Digital H&S Risk Library Use Case

Author(s): Dr Carlos A. Osorio-Sandoval; Dr William H. Collinge; Dr Clara Cheung; Dr Mojgan Hadi Mosleh; Dr Patrick Manu; Dr Andre Freitas; Dr Zili Zhou

BIM Risk Library Final Report (Phase 2)

Report approved by:	(Include qualifications)
Report authorised for issue to {client name – add name here} by:	(Include qualifications)
Date of Issue:	30/06/2021
Lead Author:	Dr Carlos Osorio-Sandoval
Contributing Author(s):	Dr William Collinge; Dr Clara Cheung; Dr Mojgan Hadi Mosleh; Dr Patrick Manu; Dr Andre Freitas; Dr Zili Zhou
Customer:	Lloyds Register Foundation
Technical Reviewer(s):	(Include qualifications)
Editorial Reviewer:	(Include qualifications)
Project number:	
Distribution list:	(Optional)

Disclaimer and acknowledgements

This report and the work it describes was undertaken by The University of Manchester as part of Discovering Safety. Its contents, including opinions and/ or conclusions expressed, or recommendations made, do not necessarily reflect the policy views of the HSE or Lloyd's Register Foundation.

Executive Summary

This report presents the work undertaken during the Phase 2 –from June 2020 to June 2021 – of the BIM Risk Library Use Case of the Discovering Safety programme. Continuing with the work reported in the *BIM Risk Library Final Report (Phase 1)*, the new version of SafetiBase, which was developed by 3D Repo based on collaboration with the Discovering Safety team, moved from a prototype to a piloting phase. The Risk Library, implemented within this platform, was deployed across multiple projects from the industry partners that joined the Community of Practice established by the Use Case team.

These industry partners piloted the new version of SafetiBase in a range of construction projects of different types, including residential, commercial, industrial, and infrastructure projects. New risk scenarios and treatments were retrieved from the pilot projects to improve the Risk Library knowledgebase.

Concurrently, risk scenario were also retrieved from the UK Health and Safety Executive's (HSE) Archive data, for which a Natural Language Inference (NLI) approach was proposed to automatically classify free text into target categories.

This Phase of the Use Case evaluated the implementation of the Risk Library within a 3D BIM environment, as well as the data structure that underpins it, through qualitative research methods. Furthermore, the Use Case also conducted an investigation into the applicability of these concepts in a 4D BIM environment, focusing on how 4D can be exploited to enhance the safety risk identification process. The main findings of this Use Case are outlined below:

1. Upon completion of the pilot project phase, the Risk Library contain 401 treatments for 31 different risk scenarios related to 11 risk categories.

2. The NLI experimental work demonstrated that the approach is a viable way for automatically classifying accident records into an accident kind category, which is related to the Risk Library schema.

3. Evaluation of the new version of SafetiBase using expert survey and individual interviews showed that the tool has practical utility in industry. Suggestions to further enhance the practical utility of the tool were also offered by the participants in the evaluation activities.

4. Overall, the Phase 2 study shows the benefits that the Risk Library can bring to the implementation of design for safety in the construction industry. Design professionals should be encouraged to engage with it to promote its wide use in industry and its continuous development.

Contents

Abbreviations

BIM	Building Information Modelling
CSV	Comma-separated values
DfS	Design for Safety
DSP	Discovering Safety Programme
ERIC	Eliminate, Reduce, Inform, Control
NLI	Natural Language Inference
HSE	Health and Safety Executive
H&S	Health and Safety
PtD	Prevention through Design
PD	Principal Designer
RIDDOR	Reporting of Injuries, Diseases and Dangerous Occurrences Regulations

1 Introduction

1.1 Background

The construction industry continues to be one of the most hazardous industries worldwide. It is believed that most accidents in construction can be prevented if addressed during the design phase of work (Kasirossafar et al., 2012). Under this premise, the concept of *Design for Safety* (DfS), also referred to as *Prevention through Design* (PtD), has gained popularity and traction in the past decades. DfS emphasises the importance of designers anticipating potential risks as early as possible, and proposing means to eliminate, reduce, inform, or control (ERIC) the risks. Effective implementation of DfS requires appropriate technologies and tools that enable designers to engage with construction designs in order to apply their knowledge and skills in safety management (Hossain et al., 2018; Yuan et al., 2019).

For a safer built environment sector, The Lloyd's Register Foundation and the Health and Safety Executive, supported by the Thomas Ashton Institute and The University of Manchester are jointly undertaking the Discovering Safety programme. This report provides a review of work undertaken in Phase 2 of the Digital Health and Safety Risk Library Use Case (June 2020 – June 2021); this work building upon work completed in Phase 1 (January 2019 – June 2020).

Central to the Digital Health and Safety Risk Library Use Case is the Risk Library knowledge base that maps construction risk scenarios to treatment prompts to assist designers implementing DfS. As noted in the Phase 1 report, the Risk Library is formulated on a risk scenario/treatment ontology: a *risk scenario* being characterised by six data points: (1) construction scope, a concept based on CIRIA C755 CDM 2015 (Ove Arup and Partners and Gilbertson, 2015) that describes the type of construction work; (2) risk category, a concept based on PAS 1192-6:2018 (British Standards Institution, 2018) that identifies the type of risk that could occur; (3) building element, a concept that enables classifying the scenario by related building design discipline; (4) location relative to the risk; (5) activity, a concept sub-classified at a high level that determines the stage in the life cycle of the asset during the possible eventuation of the risk; and (6) risk factor, a concept that identifies the reason behind the risk eventuation. Holistically, the risk scenario then requires an appropriate treatment; the ontology is visualised in Figure 1below.
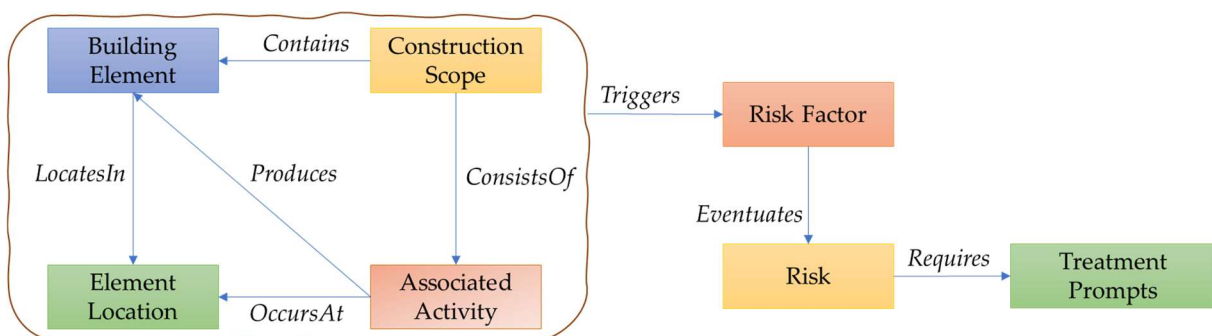

Figure 1 Risk scenario ontology

Each scenario is mapped to one or more *treatment* prompts. A risk treatment is defined (ISO Guide 73, 2009) as a process that modifies the risk. It can involve: (1) avoiding the risk by deciding not to start or continue with the activity that gives rise to the risk; (2) taking or

increasing the risk in order to pursue an opportunity; (3) removing the risk source; (4) changing the likelihood; (5) changing the consequences; (6) sharing the risk with another party or parties; and (7) retaining the risk by informed decision. The word *treatment* is used in preference to the widely used *mitigation* term because it includes the option of eliminating the risk.

As noted in the Phase 1 report, The Digital Health and Safety Risk Library is grounded in industry standard PAS 1192:6 – Specification for collaborative sharing and use of health and safety information using building information modelling (BIM) and builds upon CDM 2015 guidelines advocating for the prevention of accidents, the sharing of risk information and better coordination between construction project designers, engineers, contractors and clients. Specifically, the structuring of risk data using an ontology/framework of analysis that captures tacit and explicit expert knowledge and then mobilises that data using an interactive, interface for designers using their models was viewed as a distinct achievement in Phase 1.

Intrinsic to this was 3D Repo, an existing commercial cloud-based BIM application that developed a new version of SafetiBase based on collaboration with the Risk Library team. The new version of 3D Repo's SafetiBase hosts the Risk Library in the form of a comma-separated values (CSV) file. An important new feature enables users to leverage the Risk Library's knowledge by suggesting existing treatments upon risk scenario identification and characterisation based on the Risk Library's ontology.

1.2 Project aim
Phase 2 of the Risk Library Use Case had several key aims. These were: (1) to expand the Risk Library developed during Phase 1 by including more risk scenarios and treatments; (2) to evaluate its implementation within a BIM environment through a series of Pilot projects with industry; (3) to explore implementation of the risk scenario/treatment ontology and Risk Library within a 4D BIM environment; (4) to explore methodologies for capturing risk scenarios from the HSE archive using automated data processing techniques.

To achieve the above aims the following objectives were established:

1. Expand the Risk Library from HSE Archive data

2. Engage with industry partners to pilot the implementation of the Risk Library within the new version of 3D Repo's SafetiBase

3. Prepare a user manual for pilot project participants

4. Expand the Risk Library from pilot projects' data

5. Hold a series of 4D workshops with industry to explore implementation of the Risk Library within a 4D BIM environment

6. Evaluate the new version of 3D Repo's SafetiBase from practitioners' perspective

1.3 Project outline
Figure 2 depicts the various tasks and outputs of the work undertaken during Phase 2 of the Risk Library Use Case. The work was organised across five concurrent work-packages (WP) with interlinked work streams that lead to the completion of the research aim and objectives outlined above.
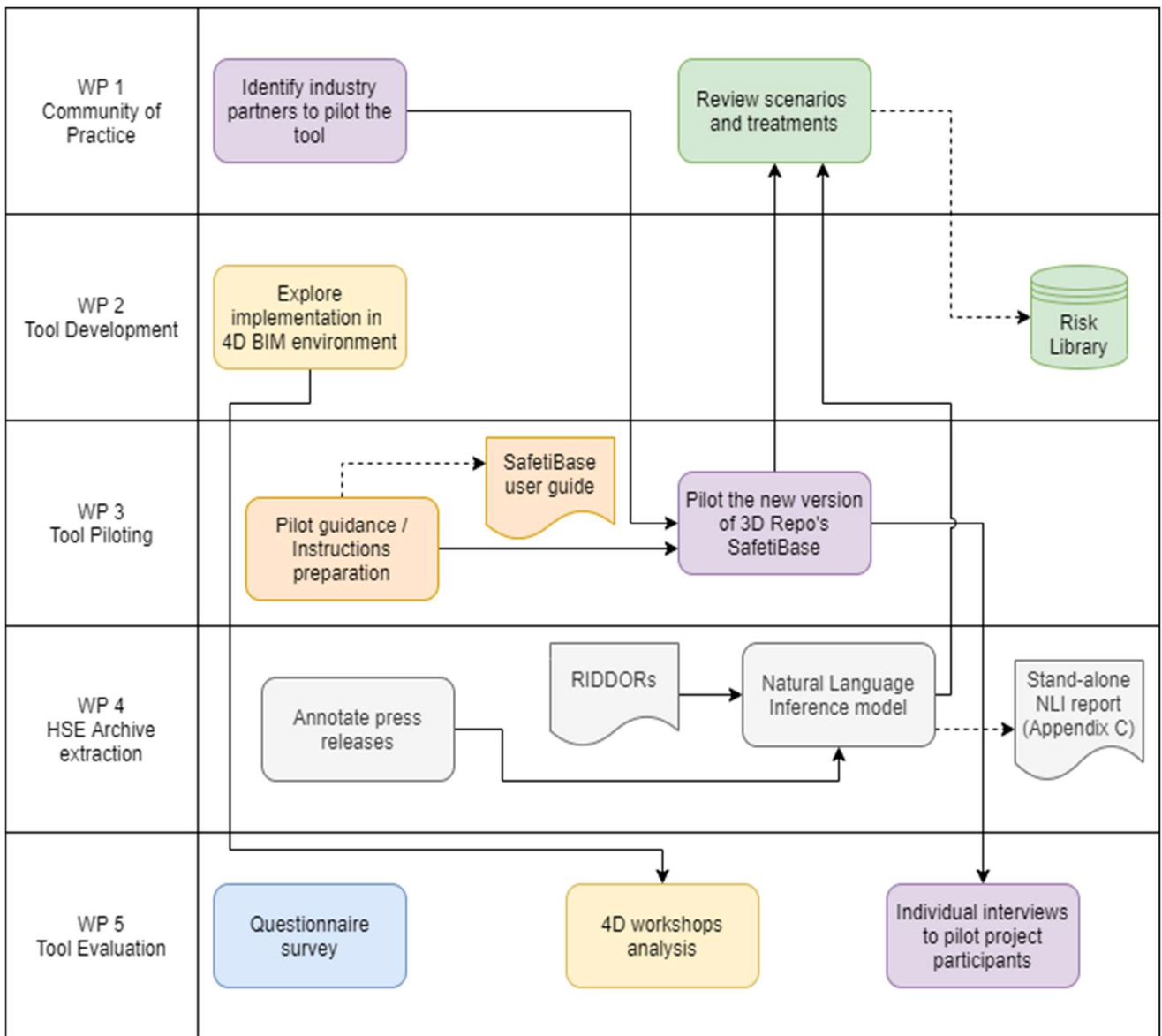
Figure 2 Project outline

1.4 Research team

Table 1 shows the researchers involved in Phase 2 of the Risk Library Use Case. The team consisted of 1 Principal Investigator (PI), 4 Co-Investigators (Co-I`s) and 2 Research Associates.

Table 1 Research team and roles

Name	Role	Email
Dr William Collinge	Principal Investigator	william.collinge@manchester.ac.uk
Dr Carlos Osorio-Sandoval	Research Associate	carlos.osoriosandoval@manchester.ac.uk
Dr Clara Cheung	Co-Investigator	clara.cheung@manchester.ac.uk
Dr Mojgan Hadi Mosleh	Co-Investigator	mojgan.hadimosleh@manchester.ac.uk
Dr Patrick Manu	Co-Investigator	patrick.manu@manchester.ac.uk
Dr Andre Freitas	Co-Investigator	andre.freitas@manchester.ac.uk
Dr Zili Zhou	Research Associate	zili.zhou@manchester.ac.uk

2. Methods

2.1 Pilot testing the new version of 3D Repo's SafetiBase

The new version of 3D Repo's SafetiBase, as developed based on direct collaboration between 3D Repo and the Risk Library project team during Phase 1, was piloted by a number of industry partners to obtain feedback on its functionality, utility and overall value. Additionally, risk scenarios and treatments were obtained from the pilot projects and added to the Risk Library database following a review by the project team.

Six construction projects at an early design stage of the construction project lifecycle participated in the pilot testing phase. The participating pilot projects are outlined in Table 2.

Table 2 Participating pilot projects

Project name	Company	Project description	Pilot duration (to date)
QA Archive Facility Expansion	AstraZeneca	Pharmaceutical cold storage	5 months
Western Yards (Building 3)	Multiplex	18-storey office block	6 months
SeedPod	Atkins – Faithful & Gould	Specialist facilities for business	3 months
TP Paddington	Atkins – Faithful & Gould	2-storey shop and 16-storey student accommodation	3 months
Un-named project	Atkins – Faithful & Gould	Groceries retail store	3 months
Cargo tunnel	Heathrow	Road tunnel underneath airway	3 months

A user guide to the new version of 3D Repo's SafetiBase (Appendix A) was prepared by the research team and distributed to the pilot project participants following a briefing meeting in which details on the workflow of the pilot project were also provided.

Pilot project licences

In order to facilitate pilot project engagement and active use of the Risk Suggestion Tool, pilot project participants were offered a University of Manchester-funded 3D Repo license for up to ten users covering up to eight months, depending on the start date of the pilot. Pilot project participants could extend the length or increase the number of users of their license liaising directly with 3D Repo.

Pilot project workflow

A prototype version of the Risk Library was loaded into each pilot projects' *Teamspace* in 3D Repo as a CSV file before the pilot started. This prototype contained the risks and treatments of work completed during Phase 1 of the research project: primarily focusing upon falling from height risks. While piloting 3D Repo's SafetiBase, the participants would be able to view the treatments contained in the database if they identified a risk scenario that was already in the Risk Library. Pilot project participants inputted their own treatments when appropriate treatments were not available for an identified risk scenario.

