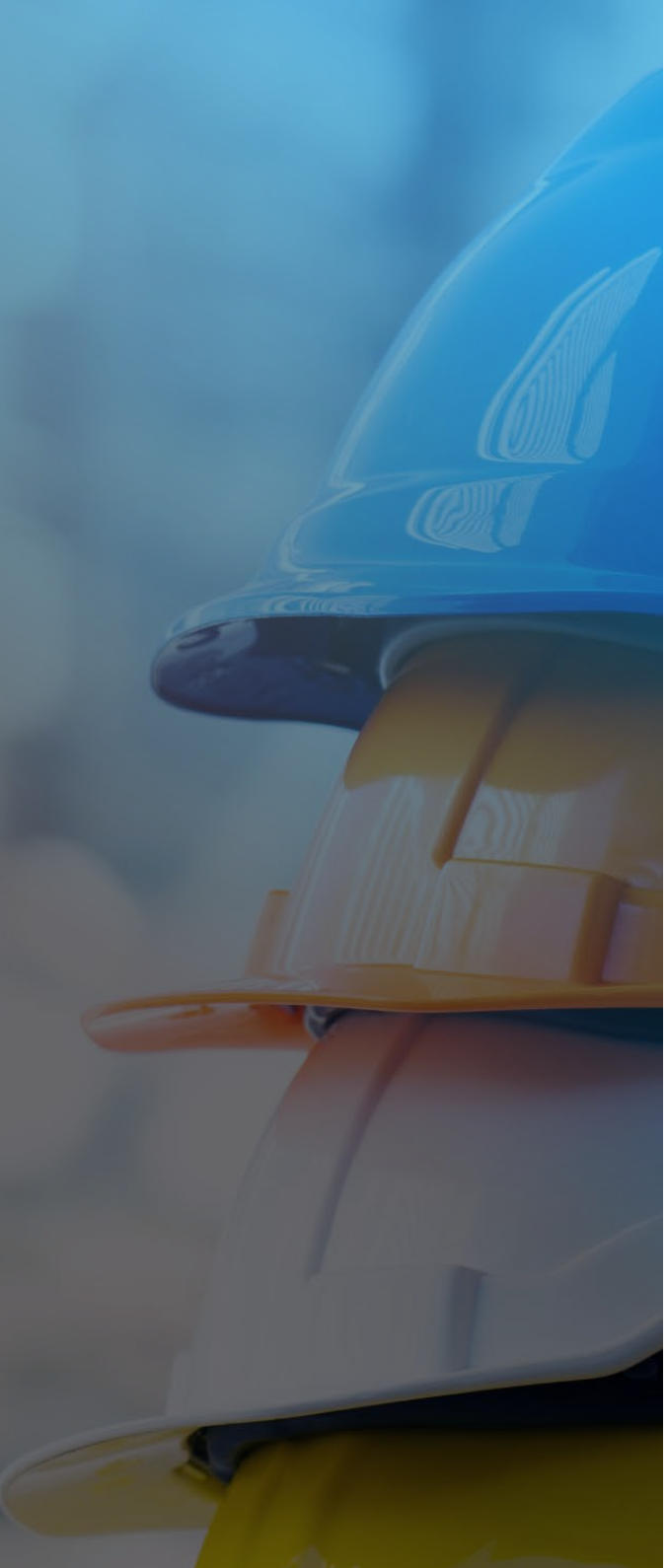
About This Study

HSE in collaboration with BAM Nuttall and two Oxford University spinouts, Mind Foundry and Oxford Semantic Technologies, deployed AI to detect hidden patterns in health and safety incidents on UK construction sites.

mind FOUNDRY

Oxford Semantic Technologies

TABLE OF CONTENTS



EXECUTIVE SUMMARY

Patterns of hazardous conduct have become increasingly more difficult to find and therefore more difficult to target and prevent. Health and safety experts need to adapt and innovate—to apply the latest in data science and Artificial Intelligence to continue to push the trend of health and safety incidents downward.

To this end, BAM Nuttall, in collaboration with the Health and Safety Executive (HSE), formed a new joint project delivered as part of HSE's Discovering Safety Programme (DSP). The DSP, supported by a grant from Lloyd's Register Foundation, aims to bring about step-change improvements in global health and safety performance through the application of advanced data analytics and AI on routine data.

As part of the DSP, HSE invited two leading technology partners, both spinouts from the University of Oxford, Mind Foundry and Oxford Semantic Technology (OST), to tackle the complex technical issues at hand. OST would apply its proprietary knowledge graphing technology, RDFox, to curate a master knowledge graph from complex data sets that health and safety experts could interrogate using Mind Foundry's end-to-end Machine Learning Platform. In a matter of weeks, Mind Foundry enabled health and safety users to uncover meaningful insights in their data and develop predictive models related to project inspection performance.

ABOUT:

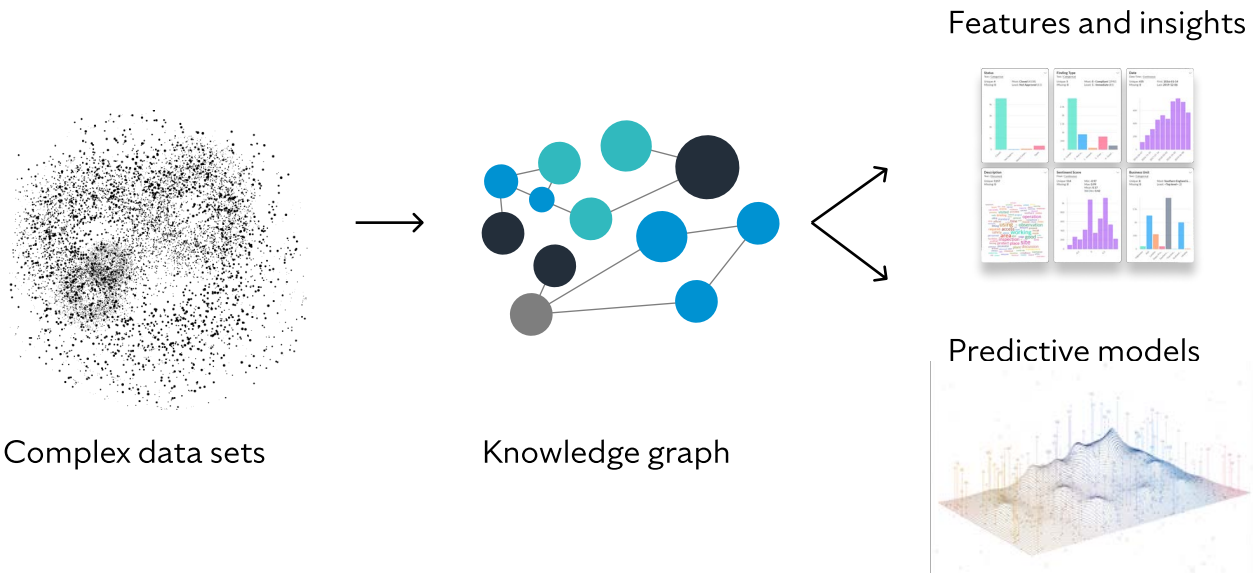
Mind Foundry is an Oxford University spinout, founded by Professors Stephen Roberts and Michael Osborne, pioneers in the field of AI and Machine Learning.

With a deep understanding of how important these technologies are to our future, the mission of Mind Foundry is to enable Humans and AI to work together to solve the world's most important problems.

Mind Foundry has developed technology and products that help people bring machine learning closer to their work. Built upon a foundation of scientific principle, organizations use Mind Foundry to empower their teams in entirely new ways.

mindfoundry.ai

These insights revealed where action could be taken to strengthen H&S culture in construction sites. Likewise, they suggested which data would be valuable to collect in order to unlock more insights.



KEY INSIGHTS FROM THE DATA

Though the dataset was limited, many valuable insights and takeaway messages were uncovered.