The research team retrieved, anonymised and reviewed risk scenarios and suggested treatments input by participants periodically following the Figure 3 workflow plan. Subsequently, the Risk Library was updated and loaded back into the projects' 3D Repo *Teamspace* so that the new scenarios and treatments could be leveraged by participants. The workflow of the pilot projects is outlined in Figure 3.
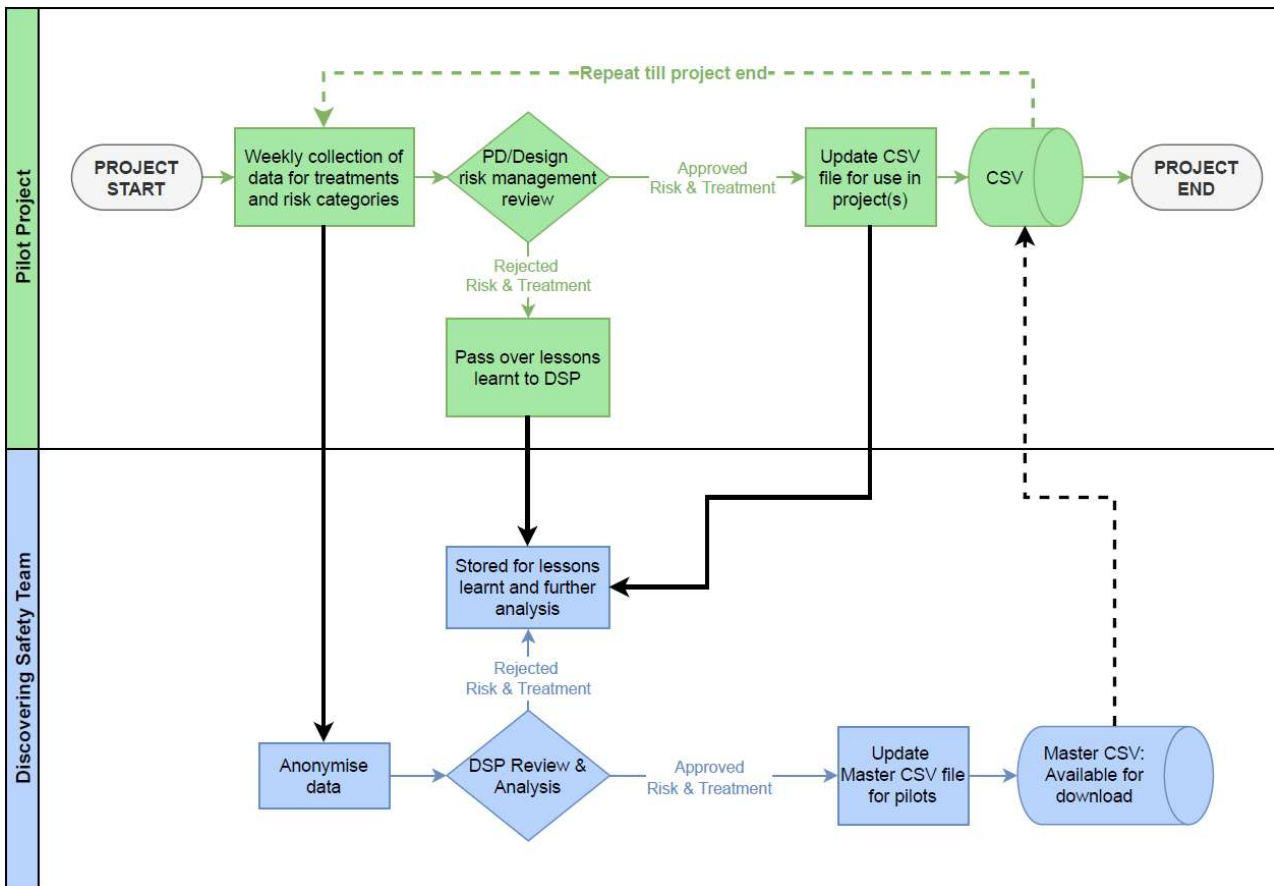
Figure 3 Pilot project workflow

Pilot collaborative agreements

Prior to piloting the new version of 3D Repo's SafetiBase and the Risk suggestion feature, each pilot project company signed a tri-partite Collaborative agreement with University of Manchester and the HSE. An unsigned copy of this Agreement is attached at Appendix B. In essence, this agreement provided each pilot project with assurance that the data they shared with the HSE and the University of Manchester would be held securely, anonymised and used solely for the further population of the Risk library database of risk scenarios and treatments.

Pilot project insights

Each one of the industry partners participating in the piloting phase incorporated the new version of 3D Repo's SafetiBase into their day-to-day activities and process to manage safety in a different way, based on their own preferences and on the stage of their projects.

One of the pilot project teams used SafetiBase as a team following a systematic approach to identify risk scenarios and propose treatments collaboratively. Subsequently, they used the tool to communicate the identified risks with other project stakeholders in their tendering process. Notably, they used SafetiBase after they had already populated a risk register following their usual approach. The use of SafetiBase lead to identifying new risks and a redesign prior to the tendering process.

On the other hand, other pilot project teams used SafetiBase in parallel to their usual safety management approach based on the argument of disrupting their day-to-day activities and

usual processes as little as possible. Only one pilot project team incorporated SafetiBase as part of their normal safety management process.

Section 3.4 collates further insights drawn from interviews with pilot project participants, as described in Section 2.6.

2.2 Expanding the Risk Library from HSE Archive data

A key aim of the Use Case was to investigate how risk scenarios could be identified, extracted and mobilised in the Risk Library from the HSE archive of data relating to accidents in construction. Appendix E contains a stand-alone report on the text inference approach adopted by the University of Manchester team to characterise risk scenarios from HSE data.

Two different approaches were adopted for undertaking this task: (1) Manual approach, and (2) Natural Language Inference (NLI) approach.

Manual approach

Press releases related to construction accidents were manually annotated using the below protocol:

- Step 1. Identify the event reported in the press release.
- Step 2. Choose the risk category that best describes the reported event. The chosen category should respond to the question 'what happened to the subject of the press release?'
- Step 3. Choose the risk factor category that is most strongly linked to the previously selected risk. The chosen category should respond to the question 'what circumstance eventuated the selected risk?'
- Step 4. Choose the location category that best describes the location of the subject of the press release at the moment of the risk eventuation. The chosen category should respond to the question 'where was the subject of the press release when the risk eventuated?' If the question is not clearly responded from the information in the text, and the identified risk could have been eventuated at different locations, leave blank.
- Step 5. Choose the construction scope category that best responds to the question 'what type of construction work was being carried out at the time of the reported event?'
- Step 6. Choose the activity category that best describes the stage of the life cycle of the asset at the time of the reported event.
- Step 7. Choose the building element category that best describes the building element involved in the reported event. Consider the element that would be most likely present in the asset's design model. If more than one element is identified, follow this priority: element that collapsed or failed (exclude excavation walls collapsing; include temporary structures like scaffolding), element being installed or removed, element being transported or handled.
- Step 8. If applicable, choose more risk categories that could have happened but were not eventuated. Only choose risks related to the location selected in Step 4.
- Step 9. If applicable, choose other risk factors that could have eventuated risks identified in Step 8.

Newly identified risk scenarios were mapped to treatments retrieved from existing construction guidance for designers in the UK (Ove Arup and Partners and Gilbertson, 2015) or through focus group discussions with the research team. Subsequently, the risk scenarios and their treatments were added to the Risk Library.

Natural Language Inference approach

The manual extraction process to find risk scenarios described above is complex and expensive due to the amount of unstructured text derived from multiple sources, including

press releases and RIDDOR reports from the HSE archive. State-of-the-art techniques from the Natural Language Inference (NLI) domain (MacCartney and Manning, 2009) can provide good opportunities to push forward the automation of the risk scenario extraction process. Both RIDDOR reports and press releases text contain natural language text. Risk scenarios can be considered as target semantic information needed to be extracted from such texts. The NLI model can automatically learn the semantic information in the text through deep neural network and classify each text into a list of categories.

Two reference corpora, desensitised RIDDOR reports and HSE press releases reporting construction accidents, were used as input to explore the automatic categorisation of risk scenarios. Three transformer-based models were proposed to deliver this objective: (1) training a model using a common set of risk scenario categories present in RIDDOR reports, aiming to transfer this knowledge to a general risk scenario classifier; (2) leveraging the embedded textual knowledge from RIDDOR reports to support the target risk scenarios classification on press releases by fine-tuning a pre-trained model in the press releases; and (3) evaluating zero-shot learning on the press release corpus by using the previously built classifier.

Appendix E contains a more detailed stand-alone report on the text inference approach adopted to automatically classify free text into target categories.

The NLI-based classifier was applied to a random sample of 100 RIDDORs targeting the categories "kind group" and "main factor", which are closely related to the risk and main factor categories of the Risk Library ontology, respectively. The results were manually reviewed by an HSE expert on construction H&S.

2.3 Focus group discussions to review and analyse risk scenarios and treatments

The focus group discussions aimed to assure the quality of the knowledge appended to the Risk Library from the methods described in Sections 2.1 and 2.2. Two types of focus group activities were conducted to this aim: an industry workshop, and fortnightly review meetings with the Risk Library Use Case research team.

Industry workshop

Pilot project participants were invited to participate in an industry-focussed workshop to review three selected risk scenarios and complete their respective treatment matrices. The workshop was held on Friday 16th April 2021. Participation was voluntary. Participants were broken into two separate groups to review the selected scenarios and to provide treatment prompts for them using the treatment matrix format. The responses from both groups were curated and collated by the research team.

Fortnightly review meetings with the Risk Library use case team

Risk scenarios and treatments retrieved from pilot projects as detailed in Section 2.1 were reviewed by the Risk Library team in fortnightly 2-hour workshop sessions. Retrieved treatments were synthesised and organised within the corresponding treatment matrix. The treatment matrix was used in Phase 1 to develop the prototype of the Risk Library. Further treatment prompts were added to each risk scenario using the treatment matrix format based on the experience of the review panel and review of grey literature. An average of five risk

scenarios were reviewed during each session. The resulting risk scenarios and treatments were appended to the Risk Library CSV file.

2.4 Survey to evaluate the new version of 3D Repo's SafetiBase

An online questionnaire (Appendix C) was formulated and used to evaluate the new version of 3D Repo's SafetiBase. The questionnaire was applied to a sample of health and safety expert professionals from the AEC industry. Professionals with knowledge in the research topic and a minimum of five years of experience in a role relevant to the research subject were considered as suitably qualified and experienced experts (Hallowell and Gambatese, 2010; Poghosyan et al., 2020). The research team's relevant contacts that met the above criteria were directly invited to participate in the survey and were asked to refer others. Incomplete and invalid responses were excluded from the analysis (e.g. if the respondent didn't meet the selection criteria). Thirteen valid and complete responses from experts with an average of 34 years of experience were considered for analysis in the survey.

The survey was created and deployed in Qualtrics (https://www.qualtrics.com), an online software platform. The questionnaire started with general demographic questions to gain background information on the participants. Then, the participants were shown a video demonstration of the new version of 3D Repo's SafetiBase. Subsequently, the participants assessed the Risk Treatment Suggestion feature using a five-point Likert scale. The scale measured the level of agreement on statements about the feature's: usefulness, potential to leverage lessons learnt across multiple projects, ease of use, alignment with regulations and guidelines, potential to affect design decisions and management processes, potential to enable collaboration, and potential to improve health and safety in construction. Further opinion and comments from participants were also collected using open-ended questions.

2.5 Individual interviews

We conducted semi-structured individual interviews with a pilot project participant from each industry partner. We selected an individual from each pilot team who was directly using the new version of 3D Repo's SafetiBase in their pilot project. The interviews were video recorded with the participants' consent. Transcripts of the interviews were used to find recurring themes in regards to the participants' opinion on the data structure used to characterise risk scenarios used in the pilot projects. The interviews also aimed to understand the impact of using the new version of SafetiBase on the organisations' safety management processes and on the users' thoughts processes regarding safety risks. Appendix D shows the structured questions asked to the participants. Follow up questions were asked in addition to these questions to ensure optimal responses from participants where appropriate.

2.6 Workshop series to explore 4D capabilities

We adopted a workshop series research methodology to explore how 4D BIM can be utilised to support the traditional safety management process relying on 2D drawings, tacit knowledge and regulations (Choe and Leite, 2017). Workshops are considered an appropriate method to identify, articulate and explore fuzzy problems in research areas involving technology (Ørngreen and Levinsen, 2017). Participants of the workshop were selected by purposive sampling based on their knowledge and experience in the field of construction safety or 4D modelling to incorporate a wide range of different perspectives into the sessions. Researchers and participants contributed to the sessions in a mutual process controlled by the participants.

During the workshop sessions, participants performed a safety risk assessment identifying safety risks using a 4D model of a 16-storey residential building, developed to include temporary structures to ensure the availability of project-specific spatiotemporal information. The identified risks were captured within 3D Repo's SafetiBase and characterised according to the Risk Library ontology. Indicative 3D pins were placed in the model to visualise the points in the model where risks were identified. A treatment to mitigate each risk scenario was proposed. A start and end date were provided to the risk to indicate its duration in the context of the 4D model for visualisation purposes.

At the end of the session, workshop participants engaged in focus group discussions to compare the 4D BIM-supported safety risk assessment process to the traditional safety planning approach. Figure 4 outlines the basic structure of the workshop sessions. As previously noted, participants were in control of the workshop sessions, excluding the welcome and focus group discussion segments, in which the researchers adopted a moderator role. Thus, the process in between those segments did not necessarily follow a sequential order.
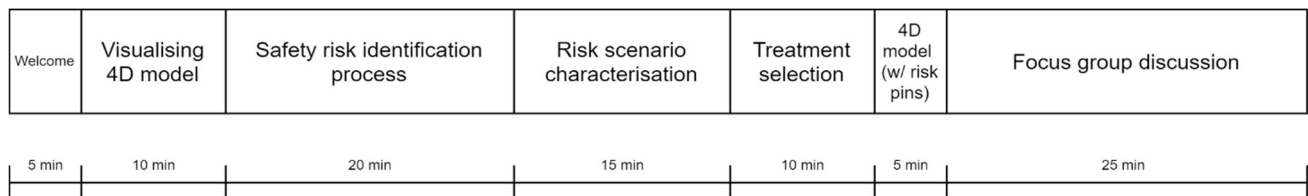
Welcome	Visualising 4D model	Safety risk identification process	Risk scenario characterisation	Treatment selection	4D model (w/ risk pins)	Focus group discussion
5 min	10 min	20 min	15 min	10 min	5 min	25 min

Figure 4 Workshop session basic structure

3. Results

This section of the report presents the results of the Use Case research work in several sections.

3.1 Risk Library expansion

Following focus group reviews as detailed in Section 2.3, the Risk Library (v 11/06/2021) contains 401 treatment prompts for 31 risk scenarios related to 11 different risk categories. The risk scenarios included in this version of the Risk Library are listed in Appendix F. The complete Risk Library is available to download as a CSV file from the Discovering Safety website (https://www.discoveringsafety.com/). The majority of the risk scenarios were sourced from pilot projects.

During the research team review meetings of the collected scenarios, the ontology had to be extended to reflect the categories input by the pilot participants. This suggests that, while the main categories of the schema are valuable, more work is needed to match the variety of risk scenarios that the industry presents.

3.2 Automatic classification of free text into Risk Library categories

Appendix E contains a stand-alone report on the automatic classification of free text into target categories related to risk scenarios. The NLI-based classifier performed with a mean accuracy of 71.3% for the kind group column on a sample of 51k RIDDORs. The kind group category is strongly related to the Risk Library's risk category.

The NLI-based classifier was applied to a random sample of 100 RIDDORs targeting the "kind group" and "main factor" categories. The accuracy of the model was calculated by comparing the predicted category to the original category reported in the RIDDOR. An HSE expert validated the results by manually reviewing the results of the model. It was observed that, in some instances, the expert agreed with the predicted category even if the prediction was not accurate compared to the reported category. Table 3 depicts the accuracy and expert agreement of the NLI-based classifier on the sample.

Table 3 NLI-based classifier results on a random sample of 100 RIDDORs

Results	Kind group	Main factor
Accuracy	70.0%	65.0%
Expert agreement	93.0%	83.0%

Table 4 shows the text of a RIDDOR in which the NLI-classifier did not predict the reported kind group or main factor categories accurately. However, the expert considered that the predicted categories were more accurate to describe the input text than the category originally reported on the RIDDOR.

This technique could be exploited to automatically identify risk scenarios from RIDDOR reports. A requirement to achieve this is to manually classify RIDDOR reports into the target categories to pre-train the model.