	Insight	Takeaway
Time of year + Type of accident	Strong correlation between accidents and seasons: lacerations in summer; eye injuries and bruises in winter.	Should Safety Summits (typically held in winter) be moved closer to summer? What else can be done to ensure workers wear the appropriate equipment at the right time of the year?
Joint ventures + Safety reporting	Construction projects undertaken by joint ventures tend to have a lower proportion of negative findings during inspections.	This seems counter intuitive to the idea that joint construction ventures would be more risky due to unfamiliar collaborations.
Near miss data + Predictive models	Near-miss data could be the key to allow for more types of predictive models to be built.	By collecting this data looking forward, accident risk sources could be more accurately identified and mitigated, creating quantifiable and actionable predictions that could help make better decisions on the site.
Text data + Sentiment analysis	Value in text data, often overlooked due to the recurring notion that it's difficult to deal with, was unlocked through NLP.	This powerful Machine Learning tool enabled the creation of trend charts to track project inspection performance over time directly from raw text written by inspectors in their reports.

TECHNOLOGY BENEFITS

Key highlights from the project include:

Assembly of a new master knowledge graph from numerous H&S data sources with poorly structured relationships between them.

Enrichment of data sets using state-of-the-art Machine Learning to calculate risk profiles, morale and sentiment scores from text, and predict inspection performance scores.

Identifying key project attributes that best estimate project risk profiles and likelihood of health and safety hazards.

Uncovering a number of health and safety issues recurring on UK construction sites, resulting in immediate opportunities for data-driven insights to improve decision-making.

Using advanced AI and data science technologies need not be overly time consuming or require specialist expertise.

LOOKING AHEAD

Beyond these specific findings, the pilot also proved that advanced AI and data science technologies didn't require additional time or specialist expertise. Mind Foundry and OST technologies offered meaningful health and safety insights, whilst also being easy to deploy and performing in a stable and reliable way, as confirmed by a cross-section of existing BAM and HSE technical and health and safety experts.

3 Key uses to take forward.
Many more to be explored over time.

Predictive insights

Predict inspection outcomes based on projects proven to be most at risk, optimising time and effort.

Pattern detection

Test assumptions and tune strategies by detecting patterns in the data quickly.

Semantic analysis

Add textual insights and quantify leading factors like morale and risk, improving predictions.

NEXT TIME
Increase the frequency & volume of data to assemble, enrich, and analyse.

As a result of this pilot, it is clear that improving data collection methods and frequency, integrating more diverse data, automating workflows, and deploying predictive analytics will build on these foundational steps and preliminary findings.

Feasible near-term benefits in the future include:

- Utilising free-text observation data to identify and resolve problems pre-inspection.
- Using these patterns to predict 'at-risk' sites and better target inspections and training.
- Change the way key decisions are made, by relying on predictive analytics.
- Collecting near-miss data to enable the creation of more powerful predictive models to prevent accidents and isolate sources of risk.
- Giving time back to inspectors and health and safety experts to innovate and focus on other health and safety effects on projects over time, such as using sentiment analysis to trend morale on-site, or discover the effects as sites grow in size or are managed jointly with partners.

Ultimately, in the future machine learning can be used before planned work to ascertain what extra risks may be associated with the task and its performance.

This will be a paradigm shift away from relying solely on safety cultures and toward informed decision making with a quantifiable value of risk understood from the planning process that helps us make better decisions about the lifecycle of construction management.