RIDDOR text	Category	Reported	Predicted	Accuracy	Agreement
The operative was cutting a roof batten without placing it on a secure surface the saw jumped and cut his thumb	Kind group	Lifting and handling injuries	Contact with moving machinery	0	1
	Main factor	Pushing, pulling	Loss of control of machinery, transport or equipment	0	1

3.3 Survey to evaluate the new version of 3D Repo's SafetiBase

The questionnaire survey described in Section 2.4 was open for five weeks and received 35 visits which resulted in 13 completed questionnaires. Twenty-two responses were excluded from the analysis because the respondents didn't meet the inclusion criteria, or because the responses were incomplete. Figures 5 to 7 depict the demographic information of the respondents.
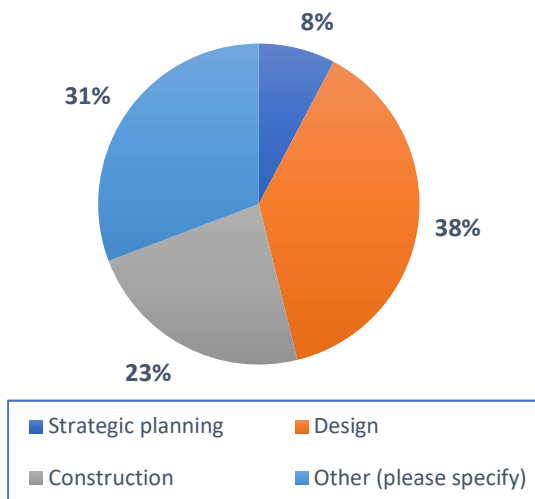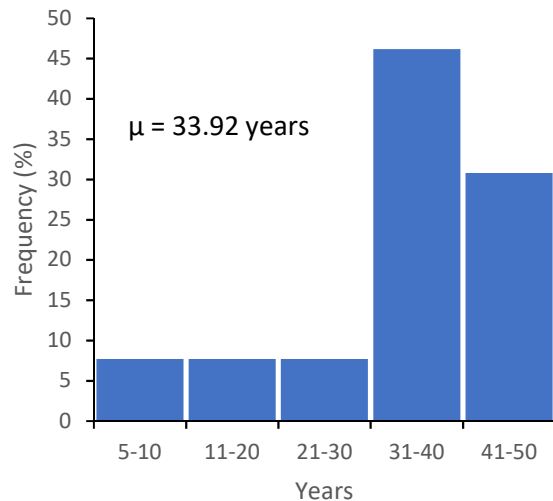
Demographic information

Figure 5 Area of work

$\mu = 33.92$ years

Figure 6 Years of experience in the construction industry
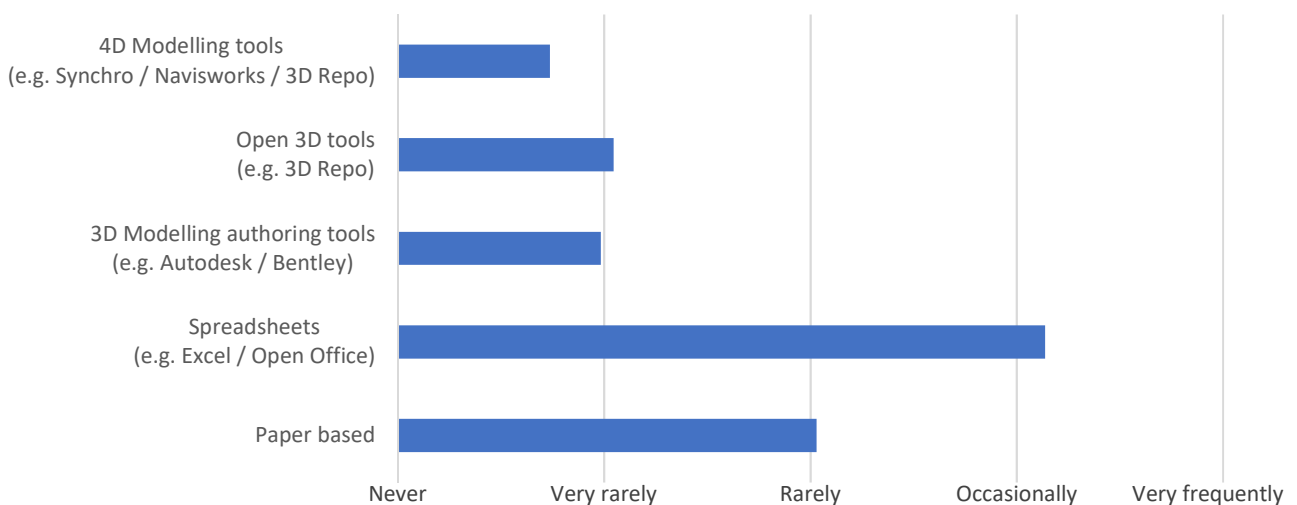

Figure 7 Frequency of use of tools to manage health and safety information

Feature assessment

Figure 8 shows the level of agreement of the survey respondents regarding the nine attributes of the Risk Treatment Suggestion feature assessed in the questionnaire. The feedback of the experts reveals that for all the evaluation criteria, at least 50% of the experts agree to some degree that the feature achieves the evaluation statement. Most notably, approximately 85% of the experts either somewhat agreed or strongly agreed that it can positively impact design decisions, supports selection of appropriate treatments to mitigate health and safety risks, and enables leveraging lessons learnt across previous projects.
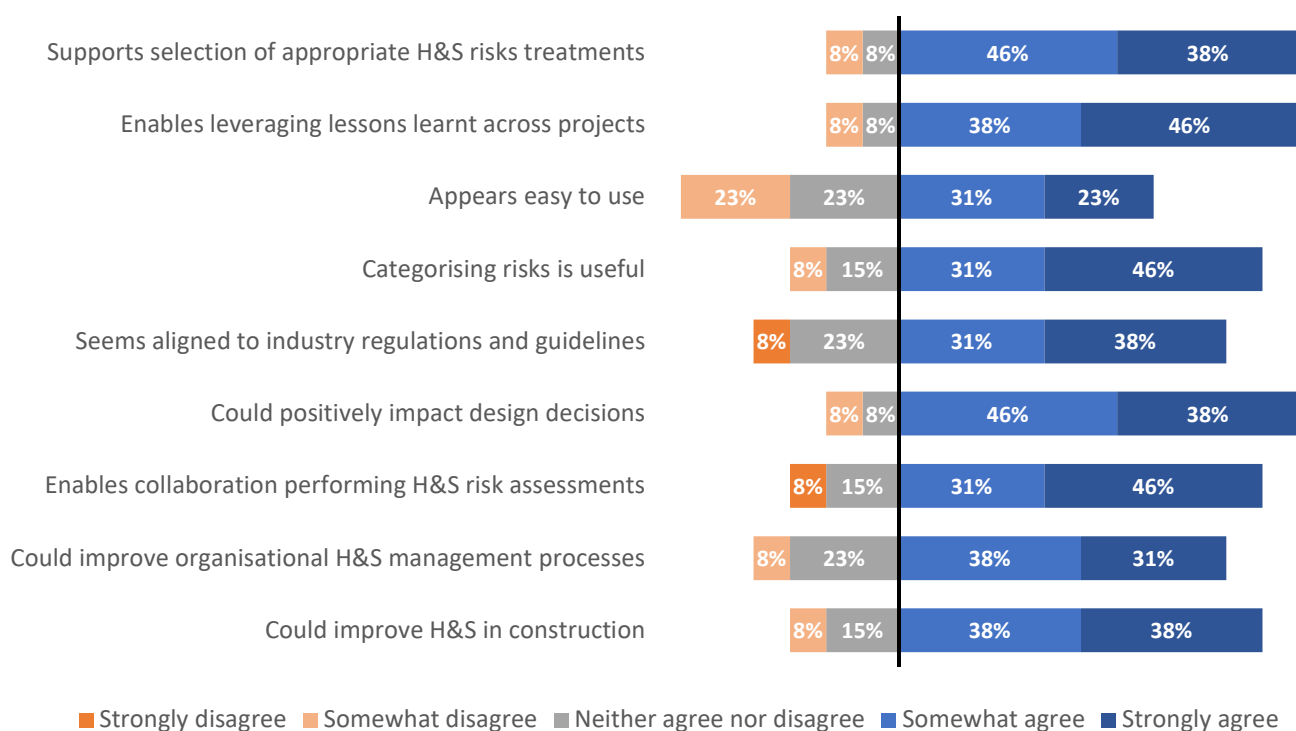

Figure 8 Results of the Risk Treatment Suggestion feature assessment

Challenges to adopting the new version of 3D Repo's SafetiBase

The responses to the open-ended question *"What are the potential challenges to adopt this tool?"* suggest that licensing cost and the required time to train personnel to use the system adequately are the main challenges to adopting the new version of 3D Repo's SafetiBase in new projects. The time required to update the risks input within the system as a project progresses was also highlighted as a potential barrier to overcome. Two respondents mentioned that the system requires having a 3D model of the project to be useful, which is often not available, especially at the early design stages.

3.4 Individual interviews

The findings from the interviews with the pilot project participants are presented below.

Current Safety Management approach

Before pilot testing the new version of 3D Repo's SafetiBase, the Safety Management approach that all the interviewees followed involved using spreadsheets to capture risks, their mitigation measures, and residual risks. Several interviewees highlighted that these

spreadsheets do not follow a consistent approach in regards to data structure. The majority of the interviewees work in teams with designers and other stakeholders to identify risks and populate the spreadsheets in workshop sessions where 2D drawings are reviewed.

Implementation of the new version of 3D Repo's SafetiBase

In general, internal and external stakeholders supported the implementation of the new version of 3D Repo's SafetiBase in the selected projects. There were challenges with external stakeholders in some instances, as noted by a health and safety manager, who said that *"Where we need to capture information from [external stakeholders] it's challenging […] There are concerns in terms of responsibilities of people involved in the [research] project […] they do not want to get involved in terms of creating a database where they are responsible and liable for creating safety treatments for common safety risks. We explained to them that [this was not the case…] but there were not much developments from external stakeholders"*.

In one of the pilot projects, there was a struggle to get involvement from both external and internal stakeholders, as stated by the principal designer, who was the only one using SafetiBase: *"We had great difficulty with getting the agreement from the client side of things for the use, [as well as] getting any engagement from the rest of the design team. Everyone is stuck in their ways to some extent"*.

Three interviewees stated that they implemented SafetiBase in parallel to their usual Safety Management process so that introducing a new system would not impact what they were currently doing. Furthermore, one of the participants mentioned that *"we are looking at every opportunity to drive some best practice back into that project"*. Another interviewee mentioned that they implemented SafetiBase into their usual process: *"[…]very similar to what we did before […] we had a workshop and I did a screen share and ran through all of the risks and issues […] we discussed the project and added more onto [the system]"*. Finally, one of the participant projects implemented the system after their usual approach had taken place. In this particular case, the system enabled the project team to identify and remove critical risks by redesigning the facility: *"we were able to design that [risk] out […] it was a much safer solution that we came up with, and a lot cheaper as well […] Definitely a big success"*.

Benefits of the implementation of SafetiBase

After using the new version of 3D Repo's SafetiBase, the interviewees perceived that being able to add safety information to a BIM model of their facilities, and pinpointing where the risks are on the model, adds value to their current Safety Management process. One of the interviewees said that *"we will definitely do more of [using 3D models] in the future"*, and another compared the two approaches saying that *"things can easily go missing when it's just a long [risk] register, just text everywhere […] Whereas with [the new version of 3D Repo's SafetiBase] you can see the problem […] I wouldn't say it has influenced our process, but there is space to make it easier"*.

Another benefit identified by the interviewees is the fact that SafetiBase has a structured approach to input risk data, as it *"is essential, […] and it just helps improve [information] consistency"*. Furthermore, characterising the risks in this structured approach *"makes you think about what the actual risk is, rather than just writing […] words that might not accurately capture what the risk is"*.

SafetiBase enables collaboration in the risk identification and treatment selection processes, which was identified as another strength of the system. A Safety, Health and Environment lead said that *"getting a number of people around the table and [using] the tool as a group, rather than as individuals, would still be a strength, and I would recommend that I definitely a part of [the system]"*.

Some interviewees highlighted that adding more risk scenarios and treatment prompt suggestions to the Risk Library would yield more benefits.

Challenges to adopting the new version of 3D Repo's SafetiBase

A recurring challenge identified by the interviewees was that characterising the risks according to the Risk Library ontology is often a difficult task, albeit it has value. Furthermore, a few of the interviewees highlighted the need to capture risks not related to safety in a similar fashion. One of the interviewees said that *"we had difficulties [...] identifying occupational health risks, although they are available in the tool. I think it's more difficult to categorise [them]"*, and another said that *"we tend to put [business risks] in the same document [...] we do very much like to put business risk treatments in there as well. I was struggling a little bit with those elements"*.

Another potential challenge to adopting the system widely is the difficulty in getting support from stakeholders, as stated by a principal designer: *"The system is great. The issue [...] is getting engagement from the full design team [...] it needs to be pushed from the client side"*.

Challenges to adopting the risk scenario classification ontology

Another frequent challenge that the interviewees found in relation to the Risk Library schema was that in some instances it was difficult to categorise identified risk scenarios into the categories provided within the new version of 3D Repo's SafetiBase. However, one participant pointed out that with increased use of SafetiBase, familiarity with the concepts facilitates the process.

3.5 Workshop series to explore 4D capabilities

The following themes derived from the focus group discussion segment of the workshop sessions described in Section 2.6: (1) 4D to support in the safety risk assessment process, (2) 4D to visualise risks, and (3) barriers to adopting 4D for safety management in construction.

4D to support in the safety risk assessment process

The majority of the workshop participants agreed that without the BIM model, some of the safety risks identified during the session would have been overlooked. Visualising the 4D model instead of a static model was considered invaluable to identify risks associated with temporary works. For example, Figure 9 illustrates a risk scenario in which workers inside an excavation would be at risk of being struck by the temporary props highlighted in yellow if they fell on them. The majority of the participants also agreed that 4D facilitates focussing on a particular point in time in the project, which could result in identifying risks that are not evident from static BIM models representing the finished product.
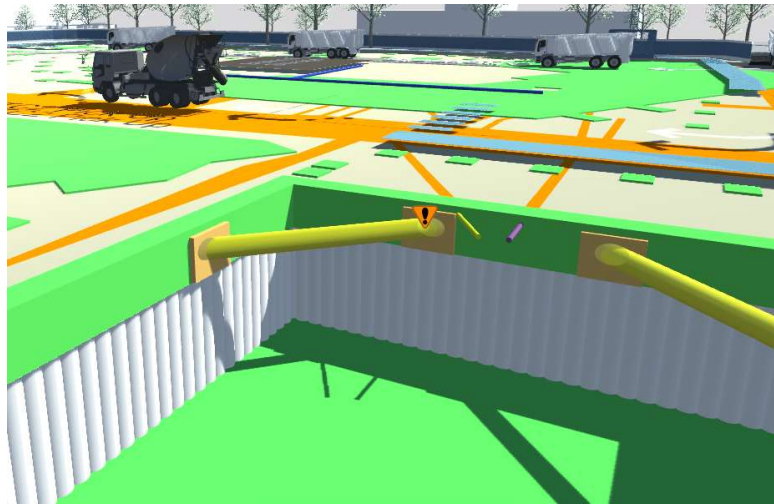
Figure 9 Identified risk scenario associated with temporary works

One participant mentioned that by identifying safety risks and marking them on the BIM model, there was potential to generate "a more meaningful" risk register, since the identified risks were project-specific rather than the more general type of risks typically included in risk registers. Another participant added that this approach provides a dynamic risk register in a 4D environment.

Participants highlighted that the results of the safety risk identification and treatment selection processes were improved when working collaboratively. One participant said that working with 4D models promoted active engagement from those involved in the activity. The researchers noticed that all the workshop participants were in fact actively engaged and participating in both the risk identification and treatment selection processes. Another participant emphasised that the combination of visualising the 4D mode with the interactive discussion surrounding selection of treatments for risks prompted thought and reflection, which was very valuable to the process.

4D to visualise risks

Some of the identified risks were not applicable throughout the construction phase of the modelled project. Motivated by this idea, the participants decided to provide start and end dates to risks to indicate when the risk was applicable. As a result, the indicative 3D pins associated to risks with start and end dates were only visible in the 4D model if the current date was in between the start and end date of the risk. The majority of the participants agreed that this approach was very useful, particularly from the perspective of temporary works planning. The option to make the duration of risk exposure visible in a 4D environment was of particular interest to participants.