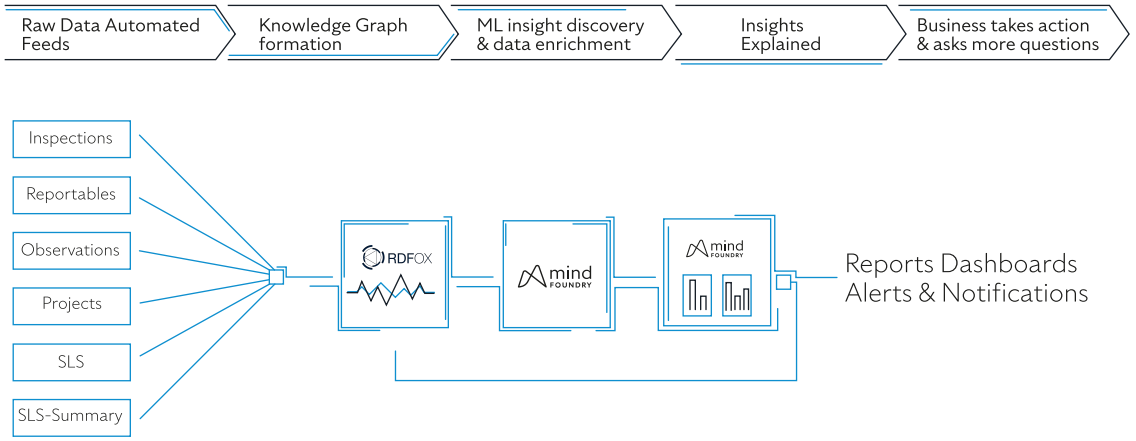
Whilst the pilot proved some of these benefits, more work is to be done to continue to unlock health and safety benefits using knowledge graph and machine learning technologies. This paper will go into further detail about the overall solution and how it has provided a unique way to enhance BAM and HSE's proactive approach to health and safety.

OVERALL SOLUTION, EXPLAINED

BAM with HSE formed a collaboration under HSE's Discovering Safety Programme. Its goals were simple: seek out and apply the best-in-class technologies in AI-driven predictive data analytics to uncover actionable insights to improve health and safety on construction sites across the UK.

This posed a number of challenges. Was there a sufficient amount of health and safety data available? Could it be structured in a way that wasn't overly time-consuming or expensive to analyse? Could business users with no background in data science be equipped with software tools discover these insights on their own? Were those insights unique unrecognizable patterns found in the data? Could these patterns improve the way we targeted health and safety inspections and training?

To approach these challenges, BAM and the DSP agreed on the need for a novel way of working with data. This new approach is illustrated below.



bam nuttall

ABOUT:
Part of the Royal BAM Group of civil engineers, BAM Nuttal is one of the UK's largest civil engineering contractors, employing over 3,000 people as well as subcontractors and joint venture staff.

Established for more than 150 years, BAM Nuttall delivers exceptional engineering, construction and investment services across private and public sector projects, in the UK and abroad. BAM Nuttall's projects range from road, rail and airports, to marine, tunnelling, and energy schemes. BAM Nuttall strives to create the best working environment for their team by constantly improving safety and working conditions.

The process above started with OST's RDFox technology integrating a variety of BAM data sets which did not have explicitly defined relationships (i.e. without necessarily being connected with a matching key or ID). These connections were established by RDFox, through health and safety expert-driven rules. In doing so, BAM benefited from a newly formed knowledge graph containing Project and Health and Safety specific data. Importantly, the knowledge graph did not require the same level of time or expense in setting up as a data lake or other traditional database repository would.

With this newly formed data set in hand, business users uploaded it into Mind Foundry's Machine Learning Platform so that they could build powerful predictive models and explore hidden insights in a friendly, no-code environment . Whilst it was important to BAM and the DSP to put AI and Machine Learning at the core of this new way of working, these advanced capabilities needed to be made possible by amplifying the know-how of health and safety experts.

Mind Foundry achieves this through a transparent and scientifically reliable platform that puts high-quality data science in the hands of the health and safety user. From data curation and enrichment, through model setup and optimisation, and finally on to rigorous and informative evaluation, Mind Foundry transparently incorporated expert knowledge at all stages of model development and refinement. Indeed, this led health and safety experts to retrieve meaningful insights in a number of areas, including

- Predicting project inspection performance
- Identifying patterns in accident occurrence records
- Spotting project attributes that strongly correlate with high-risk profiles.