Barriers to adopting 4D for safety management in construction

Identified potential limitations of the approach include that some of the identified risks need to be updated as the construction programme and 4D model evolve. Therefore, there is a requirement to review previously identified risks and manually update them in addition to identifying new risks that would arise as a result of the updated programme. This requires a significant amount of time and is a potentially error-prone process. Concerns regarding the level of skill required to use 4D technologies were also raised as a potential limitation.

4. Project outputs

This section reports the outputs produced in connection with the work presented in this report:

Data

- A CSV file of the Risk Library (v 11/06/2021) containing 401 treatment prompts for 31 risk scenarios related to 11 different risk categories available to download from the Discovering Safety website.

Conference papers

- Osorio-Sandoval, et al. (2021) 'A method to implement prevention through design using 4D BIM'. In: *38th International Conference of CIB W78, Luxemburg, LX, 13-15 October 2021. [accepted]*

Awards

- Collinge, et al. (2020) 'Discovering Safety: BIM Safety Risk Library' buildingSMART award 2020 under the professional research category for work undertaken during Phase 1 of the Use Case (https://www.buildingsmart.org/bsi-awards-2020/winners/)

Reports

- Zhou, et al. (2021) 'Characterising Risk Scenarios from HSE Archive Data-A Text Inference Approach'. Submitted to HSE as a standalone report. (Appendix E)

Community of Practice

- An active pilot project community of practice committed to continuing to work positively with the Discovering Safety Programme.

Code

- NLI classifier experimental code that can be further matured into a tool for automatic identification of risk scenarios from HSE archive data

Validation of conceptual ontology and digital tool

- Phase 2 work has validated the ontology and conceptual digital tool to assist safety work in design phase of construction projects.

5 Conclusions and Issues of Note

The work done and evidence gathered from Phase 2 of the BIM Risk Library Use Case validates the original conceptual risk scenario/treatment ontology and the premise that health and safety in construction is open to improvement via proactive use of a tool linked to a digital BIM model. The implementation of the Risk Library within an existing commercial cloud-based BIM tool, namely 3D Repo – SafetiBase, was well received by industry practitioners and enabled the exploitation of both 3D and 4D functionalities.

Piloting the tool with a wide variety of projects in industry proved that the concept of a digital tool to assist in safety management is sound. Visualising safety information within a BIM environment provides context to the identified risks, which enhances communicating risks to other stakeholders. Subsequently, this approach provides an opportunity to mitigate these risks with improved design, construction methods, and planning.

The Pilot Projects also enabled further data to be collected from industry regarding real risk scenarios and treatments from live projects: this data feeding into a growing risk library database of potential value for all industry to use and share.

The authors believe the Risk Library is a promising means to leverage knowledge and regulations on construction safety across multiple projects and organisations – the methods for expanding the Risk Library also facilitating the continuous improvement of the knowledge base. In particular, automatic classification of free text into the Risk Library categories allows gaining semantic information from documented past events.

In addition to the general conclusions, there are several issues of note of relevance for the further use/development of the tool and further development of a risk library:

- The ontology used to characterise risk scenarios provides information consistency and a way for structuring information uniformly across projects and across companies. This has the potential to create a uniformity of approach in the future. Several pilot project participants commented on the potential positive outcomes of structuring risk data uniformly, in design, maintenance and operations of buildings.

- The creation of an embryonic risk library from Phase 1 and Phase 2 work holds great potential for further development that connects with wider societal ambitions to create open, trustworthy data eco-systems for the benefit of all (c.f. the Open Data Institute: https://theodi.org/). The open sharing of data related to construction risks and their treatment would potentially advance national and international practices.

- The use of the digital health and safety tool was noted as a valuable learning tool opportunity for future designers. The learning potential of the Tool could be researched in the future.

- Several Pilots noted how the Tool is very useful for more unusual, unique risks encountered on projects, rather than routine risks.

- The use of the Tool to assist on projects in 4D generated much interest from industry (see Section 3.5). Use of a "dynamic" risk register in a 4D environment was noted as particularly exciting and innovative.

- Use of the Tool challenged and changed working practices and the role of designers in the risk/treatment identification process. Whilst different pilot projects approached use of the Tool differently (i.e. team approach / individual designer approach), there were some commonalities. For example, the use of Excel spreadsheets to capture and communicate risks to contractors was questioned, as well as the formatting/structuring of the spreadsheets.

- The contractual relationships and organisational set-up between clients and design teams on different pilot projects effected use of the Tool. Very positive use of the Tool and feedback was received from pilots where the client managed designers themselves in-house; more challenging situations arising where clients/designers were separated by contracts and distance.

- One pilot project advanced the innovative use of the tool by integrating it with their own procurement and tendering process: assessing tendering parties through their use of the digital tool in identifying solutions to design questions.

- The risk scenario and treatment validation process was time consuming. It is suggested that this process needs to be streamlined and made more efficient moving forward. One potential solution would be to have 3 treatments per scenario. Consultation with the Community of Practice on this issue is advised.

- Working at scale challenge: Phase 2 work proved the veracity of the BIM risk library concept and the value of a digital tool to assist designers working on projects with health and safety on a small number of projects from diverse sectors. A challenge now is to scale up usage across industry. This is a multi-faceted challenge requiring attention on several issues: communication/publicity; senior management support and drive; software issues (extension from 3D Repo/SafetiBase); more explicit value data generation from tool use at company/project level (metrics); tackling contractual/organisational challenges with use of tool.

- Potential HSE benefits: exploration of potential benefits to HSE from the Tool and the NLP approach to interrogating the HSE archive could be investigated. For example, the ontology concepts employed for the risk scenario definition could be integrated with HSE investigator work sheets/reports so there is an integration of concepts and more integrated workflows at multiple levels. The NLI approach employed in Phase 2 opens up options to do statistical analysis of HSE resources (e.g. RIDDORs analysis for % type of accidents, projects, etc).

- The NLP work has laid groundwork for further HSE archive interrogation: the NLI coding could be matured into a tool, with further methodological refinements made if required (annotation of data and sampling/checking for heterogeneity and accuracies).

On the basis of the work done in Phases 1 and 2 of the Risk Library use case, future work could include:

- Leveraging the contextual information within BIM models to streamline risk scenario identification.

- Aligning the proposed NLI classifier to the Risk Library ontology to identify recurring risk scenarios in documented past events in order to address them with appropriate treatment prompts that can be appended to the Risk Library.

- Developing a training tool (e.g. serious game) to familiarise new users with the categories used in the Risk Library to facilitate risk characterisation.

- Pilot testing the implementation of the Risk Library within a 4D BIM environment in a live construction project to better understand how this approach could leverage the time element of 4D models.

- Adopting a data structure approach to characterise non-safety related risks, such as business risks and occupational health risks.
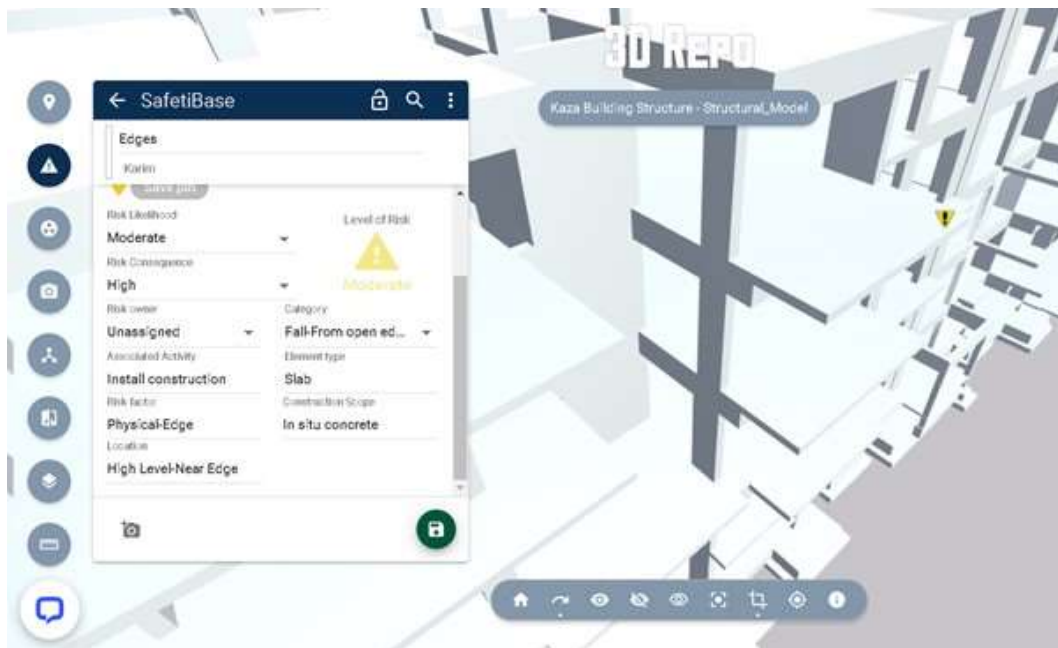
6 References

British Standards Institution, 2018. PAS 1192-6:2018 Specification for collaborative sharing and use of structured Health and Safety information using BIM. BSI Standards Limited 2018.

Choe, S., Leite, F., 2017. Construction safety planning: Site-specific temporal and spatial information integration. Autom. Constr. 84, 335–344. https://doi.org/10.1016/j.autcon.2017.09.007

Hallowell, M.R., Gambatese, J.A., 2010. Qualitative Research: Application of the Delphi Method to CEM Research. J. Constr. Eng. Manag. 136, 99–107. https://doi.org/10.1061/(asce)co.1943-7862.0000137

Hossain, M.A., Abbott, E.L.S., Chua, D.K.H., Nguyen, T.Q., Goh, Y.M., 2018. Design-for-Safety knowledge library for BIM-integrated safety risk reviews. Autom. Constr. 94, 290–302. https://doi.org/10.1016/j.autcon.2018.07.010

ISO Guide 73, 2009. Risk Management - Vocabulary.

Kasirossafar, M., Ardeshir, A., Shahandashti, R.L., 2012. Developing the Sustainable Design with PtD Using 3D/4D BIM Tools, in: World Environmental and Water Resources Congress 2012. American Society of Civil Engineers, Reston, VA, pp. 2786–2794. https://doi.org/10.1061/9780784412312.279

MacCartney, B., Manning, C.D., 2009. Natural language inference. Stanford: Stanford University, 2009.

Ørngreen, R., Levinsen, K., 2017. Workshops as a Research Methodology. Electron. J. e-Learning 15, 70.

Ove Arup and Partners, Gilbertson, A., 2015. CDM 2015 – construction work sector guidance for designers , fourth edition. CIRIA.

Poghosyan, A., Manu, P., Mahamadu, A.M., Akinade, O., Mahdjoubi, L., Gibb, A., Behm, M., 2020. A web-based design for occupational safety and health capability maturity indicator. Saf. Sci. 122, 104516. https://doi.org/10.1016/j.ssci.2019.104516

Yuan, J., Li, X., Xiahou, X., Tymvios, N., Zhou, Z., Li, Q., 2019. Accident prevention through design (PtD): Integration of building information modeling and PtD knowledge base. Autom. Constr. 102, 86–104. https://doi.org/10.1016/j.autcon.2019.02.015

Appendix A   SafetiBase User Guide

SafetiBase Risk Suggestion Tool

User guide

How to use the SafetiBase Risk Suggestion Tool for better construction health and safety

University of Manchester

Version 3.1 (November 2020)

Contents

Appendices

Part 1

1.1   Introduction

This Guide is to assist construction project designers in using the SafetiBase Risk Suggestion Tool for better health and safety management. The Tool is an extension of the existing SafetiBase platform hosted on 3D Repo. The SafetiBase Risk Suggestion Tool has been developed by 3D Repo, in collaboration with Atkins, The University of Manchester and the Health and Safety Executive (HSE) as part of the Discovering Safety Programme (DSP)[1] funded by the Lloyd's Register Foundation.

The SafetiBase Risk Suggestion Tool is designed to provide designers with appropriate health and safety information and knowledge to assist them in their work regarding health and safety risk identification and possible treatments. The Tool is also a useful aid in helping to meet obligations under the UK Construction Design and Management (CDM) regulations.

The logic underpinning the intelligence of the Tool is based on research undertaken by The University of Manchester and HSE. This is elaborated upon more in Section 2.1.

The SafetiBase Risk Suggestion Tool is currently accessible and usable via 3D Repo.

1.2   Overview and benefits of use

The SafetiBase Risk Suggestion Tool is designed to provide users with appropriate health and safety information to assist them regarding construction health and safety risk identification and possible treatments. The Tool is designed to be used by designers as they concurrently work on their project design models. Designers using the SafetiBase Risk Suggestion Tool can identify a risk scenario using a set of pre-defined categories that lead to appropriate risk treatment suggestions where available, otherwise they can insert their own.

The easy-to-use cloud-based Tool facilitates collaborative working across the asset life-cycle.

1.2.1   Benefits

- Provide a "dynamic risk register", where each risk is recorded and visualised in a structured way.

- Enable to share risk scenarios & treatments across different projects in your organisation, bringing long-term time/cost savings.

- Access to risk scenarios & treatments of other projects (national and international) via the Discovering Safety Programme.

- Access to the HSE archive of health and safety information by uploading an online-available master CSV file into the SafetiBase Risk Suggestion Tool.

- Serve as an aid in meeting obligations under CDM regulations.

- Provide learning/training opportunities for designers.

1.2.2   Where does the data come from?

The SafetiBase Risk Suggestion Tool holds a CSV file that contains data related to one of the most common risk types (falling from open/edge in concrete in-situ buildings), 9 related risk scenarios generated from 192 incidents and 162 treatment plans.

---

[1]Further information about DSP and the SafetiBase Risk Treatment Database project may be found at:

https://www.discoveringsafety.com/

https://3drepo.com/3d-repo-4-6-new-features/

https://www.ashtoninstitute.ac.uk/   2

Further risk scenarios/treatments can be:

- Added manually to the SafetiBase Risk Suggestion Tool.

- Uploaded to the SafetiBase Risk Suggestion Tool from CSV files.

- Downloaded from the HSE archive of health and safety information via CSV file uploads.

1.2.3   When should I use it on a project?

The SafetiBase Risk Suggestion Tool should be used first as early as feasibility/optioneering stages (RIBA stages 1 and 2/PCF stages 1-3/GRIP stages 1-4) but can be used in later stages too, whether that's detailed design, construction or even operations and maintenance. The Tool has the ability to enable the capture of risk scenarios at various project stages (i.e. Preliminary Design; Detail Design; Pre-Construction; Site Work/Temporary Work).

Early design phase engagement is recommended so that risk scenarios and their treatments can be identified as soon as possible. This early identification and consistent management of risks creates a continuous narrative for the collective risks for all associated parties.

1.3   Software requirements

The SafetiBase Risk Suggestion Tool is currently accessible and usable online via 3D Repo. In addition to this guide, there is also 3D Repo's user guide at the url https://3drepo.com/support/getting-started/

A **CSV master file** with risk scenarios and treatments will be made available to DSP Phase 2 pilot projects members. The CSV file consists of a collection of data in a tabular format, which can be visualised in a number of software packages, including Microsoft Excel. Data contained in this file compiles the results of research undertaken by The University of Manchester and HSE during Phase 1 of the Discovering Safety Programme. The CSV file can be uploaded into the SafetiBase Risk Suggestion Tool to maximise its potential and use.

3

Part 2

2.1 Tool concept & design

The SafetiBase Risk Suggestion Tool is rooted in a series of concepts that have industry validation. The conceptual ontology is shown in Figure 2.1.
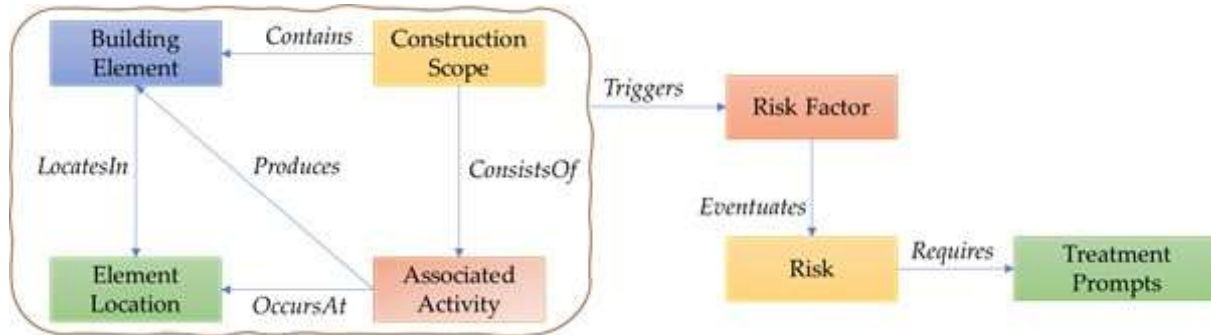

Figure 2.1: Risk scenario/treatment concepts

The concepts shown in Figure 2.1 combine together to eventuate a risk scenario that requires a treatment. Each concept has a series of sub-categories, derived from available taxonomies that appear as options on the SafetiBase Risk Suggestion Tool for designers to choose from. The root of each concept is detailed below

2.1.1 Six concepts & treatment prompts

Construction Scope: based on CIRIA C755 CDM 2015 guidelines

Construction scope concept is structured on type of construction work, divided into five main groups: General planning, excavation and foundations, primary structure, building elements and building services, and civil engineering. Each group is subdivided into several categories which categorise the construction work. In total 40 categories are identified under the five groupings, listed in Table A.1 (Appendix A).

Risk: based on PAS 1192-6 2018

Risk concept identifies different types of risks that could occur in construction sector, classified into 29 possible risk categories. Table A.2 (Appendix A) shows these categories.

Element

This concept comprises the different building elements associated with the eventuation of hazards; elements are classified by related building design disciplines to enable them to be easily assigned to the responsible designer.

Location

Location concept identifies the characteristics of the location which can be the reason why a risk arises. For prototype development, two main groups were identified: high-level location and site logistics. Highlevel group includes locations which could trigger a hazard at a high level. Site logistics group includes locations that can be critical and hazardous during construction and operation/maintenance. Table A.3 (Appendix A) shows the categories into which each group is divided.

Activity

This concept is divided into 16 categories at a high/generic level, as the method of construction is usually not known during design. The proposed 16 activity categories, listed in Table A.4 (Appendix A), cover

4

several work activities, ranging from preliminary site-investigation and testing to material disposal or re-use.

Risk Factor

This concept seeks to identify the reason behind the risk eventuation. It is divided into three groups: physical, material and task. The physical group comprises the factors related to the physical characteristics of an element. The material group comprises the type of material which could impact on the health and/or safety of workers such as lead and asbestos. It also includes the material strength which could be a reason behind structure failure. The task group covers job-steps which could eventuate a hazard such as manual handling and lifting. It also includes the work scopes which repeatedly are associated with specific hazards such as temporary works and excavation. Table A.5 (Appendix A) lists the categories into which each group is divided.

Treatment Prompt

This concept identifies the solution suggested by designers to deal with an existing risk scenario, the rationale behind using the word *treatment* instead of *mitigation* is that *treatment* is a more neutral term than mitigation.

For the prototype Tool, suggested treatment prompts were classified based on two different concepts: Type of treatment (Eliminate, Reduce, Inform and Control), and the project life cycle stage (Preliminary Design; Detail Design; Pre-Construction; Site Work/ Temporary Work).

2.1.2 Prototype Tool: Risk & Treatments

For development of the Tool, a specific risk was chosen (falling from height in in-situ concrete structures). 9 related risk scenarios were identified from an analysis of HSE data (RIDDOR reports and Press Releases). 162 Treatments for these risk scenarios were derived from Workshops and industry Guidelines, captured in a Treatment Matrix (Figure 2.2). These risk scenarios and treatments were converted by the research team into a usable interface for direct designer input in BIM environments.

Figure 2.2: Treatment matrix

An empty Treatment Matrix can be found in Appendix B.

5

Part 3

3.1 Populating/updating the CSV file

To maximise potential application of the SafetiBase Risk Suggestion Tool, users are recommended to actively and progressively populate the Tool with new risks and treatments as they develop their project design models. These risks/treatments can be saved, retrieved and edited as required.

3.1.1 Download treatment suggestion template

The CSV file template[1] to populate with risk scenarios and treatments is available on SafetiBase. To download it, follow these steps:

1. Login to 3D Repo using your personal username and password
   (*If you need support getting started with 3D Repo, please refer to their user guide at the url https://3drepo.com/support/gettingstarted/*)

2. On the *Teamspaces* screen, click on the gear icon to go to the *Teamspace Settings* menu (Figure 3.1)
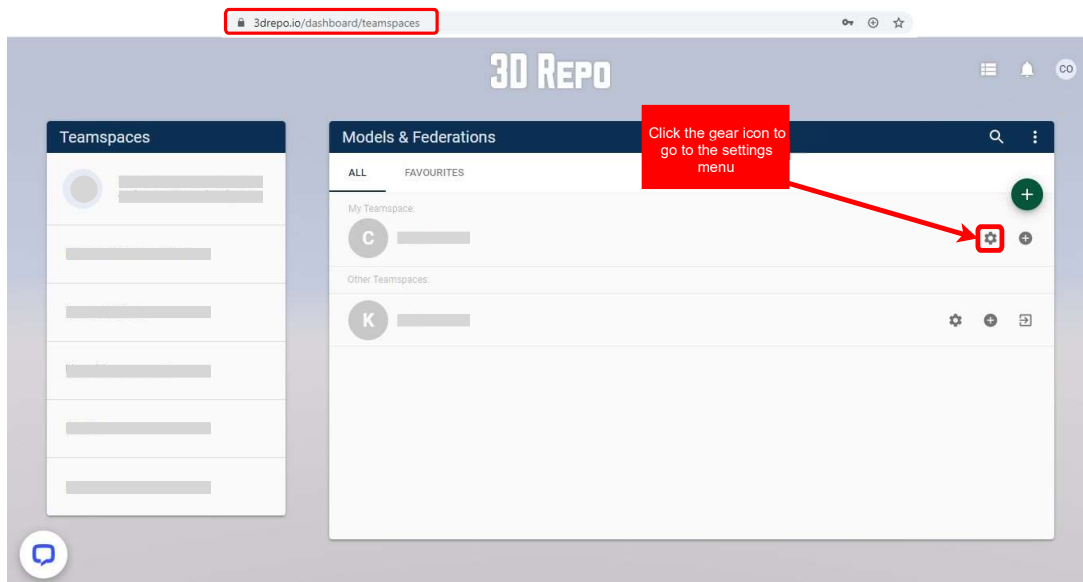
Figure 3.1: Teamspaces screen

3. Click on the *GET TEMPLATE* button on the *Teamspace Settings* menu to download the treatment suggestion template in CSV format (Figure 3.2)

---

[1] All DSP Phase 2 Pilot projects will use the existing Phase 1 CSV file as a starting point.

6
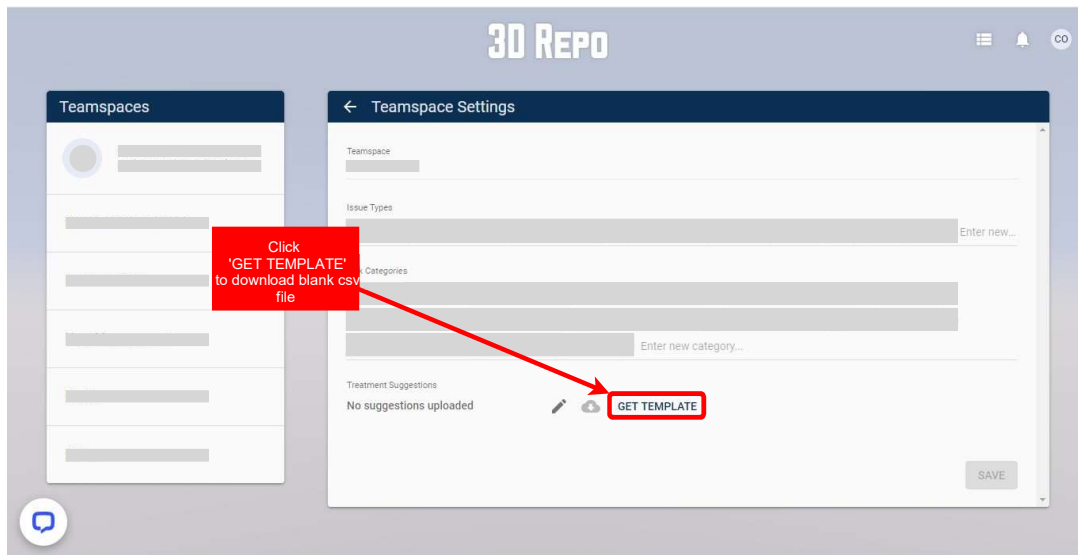
Figure 3.2: Download treatment suggestion template

3.1.2 Populate template with own risk/treatment suggestions

The risk/treatment suggestion template downloaded as detailed in Section 3.1.1 consists of a CSV file, which allows data to be saved in a tabular format. The template can be opened, populated and saved with many spreadsheet software, including Microsoft Excel. Figure 3.3 shows the structure of the data within the table.
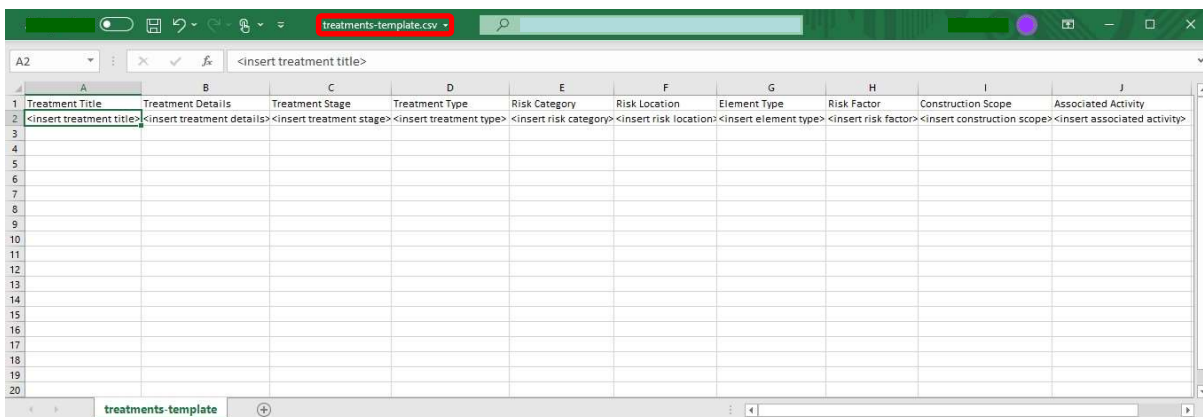
Figure 3.3: Blank treatment suggestion template

Refer to Appendix A to populate the CSV file with existing categories for each of the six concepts outlined in Section 2.1. Choosing an existing category for the *Risk* concept is of upmost importance to comply with PAS 1192-6 2018.

Use the empty Treatment Matrix (Appendix B) to support treatment identification across design phases.

3.1.3 Upload CSV file with risk/treatment suggestions

New risk scenarios/treatments can be added to SafetiBase after editing the CSV, this enables users to use knowledge from previous projects. To do this:

1. Login to 3D Repo using your personal username and password
   (*If you need support getting started with 3D Repo, please refer to their user guide at the url https://3drepo.com/support/getting-started/*)

<div align="center">7</div>

2. On the *Teamspaces* screen, click on the gear icon to go to the *Teamspace Settings* menu (Figure 3.4)
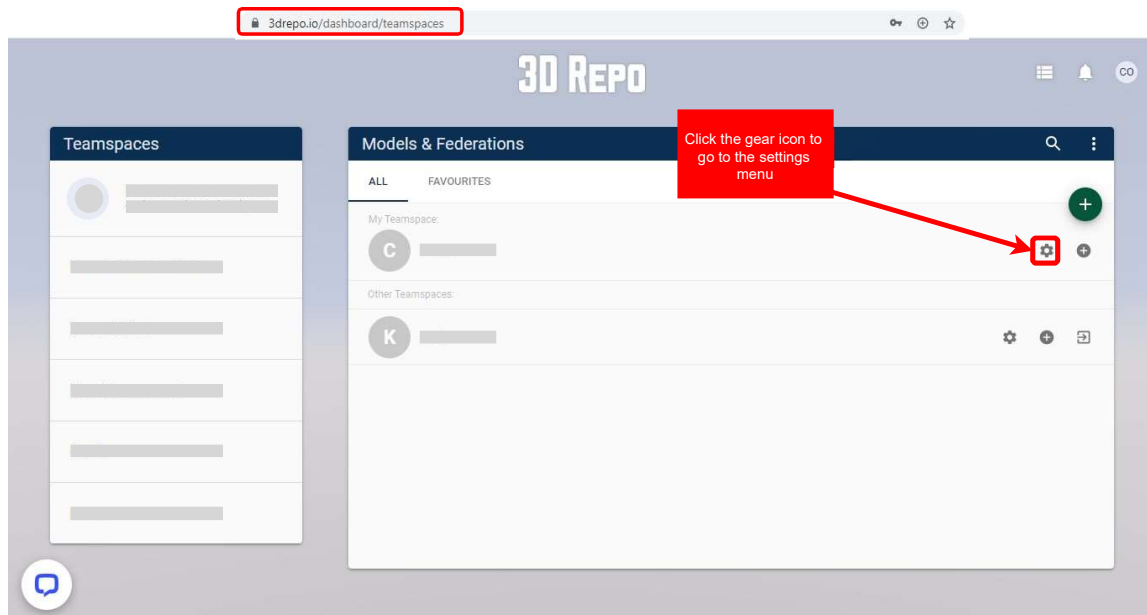
Figure 3.4: Teamspaces screen

3. On the *Teamspace Settings* menu, click on the pencil icon to upload your own risk/treatment suggestions CSV file (Figure 3.5)
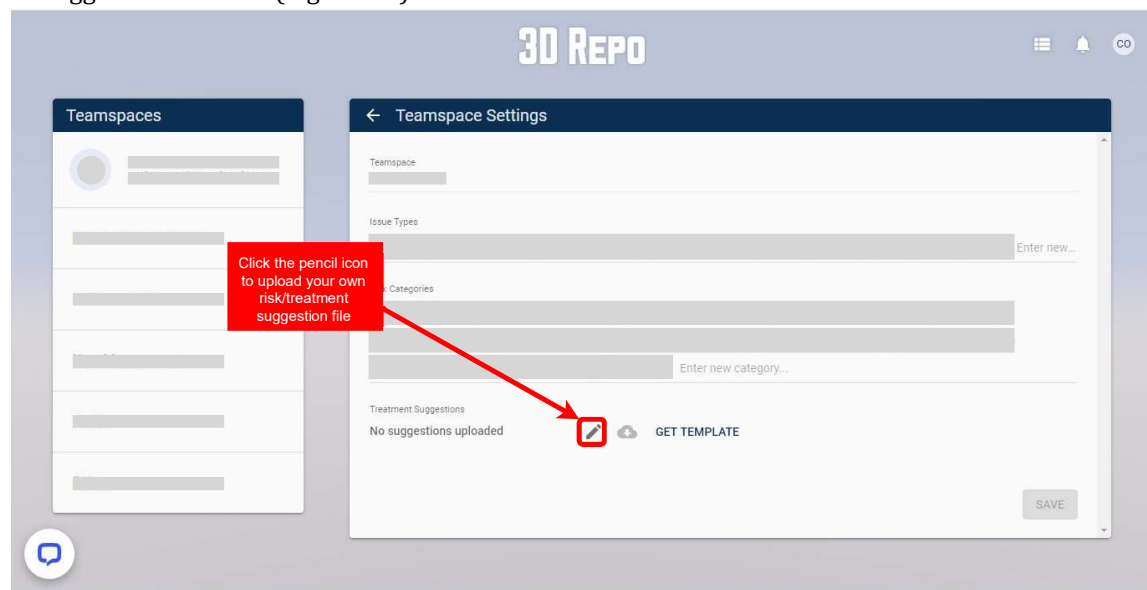
Figure 3.5: Upload risk/treatment suggestions CSV file