Steve Naylor
Technical Lead,
Discovering Safety Programme

We're keen to explore how we might extend the work, including how other additional project datasets might be brought together and used, including more unstructured datasets, and project quality and scheduling data, to enhance how we're able to use the technologies in health and safety contexts.

Going forward, tighter integration between the different components would allow an increase in the frequency and volume of data for deeper analysis. It also would serve to automate other possible downstream activities, such as triggering alerts and notifications to Inspectors, real-time updating of health and safety dashboards, in addition to other process automations in support of health and safety training. The next phase will want to weigh up the benefits versus business needs of tighter integrations between technologies and workflow steps.

BAM and HSE's DSP have found a new way of working without adding any additional overhead, accessing gained advanced capabilities to uncover new, non-intuitive insights into health and safety influencers and risk factors connected to incidents and near misses on site.

Steve Naylor, Associate Director and Technical Lead on HSE's Discovering Safety Programme, comments: "It is commonplace for routine health and safety datasets to be stored and used disparately, to be held in multiple different formats and to lack common standards for easy, collective use; the RDFox and Mind Foundry platforms enabled us to link a variety of useful health and safety datasets together as part of this work and demonstrated that it was possible to use the techniques to identify certain associations between health and safety indicators in the dataset; now that proof of concept has been demonstrated, we're keen to explore how we might extend the work, including how other additional project datasets might be brought together and used, including more unstructured datasets, and project quality and scheduling data, to enhance how we're able to use the technologies in health and safety contexts."

Phil Cullen
Health and Safety Director
BAM Nuttall Ltd.,

"Combining data sets without direct relationships is possible and has the potential to generate useful insights and help to identify previously unrecognised relationships."

Health and Safety Director at BAM Nuttall Ltd., Phil Cullen, adds: "While BAM has been making use of data from multiple sources such as observations, accidents and inspection findings, as well as a number of leading indicators, these are often disparate sets of data that limit direct comparison and can be cumbersome to identify correlations. Although this pilot made use of a sample of data that may not give an accurate representation of the real trends and correlations within the business as a whole, it did show that combining data sets without direct relationships is possible and has the potential to generate useful insights and help to identify previously unrecognised relationships. With the potential to link other data sources such as quality and environmental data and with data not limited to the sample set of the pilot then analysis in this manner could lead to improved data gathering generally and the generation of insights and relationships not previously considered."

Let's examine in more detail OST's RDFox and Mind Foundry's technology, the challenges overcome, and the opportunities they have paved for the Discovering Safety Programme.

Oxford Semantic Technologies

ABOUT:

Experts in the field of knowledge graphs, Oxford Semantic Technologies (OST) was spun out of the University of Oxford in 2017.

OST's product, RDFox, is a highly scalable in-memory RDF triple store and semantic reasoning engine, the first market-ready system designed from the ground up to incorporate reasoning natively.

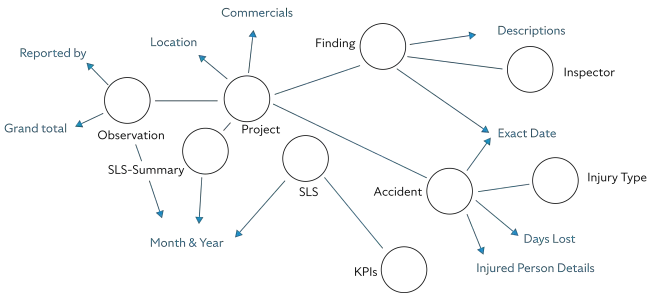
RDFox gives its users the ability to integrate data and domain expertise to then answer the questions that matter to your business. RDFox represents a paradigm shift in system design, breaking the traditional barriers between logic and data, and creating the next generation of intelligent databases.

ADVANTAGES OF KNOWLEDGE GRAPHS

OST's (Oxford Semantic Technologies) proprietary knowledge graph technology, RDFox, allowed BAM and the DSP to overcome the first fundamental step to the Discovering Safety Programme—assembling multiple independent data sets with little or no obvious relationship to link them together. Automating the necessary rules that govern relationships, filling in gaps in data such as event-driven time stamps, enabled BAM and the DSP to ask any number of probing questions.