Discovering Safety   © Crown Copyright, Health and Safety Executive 2019   36

4. On the file browser, find your risk/treatment suggestions CSV file and click *Open* (Figure 3.6)
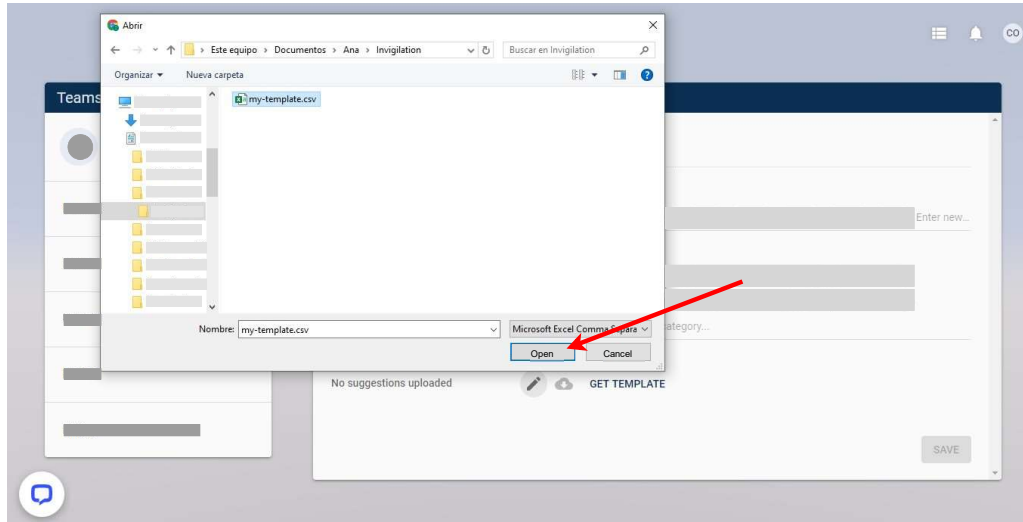
Figure 3.6: Select which CSV file to upload

5. On the *Teamspace Settings* menu, verify the name of the file that you wish to upload and confirm by clicking on the *SAVE* button (Figure 3.7)
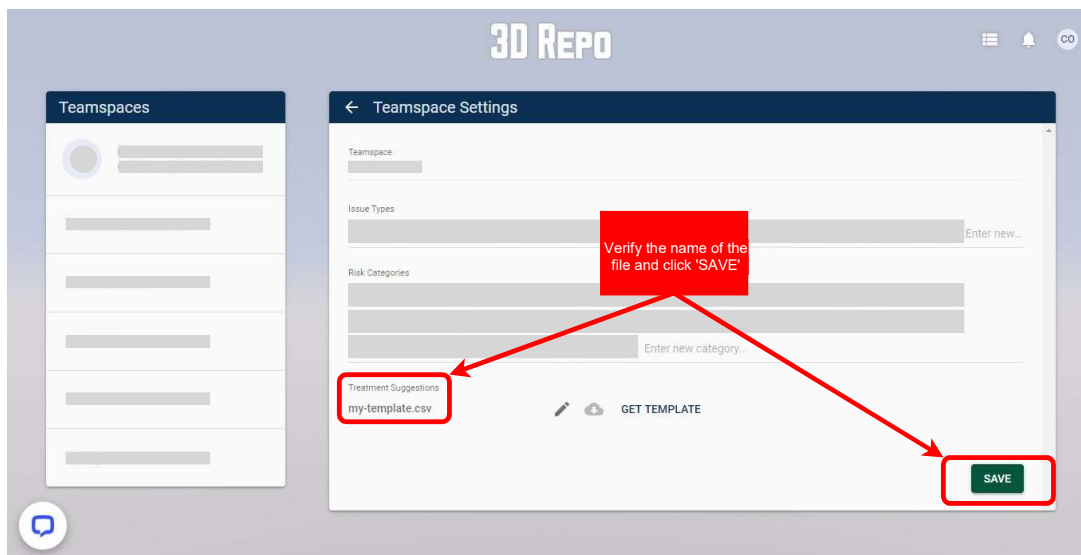
Figure 3.7: Confirm your upload

6. Once the CSV file is uploaded, users can open their model in 3D Repo and find treatment suggestions for identified risk scenarios (as detailed in Section 3.2).

9

3.2   Using the Tool

3.2.1   Identify risk scenarios

Follow these steps to identify risk scenarios using SafetiBase:

1. Login to 3D Repo using your personal username and password
   (*For more information go to 3D Repo's* user guide *at https://3drepo.com/support/*)

2. Open your BIM model in the 3D Repo platform

3. Enable the SafetiBase Add-in by clicking the SafetiBase icon (Figure 3.8)

Figure 3.8: Enable the SafetiBase Add-in

4. On 3D Repo, navigate the model and display an area of the BIM model where a risk can be identified. 3D Repo's functions (measures, clip can be useful to achieve this)

5. On SafetiBase, click on the *plus sign* to add a new risk (Figure 3.9)
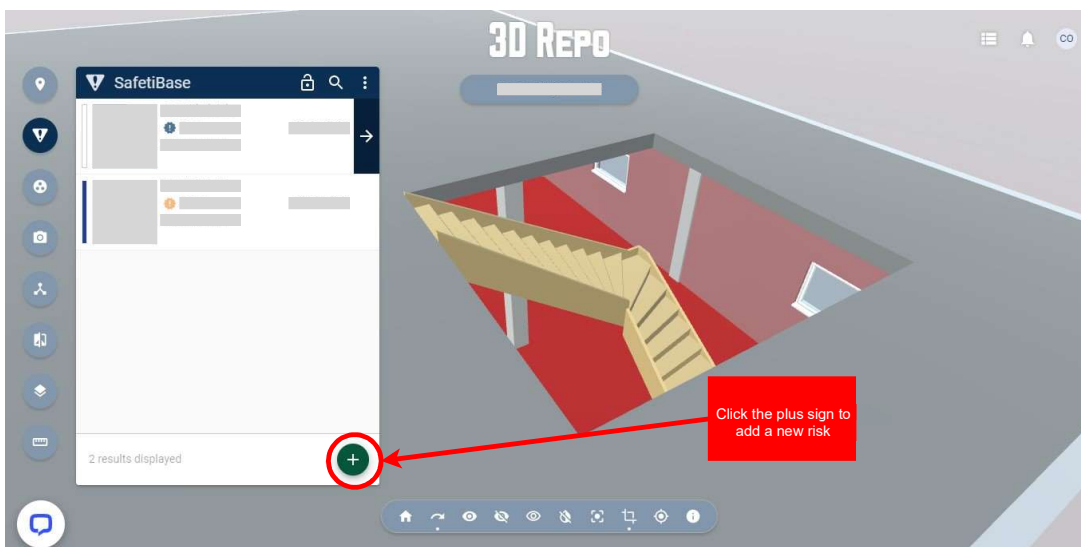
Figure 3.9: Add new risk

10

---

6. On the *RISK* tab, complete the required information (Figure 3.10):

   •   Add a name and description for the risk

   •   Add a *Pin* to the 3D model to show where the risk is located

   •   Add a screen shot of the model for better communication

   •   Identify the likelihood and consequence of the risk (the colour of the warning sign and *Pin* are based on the level of the risk)

   •   Assign the risk to a specific stakeholder

   •   Identify the category for each one of the 6 concepts related to the risk scenario (refer to Section 2.1.1). The drop-down menus will suggest appropriate categories for each concept as defined in the uploaded risk/treatment suggestion CSV file (Section 3.1.3)
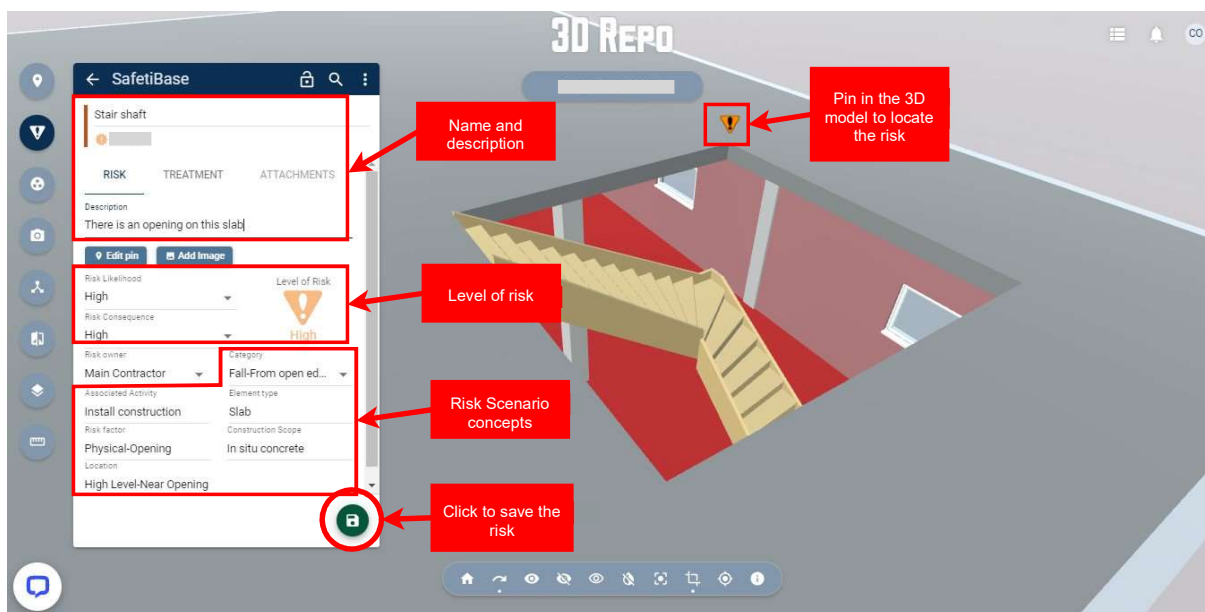
Figure 3.10: Risk information

3.2.2   Save risk/treatment information to project

Click the *save* button (Figure 3.10) to save the risk in the project. Further changes will be automatically updated.

3.2.3   Select treatment from list of suggestions

The SafetiBase Add-in can suggest appropriate treatments for an identified scenario based on the uploaded risk/treatment suggestions CSV file (see Section 3.1.3). To select a treatment follow these steps:

1. On the *TREATMENT* tab of the SafetiBase Add-in, click the *Suggest* button (Figure 3.11)

2. On the *Suggested Treatments* window, click the *Select* button of the selected treatment (Figure 3.12)
[Tip: Use the *Stage* and *Type* drop-down menus to filter the suggestions]

3. Back on the *TREATMENT* tab, complete the treatment details, if required, and save

11
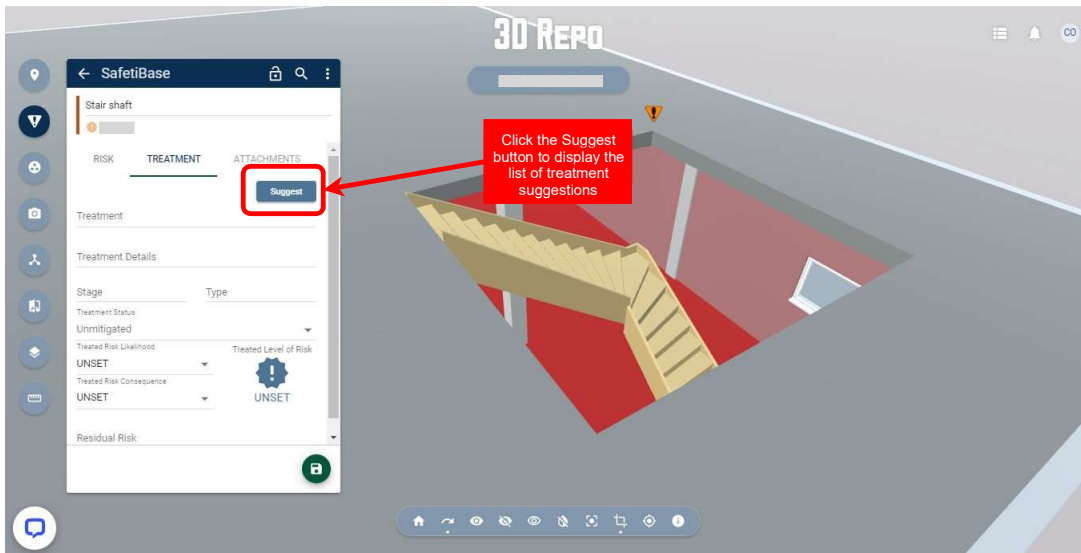
Figure 3.11: Display treatment suggestions
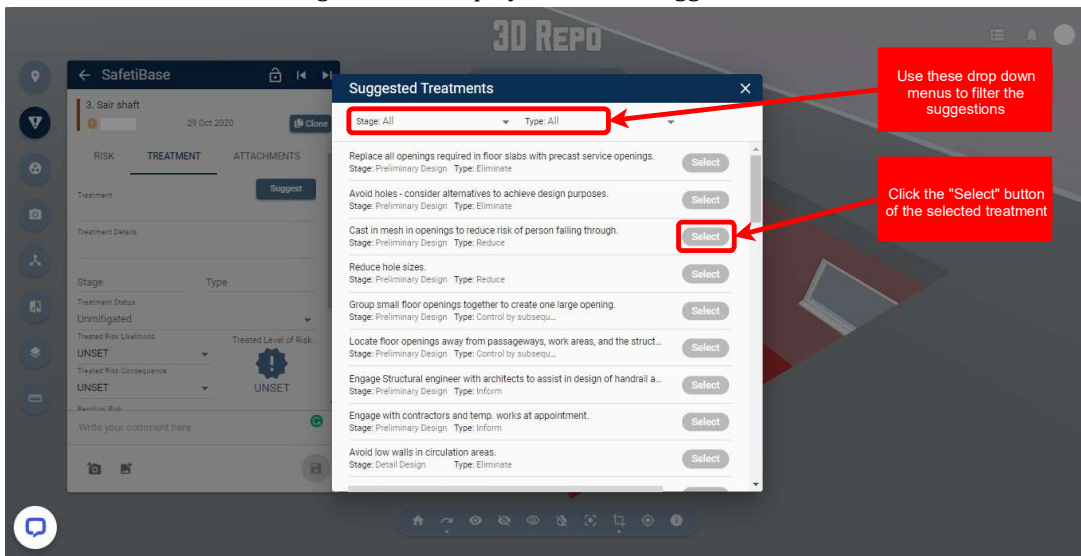
Figure 3.12: Suggested Treatments window

4. Users can also attach files and links to provide more information related to the risk/ treatment on the *ATTACHMENT* tab

If no suitable treatments are available, new treatments can be input directly into the SafetiBase Risk Suggestion Tool by completing the information on the *TREATMENT* tab. Use the empty Treatment Matrix (Appendix B) to support treatment identification across design phases.

12

3.3   Managing and exporting information

Once all the risks have been identified, the user can export them:

- As a JSON file to be re-used in other tools/applications
- Within an Excel file
- Within a Power BI dashboard
- As a report and print them as a PDF

The most preferred solution is to maintain the risks and their treatments throughout the lifetime of an asset. Users can tag other users using the '@' symbol within the comments section and the '#' symbol to tag other risks to bring to others attention (see Figure 3.13).
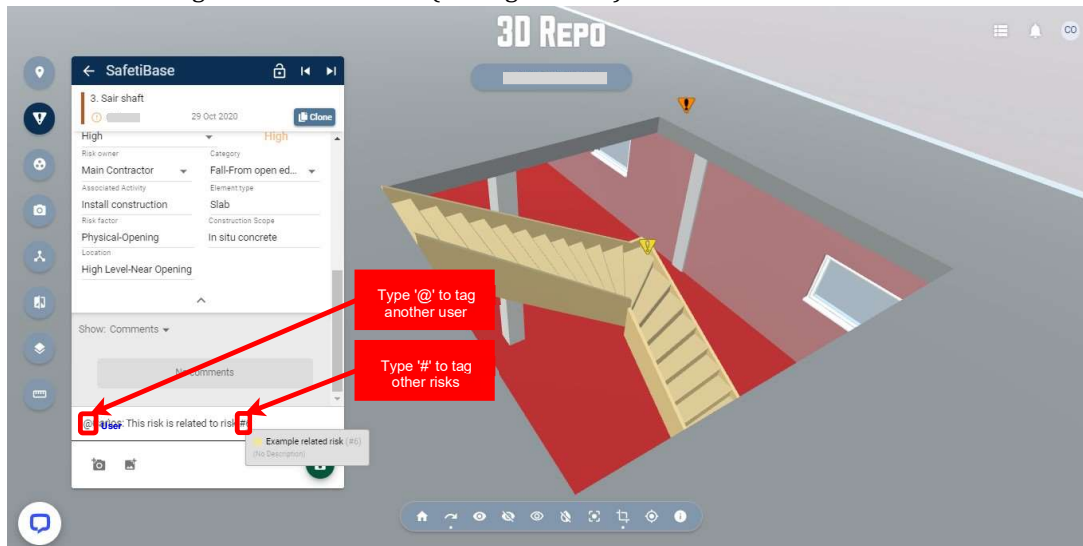
Figure 3.13: Tagging users and risks

3.3.1 Exporting JSON file

On the SafetiBase Add-in, click on the three dots at the right of the window menu bar to display commands and select *Download JSON* to export the data into a JSON file (Figure 3.14)
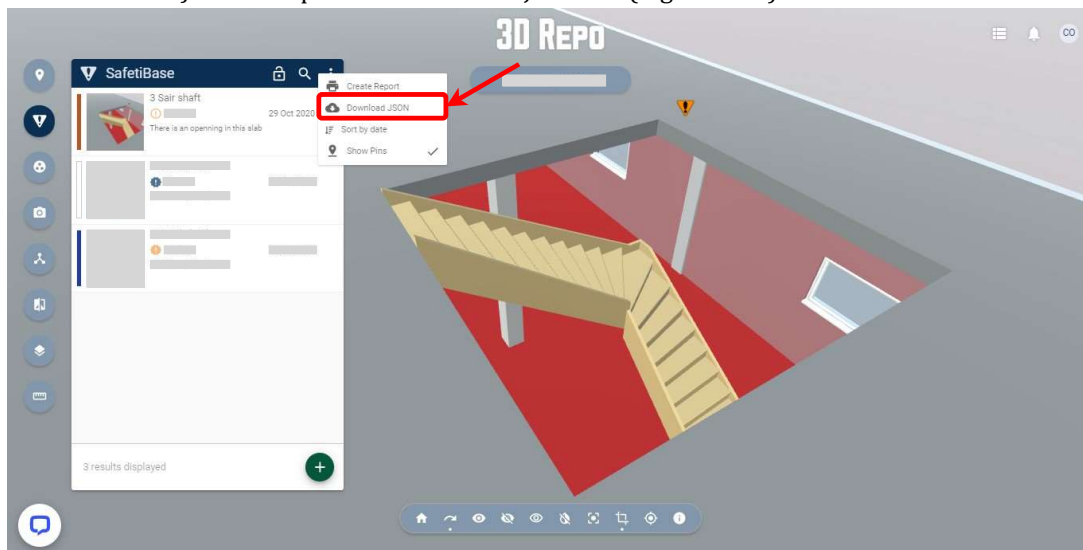
Figure 3.14: Export JSON file

13

3.3.2 Exporting into Excel file

To export risk information into an Excel file, follow these steps:

1. Download risk JSON file (see Section 3.3.1)

2. Download the *SafetiBase Risks Table Template* Excel file from 3D Repo downloads

3. Open the *SafetiBase Risks Table-Template* Excel file

4. Edit the data source settings in your Excel file as shown in Figure 3.15
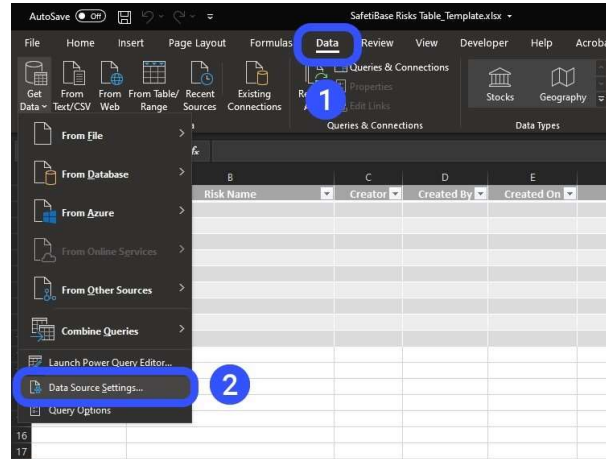
Figure 3.15: Data source settings in Excel

5. Click on *Change Source* and locate the JSON file (Figure 3.16)
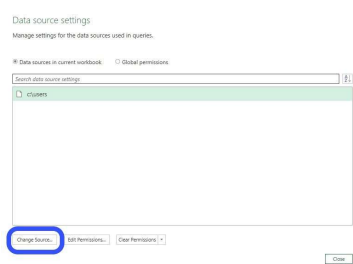
Figure 3.16: Change Source

6. Select the JSON file to remember its location (Figure 3.17)
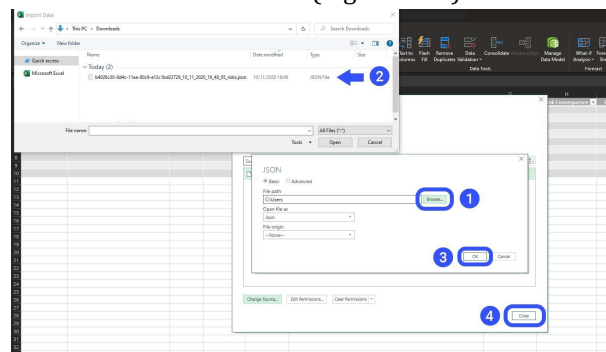
Figure 3.17: Select JSON file

7. Refresh table to see the risks from 3D Repo

14

3.3.3   Exporting into Power BI dashboard

Refer to 3D Repo's user guide on issues and risks dashboard in Power BI to export information into a Power BI dashboard.

3.3.4   Printing PDF report

To print a PDF report, follow these steps:

1. On the SafetiBase Add-in, click on the three dots at the right of the window menu bar to display commands and select *Create Report* (Figure 3.18)
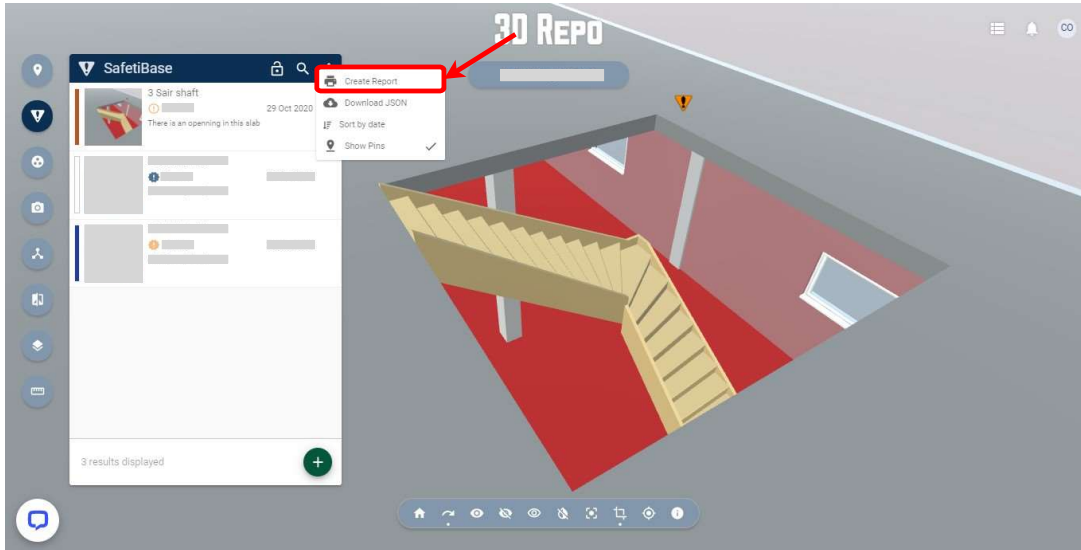
Figure 3.18: Create Report command

2. The report will be displayed in a new window. On the report window, click the *PRINT* button to save to PDF (Figure 3.19)
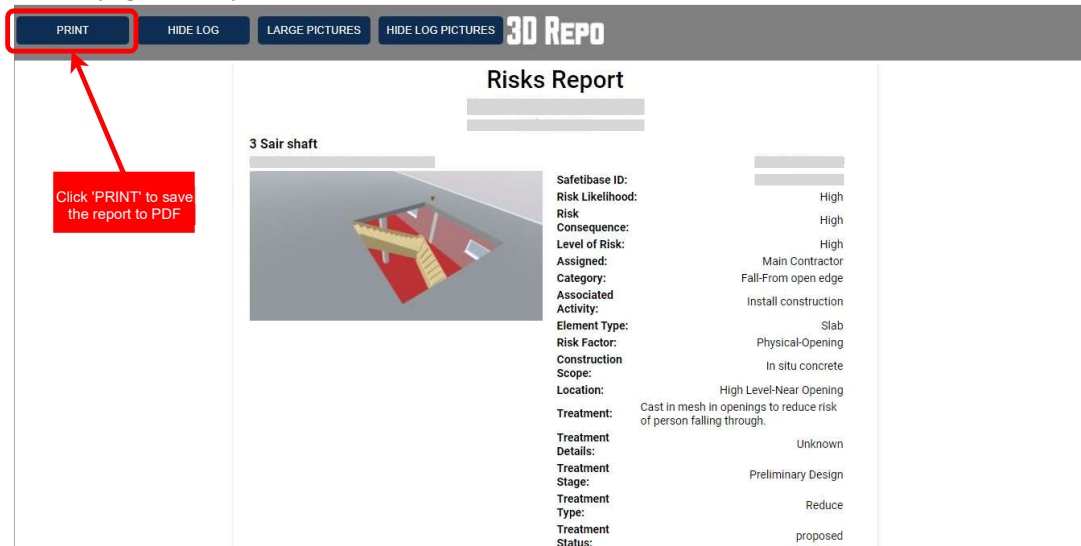
Figure 3.19: Report window

15

Part 4

4.1 Pilot project workflow

The assessment of the SafetiBase Risk Suggestion Tool is one of the key objectives of the DSP. To achieve this objective, the Tool will be piloted by key industry partners in 5 live projects in the design phase. Tool piloting will have a duration of up to five months.

The workflow of this component of the research programme is described below.

4.1.1 Overview

A master CSV file containing a database of risk scenarios and treatment suggestion prompts is maintained by the DSP team online. The DSP team will collect data related to new risk scenarios and treatments identified by pilot project users as they develop their project design models. The master CSV file will be updated based on review and analysis of this data. Updated versions of the master CSV file will be available for download periodically to be fed into the Tool, as shown in Figure 4.1.
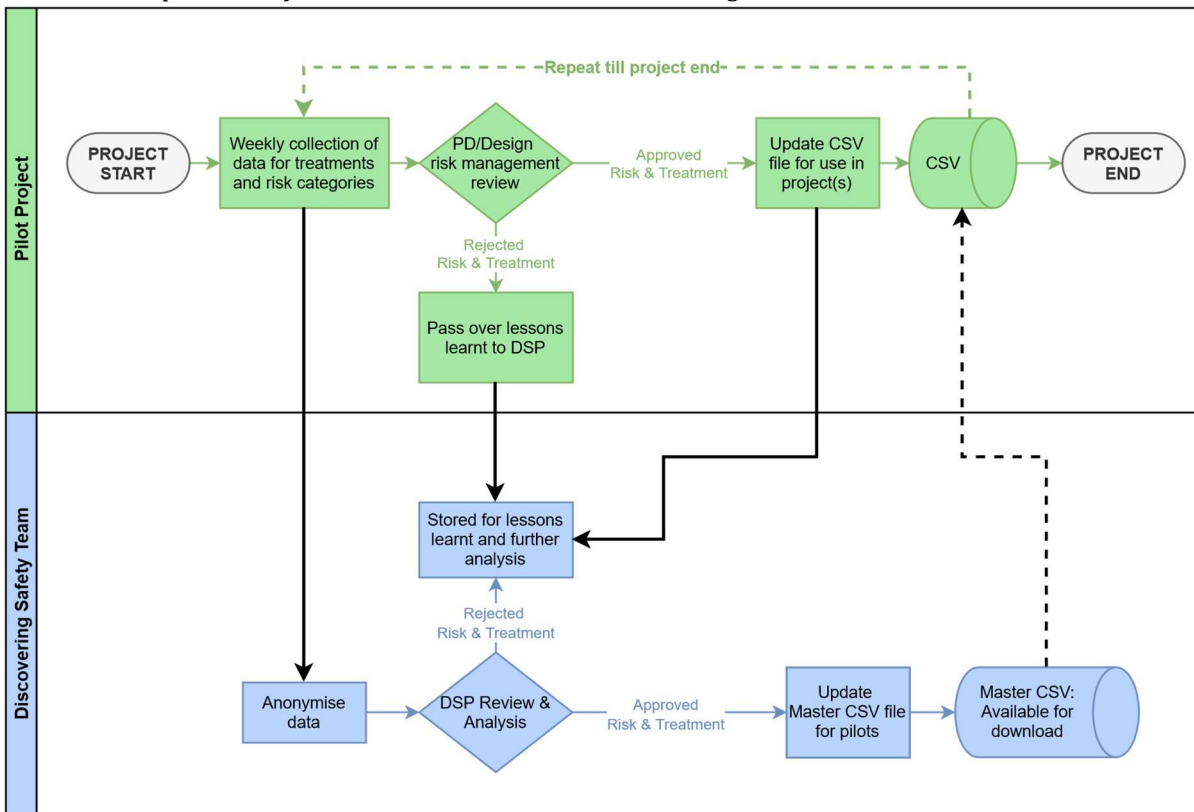
Figure 4.1: CSV management workflow

4.1.2 Data collection and management

A member of the DSP team will have access to pilot projects' 3D Repo Teamspaces. The DSP team member will have the solely responsibility of collecting data from the model as follows:

- Risk and treatment data will be downloaded by the DSP team member with a previously agreed frequency (weekly, biweekly, monthly).

- The data will be downloaded in JSON format.

- The data will be anonymised removing all project-specific information.

16

Subsequently, the data will be shared among the DSP team, who will:

- Review and analyse the anonymised data.

- Update the master CSV file with new identified risk scenarios and treatments where appropriate.

- Notify pilot project users when a new version of the master CSV file is released so that they can download it and feed it into the Tool at their discretion.

4.1.3   Feedback

An online Community of Practice (CoP) and website will be established for Pilot projects to exchange experiences and opinions of the SafetiBase Risk Suggestion Tool. This is to increase learning and sharing of experiences.

During the Tool piloting period, the DSP team will meet regularly with pilot project participants to review new identified scenarios, address any encountered issues and collect direct feedback to inform overall data integrity and quality assurance issues.

After the Tool piloting period ends, the DSP team will request feedback from pilot project users in terms of:

- Tool functionality.
- Risk scenarios and treatments in the master CSV file.

- Overall Tool utility and value for company/industry.

Feedback from the pilot stakeholders will be analysed quantitatively and qualitatively to identify the areas of strength as well as areas for improvement of the Tool. This information will provide a basis for further refinement of the Tool and further development work in Phase 2 of the DSP.

17

Further help & guidance

Use of the Tool is considered here:

- A detailed YouTube video by 3D Repo demonstrating the use of the SafetiBase Risk Treatment Tool.

- A YouTube video webinar by 3D Repo presents the SafetiBase Risk Treatment Tool.

- A background of SafetiBase by Zane Ulhaq, Associate Director of Atkins and chair of SafetiBase.

Background Research information for the SafetiBase Risk Treatment Tool can be found here:

- A video tutorial (6 minutes 40 seconds) is available online at:

https://videoreview.techsmith.com/review/GeMGqqrewIAtWQPfO6a9Fg/v1

This video is also available as an MP4 video from the DSP website and can also be obtained directly from William Collinge (The University of Manchester) or Gordon Crick (HSE).

- A longer video tutorial (10 minutes 15 seconds) is available online at:

https://videoreview.techsmith.com/review/KgSoVrxKE4M5aafapbnE2g/v1

If you have further questions/issues concerning the Tool, please contact Dr Bill Collinge (The University of Manchester) or Gordon Crick (HSE).