To this end, different types of unrelated data sets needed to be intelligently ingested, including project data, safety inspection data, company safety score, major and minor accident data, observations sourced from the health and safety mobile app, and key performance indicators.

In doing so, graph databases are an ideal way to ingest such varied sources of information because they are free from the need for matching IDs that often constrain the formation of a structured table. Instead, data points are simply expressed as nodes which can be connected with edges, if they have some form of relationship.

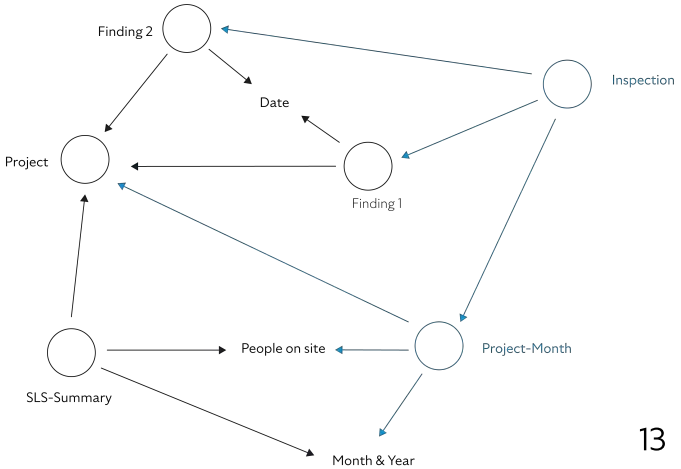


This process enriched the data set for machine learning by making more data available.

Graph databases are therefore more expressive and flexible than relational databases since they don't need to comply with a rigid, predefined structure. This is in part achieved because knowledge graphs can reason, meaning that rules can be applied to the data to identify or establish relationships between the nodes. For example, rules can help establish if data points expressed in different formats are actually the same.

Once the data is ingested, the knowledge graph starts to form a complex yet meaningful network of data. It is here where the expertise of BAM's health and safety leaders helped refine such rules and relationships, and surface implied knowledge that did not naturally exist in the raw data. For instance, the inconsistencies and missing data across the newly formed knowledge graph provided a challenge. Time is an important feature, but was often recorded differently or not at all depending on the original source of data. Another challenge was that the statistics for Projects were recorded on a monthly basis in some spreadsheets and on an inspection-by-inspection case in others. In a traditional relational database, this would be a costly problem to overcome. With RDFox, however, it was possible to automate a simple rule to work out the 'project-month' (the month when a given project was running),correct for inconsistencies, and fill in the blanks where a data point was missing. This process enriched the data set for machine learning by making more data available.

RDFox's automated approach to forming these relationships was shown to be both approachable and quick. The ingestion and mapping of the BAM data sets took only a few hours to code and a couple of seconds to process.



Integrating project health and safety, quality, and scheduling data empowers the exploration of associations between inadequate health and safety practice and the extent to which projects are delivered to time and quality, potentially providing early warning signs of when injuries to workers might be more likely on a project.

Once configured, querying the raw graph was instantaneous and expanding the database, such as adding additional data sets and correcting for the 'project-month' inconsistency, was also efficient. The rules that filled in various data gaps and generated useful new nodes (e.g. Inspections, Project-Months) were easy and took about a day.

Having few technical barriers during the initial ingesting and configuring of the knowledge graph, and the ability to make changes to the graph quickly and easily, gave BAM and HSE's DSP the confidence to ask more probing questions and seek out additional data sets to augment the initial findings. Whilst, the initial version of the knowledge graph proved sufficient for the pilot, it will be important to continue to train BAM's technical teams to maintain and expand the knowledge graph, as well as invest in efforts to be able to automatically integrate data as it is generated as part of BAM operations.

Combining RDFox's flexible knowledge graph with Mind Foundry's Machine Learning Platform offered health and safety experts a platform with unlimited possibilities to the questions which could be asked.

For example, integrating project health and safety, quality, and scheduling data empowers the exploration of associations between inadequate health and safety practice and the extent to which projects are delivered to time and quality, potentially providing early warning signs of when injuries to workers might be more likely on a project.

PREDICTIVE POWER OF MACHINE LEARNING

Mind Foundry's machine learning platform became the front-end user interface for the Discovering Safety initiative solution. In doing so, it enabled BAM and HSE with new advanced data analytic capabilities without the need for specialist data science experts. Instead, Mind Foundry put predictive analytics directly in the hands of health and safety experts and augmented their abilities to uncover hidden patterns of incidents and risks.

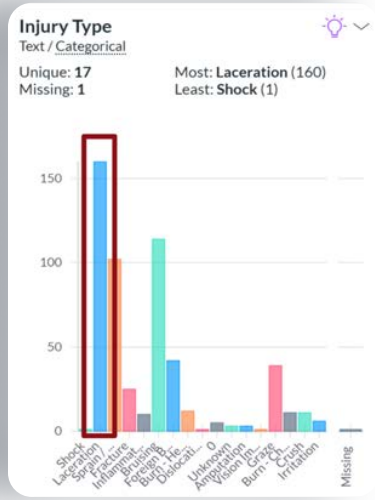
One example of this was the ability to use the long description fields of an Inspection to calculate a score based on morale and risk sentiment through in-app sentiment analysis features. Health and safety experts could then reliably select the correct machine learning model and, with in-app guidance, report on predicted results with opportunities to drill down and have the science explained. It did so without the health and safety expert needing any prior machine learning or data science experience.

mind FOUNDRY

Empowers our users to quickly find patterns in their own data.

These patterns may then translate into actionable insights to refine the corporate strategy.

Mind Foundry enabled H&S users to ask questions that would have otherwise been difficult to explore due to the size and nature of data sets.



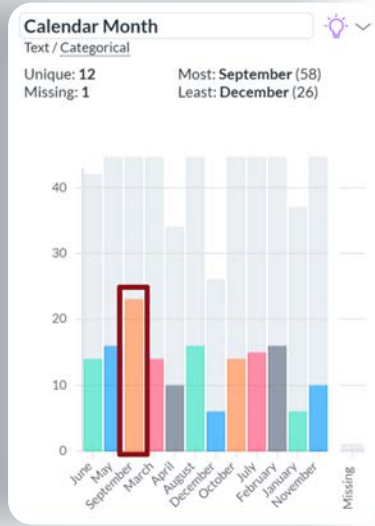


Figure X: Lacerations occur most frequently in late summer/early autumn

Indeed, Mind Foundry enabled H&S users to ask questions that would have otherwise been difficult to explore due to the size and nature of data sets. This included whether project characteristics, such as the growing number of people on-site, affected the safety over time, or if these features were linked with more observed health and safety risks. Another question was whether a project undertaken as a joint venture with other construction companies, rather than for BAM Nuttall on its own, was associated with different inspection findings, or whether this also had any influence over health and safety risks.

A primary challenge was to spot accident patterns using a variety of data sources. Of the 17 types of injuries analysed, Mind Foundry found "time of year," "level of experience," and "age" to be the leading influencers to most injuries. Fractures were three times more likely to happen with employees that had been in the role between 6 and 12 months. And lacerations were the most frequent accident in the summer and early autumn months, occurring nearly twice as frequently as during winter months.

Mind Foundry helped surface these otherwise hidden data patterns quickly and easily, providing an opportunity for the health and safety expert to confidently take action.

Once uncovered, these patterns provide an opportunity for targeted interventions to be put in place. Crucially, Mind Foundry helped surface these otherwise hidden data patterns quickly and easily, providing an opportunity for the health and safety expert to confidently take action, for example, by adjusting current training strategies or the timings of safety awareness campaigns through the year. Indeed, a training that normally happens before Christmas and involves inputs on preventing lacerations was recommended to be moved to the beginning of summer to reduce those incidents when they most occur.

Beyond visualising quantitative figures such as the number of lacerations, Mind Foundry was also able to score descriptive or textual data using sentiment analysis. This could enable any number of uses, but one meaningful comparison was trending the sentiment, which can be interpreted as morale over time. As the mean sentiment score trends down, as was the case for Accident Books in Project "39" and "42", health and safety experts could plan more targeted interventions.

Our sentiment analysis would flag Projects 39 and 42 for “Accident book”, which could bring them to the attention of BAM management.

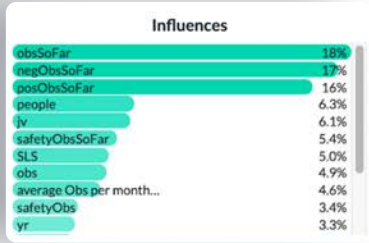
An appropriate training plan for these projects and finding categories could then be tailored accordingly.

Figure X: Trending sentiment score over time to assess project risk and inspection outcomes from descriptive data

These analytical capabilities of Mind Foundry go beyond only looking at historical trends. It can also be used to predict future outcomes, for example, predicting the outcome of an inspection before the inspection has happened.

QUESTION

Could we predict the outcome of an inspection before the inspection has happened?



Yes.
Mind Foundry can predict the outcome of an inspection with healthy accuracy.

With a target score of 1, this model was sufficiently accurate to give a clear indication of the predicted outcome of an inspection. Even more importantly, the model provided an intuition for what factors drive the accuracy of that prediction. In this case, the observations a project had received at that point in time, as well as the number of people working on the project, were leading indicators of the outcome of the inspection. On a practical basis, this means that we can now build an early warning system, powered by a machine learning model, which would flag potential issues to the project management team before the involvement of an inspector.

Taken together, uncovering hard to find patterns, generating new patterns, along with opportunities to predict future outcomes of inspections leads us to new possibilities in how we plan, inspect and even report on health and safety concerns.

Although these initial findings are interesting and valuable, the true potential and future benefits of AI and machine learning need to move beyond the limits placed around the initial pilot phase of this project. A next phase should encourage more business user engagement, access to more frequent data, and a more tightly integrated end-to-end process to surface greater opportunities to exploit the benefits of AI and machine learning.

FUTURE OF HEALTH & SAFETY TODAY

Going forward, the team will progress by bringing predictive models to a production stage and converting the discovered insights into tangible actions. The targeting of training sessions and re-design of how inspections are carried out, especially in joint venture projects, are of particular importance for the Discovery Safety Programme and the safety team at BAM Nuttall.

In a similar manner, the current analyses could be extended to larger and more varied types of data. For example, further semantic analyses could also be carried out on other sources of raw text data, such as observation data. This could help to curate morale trends in specific sites, which could be targeted for improvement as needed.

Dr. Helen Balmforth
HSE Head of Data Analytics,
Discovering Safety Programme Director

We're keen to explore as part of follow-on work the value of bringing in other project data, for example, project scheduling and quality data. We're also keen to shift the focus to an individual, current joint venture project and explore how we might bring together data from multiple contractors working on the project to support health and safety decision-making across the entire project.

Additionally, collecting near-miss data could be key to building more powerful predictive models to prevent accidents. By comparing and contrasting near-misses with effective accidents, hidden-in-plain-sight accident risk sources could be more timely isolated and acted upon. Furthermore, by combining this with employee and contextual data (such as time of the day), more specific training can be tailored to address the right audiences at the right time.

Dr Helen Balmforth, HSE Head of Data Analytics and Discovering Safety Programme Director, suggests a way forward: "We're keen to explore as part of follow-on work the value of bringing in other project data, for example, project scheduling and quality data. We're also keen to shift the focus to an individual, current joint venture project and explore how we might bring together data from multiple contractors working on the project to support health and safety decision-making across the entire project."

DISCOVERING SAFETY

bam nuttall

mind FOUNDRY

Oxford Semantic Technologies