18

Guide authors

For queries/questions regarding the BIM Safety Risk Library tool, please contact the Guide Authors:

Dr William Collinge: BIM Safety Risk Library Principal Investigator (PI)
The University of Manchester william.collinge@manchester.ac.uk

Carlos Osorio-Sandoval: BIM Safety Risk Library researcher
The University of Manchester
carlos.osoriosandoval@manchester.ac.uk

Gordon Crick
Health and Safety Executive gordon.crick@hse.gov.uk

Zane Ulhaq
Atkins/SafetiBase zane.ulhaq@atkinsglobal.com

19

A.1 Concepts classification

Table A.1: Construction Scope

General planning	Excavation and foundations	Primary structure	Building elements and building services	Civil engineering
Site investigation and remediation	General excavation	General concrete	External Cladding	General civil engineering, including small works
Surrounding environment	Deep basements and shafts	In situ concrete	Roof coverings and finishes	Roads, working adjacent to, maintenance of
Site clearance and demolition	Trenches for foundations and services	Precast concrete	Atria	Railways, working adjacent to, maintenance of
Access (onto and within site)	Retaining walls	Prestressed, post tensioned concrete	Windows/glazing including windows cleaning	Bridge construction
Site Layout	Ground stabilisation	General steelwork	Surface coating and finishes	Bridge maintenance
	Pilling	Stability and erection of structural steelwork	Cleaning of buildings	Working over/near water
	Under pinning	Masonry	Mechanical services	Pipes and cables
		Timber	Electrical services	Work in coastal and maritime waters
		Refurbishment of existing buildings	Public health services	
			Lifts, escalators and auto walks	

Table A.2: Risk

Material effect	Mechanical effect	Fall	Trapped	Event	Handling	Struck
Asbestos	Noise	From ladder	Confinement	Electric shock	Materials handling	By falling object
Lead	Loss of control using hand or power tool	From open edge	Crushed by excavation	Fire or explosion	Mechanical lifting operation	By moving vehicle
Wood dust	Vibration	From scaffold	Unintended collapse	Machinery guarding	Working overhead	By machinery or part
Silica dust		Through fragile material	Drowning and flooding	Loss of control		Overturning plant or moving machinery
Chemical		Slip or trip on the same level	Asphyxiation			

20

Table A.3: Location

High level	Site logistics
Between Joist	Traffic route
Near Edge	Crane area
Near Opening	Pump area
Near Opening – Shaft	Confined area
Near Opening – Stairwell	Exposed area
Scaffolds	Excavation area

Table A.4: Activity

Preliminary investigation, test & prototypes	High impact events
Material sourcing	Modification
Component manufacture	Ageing
Storage, transport, logistics	Life extension
Install construction	Demolition, removal
Commission; site tests	Post processing
Use	Material disposal or re-use
Operation	Not Applicable
Maintenance	

Table A.5: Risk factor

Physical	Material	Task
Size	Asbestos	Excavation
Connection	Lead	Manual handling
Weight	Dust	Welding
Length	Strength	Lifting
Opening		Cleaning machinery
Fragile		Scaffold
Collapse		Site management
Collapse-Design		
Collapse-Incomplete		
Edge		
Unknown Design Change		

21

A.2 Empty Treatment Matrix

	Scenario		Risk	Construction Scope
	Element	**Location**	**Activity**	**Risk Factor**
	Eliminate	Reduce	Control by subsequent design	Inform
Preliminary Design				
Detail Design				
Pre construction				
Site work, Temp Works, Change control				

Figure B.1: Empty Treatment Matrix

22

Appendix B   Tripartite collaborative agreement for pilot project participants

Industry Partner Collaboration Agreement

Discovering Safety Programme –Safety Risk Library for Construction Pilot Project

This Collaboration Agreement is made on [insert date here] between the following Parties:

[INSERT COMPANY HERE, address, etc.] (hereinafter "the Industry Partner");

THE UNIVERSITY OF MANCHESTER (a Royal Charter Corporation registered under number RC000797, an exempt charity) of Oxford Road, Manchester, M13 9PL (hereinafter "the University"); and the

HEALTH AND SAFETY EXECUTIVE (a non-departmental public body of the UK Government) acting through its Science Division of the HSE Science and Research Centre, Harpur Hill, Buxton, Derbyshire, SK17 9JN (hereinafter "HSE");

each a "Party" and together "the Parties".

Whereas,

- the University and HSE have entered into a separate contractual arrangement (dated 26th August 2020; to which this Agreement is to be annexed) to jointly collaborate on the "Safety Risk Library for Construction" Pilot Project in order to advance learning and professional practices for industry in the area of construction project health and safety design;

- this Collaboration Agreement allows for the accession of the Industry Partner in order that it may participate in the Pilot Project, gain access to the Community of Practice group via the HSEs Discovering Safety programme website, submit information and data into the Pilot Project, access and evaluate tools to interrogate the evolving risk/treatment library and provide feedback and comments on their application and usefulness.

The Parties agree to the following terms:

1. The University will provide clear guidance and information to the Industry Partner regarding the Pilot Project, the Community of Practice, what is required and the benefits of participation.

2. Signature of this Agreement provides permission from the Industry Partner for the University to process anonymised information and data submitted by the Industry Partner and for it to be included (after processing) into a collated risk scenario and treatment library which will then be shared by HSE to the wider construction industry through the Community of Practice.

3. The Industry Partner will gain access to the Community of Practice group and will be invited to actively contribute to discussions and debate on the Pilot Project and to evaluate prototype tools for the interrogation of the evolving collated risk library dataset.

4. The Industry Partner will be asked to regularly provide risk scenarios / treatments (via upload through the Community of Practice) which will be processed and collated into the collated risk library by the University, together with those from other Industry Partners, provided that nothing in this Agreement shall oblige or commit the Industry Partner to provide any minimum contribution and all information submitted shall be at the Industry Partner's absolute discretion.

5. The Industry Partner will be asked to review and comment on the use and applicability of the prototype tools for interrogation of the collated risk library and its usefulness when applied to its ongoing construction design projects together with the other information within the shared Community of Practice group site.

6. The Parties acknowledge that the Industry Partner's participation in the Pilot Project and the Community Practice is voluntary and the Industry Partner accepts no responsibility and makes no warranty or representation as to the accuracy or completeness of the information disclosed by it or on its behalf and the University and HSE each acknowledge and agrees that it places no reliance on the information and will be responsible for making its own decisions in connection with the Pilot Project.

Obligations of the University and HSE to the Industry Partner:

7. The University and HSE will follow institutional Research Ethics and Data Management procedures in carrying out the Pilot Project.

8. The Industry Partner shall retain ownership and shall be the Data Controller for all of its information and data shared in the Pilot Project (including any Personal Data);

9. All information and data shared into the Pilot Project by the Industry Partner will be held and stored securely, under obligations of strict confidentiality, and only processed in accordance with Data Processing legislation for the purposes of the Pilot Project.

10. The University will consult regularly with the Industry Partner on the processing and use of its data in the Pilot Project and any further use of data once proof-of-concept has been demonstrated.

11. The University shall fully indemnify and hold both HSE and each Industry Partner harmless in respect to its statutory obligations as a Data Processor under the relevant legislation.

12. The University and the HSE each acknowledge that it shall not be entitled to any right or licence in respect of any information or data received from the Industry Partner unless otherwise set out in this Agreement.

Obligations of the Industry Partner to the University and HSE:

13. The Industry Partner shall encourage its staff to engage with and contribute to the Community of Practice and to evaluate the prototype tools in their construction design projects. They should use reasonable endeavours to provide constructive feedback and comments on the utility of the Community of Practice, prototype tools and to regularly contribute information and data in the form of risk scenarios and treatments for inclusion in the evolving collated risk library. All feedback shall be treated in strict confidence and not shared beyond the Community of Practice group.

Data Processing:

14. Each Party shall be solely financially liable to any relevant authority in their respective role as a Data Controller or Data Processor and for their failure to protect personal data and this liability shall not be limited in any way and overrides any other clause which may limit liability.

Table of Information and Data (including any Personal Data) to be shared and processed:

Subject matter of Processing	Construction project risks and treatments
Nature/purpose of Processing	To contribute to development of BIM Risk Safety Library
Duration of Processing	5 years
Type of Personal Data	None: data is anonymised
Categories of Data Subjects	Construction project design risks and treatments
Third countries or international organisations Personal Data will be transferred to	None
Sub-Processors	None

General:

15. This Agreement may be supplemented by a specific confidentiality agreement if the Parties so wish in order to cover any disclosure and sharing of information and data of a proprietary nature.

16. This Agreement will come into effect on the date of last signature below and will continue for a period of one (1) year. Any Party may withdraw from this Agreement by providing one (1) months written notice to the other Parties prior to such withdrawal.

17. This Agreement shall be governed by and construed in accordance with English law and each Party agrees that the English courts shall have jurisdiction to settle any claim or dispute and that their judgements will be binding, conclusive and enforceable by the courts of other jurisdictions.

The terms and conditions of this Agreement are hereby agreed to and accepted:

on behalf of **[Insert Industry Partner name]**
by:

        Signature:……………………………….

        Name: …………………………….

        Position: ……………………………

        Date:…………………………………….

on behalf of **The University of Manchester**
by:

        Signature:

        Name: …………………………….

        Position: ……………………………

        Date: ………………………………………

on behalf of the **Health and Safety Executive**
by:

        Signature:……………………………….

        Name: …………………………….

        Position: ……………………………

        Date:…………………………………….

Appendix C Questionnaire to evaluate the new SafetiBase

You are being invited to take part in a research study to evaluate the Risk Suggestion feature hosted in SafetiBase in 3D Repo using an online survey and video demonstration. This feature of an existing industry tool being piloted on a number of projects in the UK was developed as part of the Risk Library project under the Discovering Safety Programme (DSP) in collaboration with the Health and Safety Executive (HSE).

Please, read through, download and keep this [Participation Information Sheet](#) and this [Consent form](#). If you agree to the terms, select 'Yes' and proceed to the rest of the survey.

Have you read the Participant Information Sheet and Consent Form and agree to participate in the survey?

- o   Yes (1)
- o   No (2)

The following questions will help us understand more about your role and experience in relation to construction health and safety:

Q1 Please specify which area of work you are typically involved in

- o   Strategic planning (1)
- o   Design (2)
- o   Construction (3)
- o   Operations and Maintenance (4)
- o   Other (please specify) (5) _____

Q2 Please indicate your years of experience in the architecture, engineering and construction industry

_____

Q3 Please provide your professional qualifications/affiliations (e.g. ICE, RIBA, CIOB member, etc.)

_____

Q4 Please provide your educational qualifications (e.g. HNC building, BSc Architecture, BSc Health and Safety, etc.)

_____

Q5 Please choose an appropriate score to show which tools are used within your projects/assets to **manage health and safety information**. [Rate the tool as zero if it is not used and five if it is used all the time]

	0	1	2	3	4	5
Paper reporting ()						
Spreadsheets (e.g. Excel / Open Office) ()						
3D Modelling authoring tools (e.g. Autodesk / Bentley) ()						
Open 3D tools (e.g. 3D Repo) ()						
4D Modelling tools (e.g. Synchro / Navisworks / 3D Repo) ()						

Q6 Do you typically perform the health and safety risk assessment of your projects on your own or as part of a team?

- o   On my own (1)
- o   As part of a team (2)
- o   Not applicable (3)

Q7 Ideally, what improvements from the perspective of designing for health and safety would you like to see on projects? (e.g. technical, processual, managerial changes)

_____

_____

_____

_____

_____

End of Block: Demographic information

Start of Block: Video

You will now be shown a video. Please watch carefully.

Video embedded (https://youtu.be/gAhNANwrp2Y)
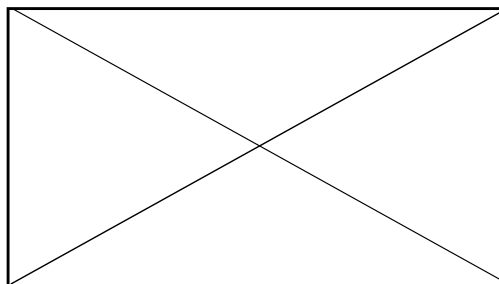
End of Block: Video

Q1 Were you aware of SafetiBase before watching the video?

- o Yes (1)
- o No (2)

Q2 Feature evaluation and assessment

In relation to the Risk Suggestion feature presented in the video that you just watched, please provide your views against the areas highlighted below

	Strongly agree (1)	Somewhat agree (2)	Neither agree nor disagree (3)	Somewhat disagree (4)	Strongly disagree (5)
Supports the selection of appropriate treatments to mitigate H&S risks (1)	o	o	o	o	o
Enables leveraging lessons learnt across previous projects (2)	o	o	o	o	o
Appears easy to use (3)	o	o	o	o	o
Categorising risks is useful (4)	o	o	o	o	o
Seems aligned to industry regulations and guidelines (CDM 2015 / PAS 1192-6) (5)	o	o	o	o	o
Could positively impact design decisions (6)	o	o	o	o	o
Enables a more collaborative environment to perform H&S risk assessments (7)	o	o	o	o	o
Could improve organisational H&S management processes (8)	o	o	o	o	o
Has the potential to improve H&S in construction (9)	o	o	o	o	o

Q3 Do you have any comments in relation to the statements above? (This includes comments that sit outside of the scope of the statements)

_____

Q4 What are the potential challenges to adopt this tool?

_____

Q1 Would your organisation like to engage with future research work in this space?

- o   Yes (1)
- o   No (Why?) (2) _____

Q2 Would you like to engage in the industry Community of Practice group set-up to review and improve this Tool?

- o   Yes (1)
- o   No (2)

Q3 Are you interested in using 3D Repo's SafetiBase in your projects?

- o   Yes (1)
- o   No (2)

Appendix D Interview questions

General questions

- Before the pilot study, what was the typical process and activities that you and your organisation followed to manage safety risks in your projects?
- How did your organisation incorporate 3D Repo's SafetiBase into its day-to-day activities and processes for your pilot project?
- How did other stakeholders in your project facilitate or hinder the implementation of SafetiBase in your pilot project if at all?
- How has the implementation of SafetiBase in your pilot project influenced or changed the safety management process in your organisation?

Questions about the use of the tool

- How has using 3D Repo's SafetiBase to categorise risks (using a scenario/treatment ontology) influenced or not influenced the way in which you think about safety risks in your projects?
- Tell me about your experience using SafetiBase to identify and categorise safety risks.
- Which features of SafetiBase did you find to be very useful?
- Which features of SafetiBase would you suggest should be improved and what would the improvements be?
- Are there any other areas of construction health and safety work that may benefit from digitalisation and use of a data structuration approach?

Closing question

- Is there anything else you'd like to add before we end?

Appendix E   Characterising Risk Scenarios from HSE Archive Data

A Text Inference Approach

Zhou, Zili
zili.zhou@manchester.ac.uk

Osorio-Sandoval, Carlos
carlos.osoriosandoval@manchester.ac.uk

Freitas, Andre
andre.freitas@manchester.ac.uk

Collinge, William
william.collinge@manchester.ac.uk

1st June 2021

Motivation

The construction industry is one of the most hazardous industries worldwide. Prevention through design (PtD) is a concept that encourages designers to identify safety risks that can occur throughout the lifecycle of facilities as early as possible and to propose means to mitigate such risks. The Risk Library (RL) is a database that captures expert knowledge and UK construction safety regulations by mapping construction safety risk scenarios to existing treatment suggestions that can eliminate, reduce, inform or control risks at relevant design phases. The RL is the core of a BIM-based tool that supports designers to select an appropriate treatment upon identification or a related risk scenario within a 3D model of a facility.

In the context of the RL, a risk scenario is characterised by (1) a construction scope that determines the type of construction work being carried out when the risk could arise; (2) a building element associated with the eventuation of the risk; (3) a location that indicates topological characteristics that can be critical and hazardous; (4) an associated activity; (5) a risk category, based on PAS 1192-6 [1], that classifies the type of risk; and (6) a risk factor that identifies the reason behind the risk eventuation.

The construction sector repository of the Health and Safety Executive (HSE), the UK regulator for health and safety, contains incident reports, enforcement activity documents and safety guidance that provide important supporting evidence, such as details of previous incidents, consequences and mitigation strategies [6]. Among these documents, press releases reporting construction accidents and RIDDOR (Reporting of Injuries, Diseases and Dangerous Occurrences Regulations) reports contain sufficient information to extract risk scenarios from real world events. The extracted risk scenarios can be reviewed by construction safety experts to map them to appropriate treatment suggestions to be included into the RL.

With this aim, a Natural Language (NL) based inference approach for risk classification is proposed to be applied on the RIDDOR reports and press releases related to construction accidents. Considering the characterisation of risk scenarios as a predictive problem, textual data from RIDDOR reports and press releases is used as input, the target task is to predict the categories of multiple risk scenario elements, such as location, risk factor, activity, building elements, etc.

Using a state-of-the-art machine learning model based on transformer-based deep neural networks (as a natural language encoder component of NL predictive model) we leverage the high generalisation ability of transformers to implicitly encode construction domain knowledge. Using these embeddings as a

representation support, the model is capable of performing the inference of a risk scenario from text (instantiating these categories).

The BIM Risk Library

Effective implementation of PtD requires appropriate technologies and tools that enable designers to apply their knowledge and skills in safety management while they engage with construction designs [3; 7]. While Building information modelling (BIM) technology has been increasingly adopted by designers in the UK [5], its implementation in the PtD context is still limited.

An existing commercial cloud-based BIM application (3D Repo) developed a feature (the new version of SafetiBase) based on the RL and its risk scenario data structure to implement PtD by enabling users to leverage the knowledge contained in the database within a BIM environment. Such a feature displays treatment suggestions related to risk scenarios upon their identification in the model by designers. Therefore, it is important to continue expanding the RL to include treatment suggestions for as many risk scenarios as possible. Extracting risk scenarios from real world events reported in text corpora is an appropriate means to enable the continuous improvement of the RL and its subsequent deployment within a BIM environment. However, due to the large amount of unstructured text from multiple sources in the construction sector repository of the HSE archive, the manual extraction of risk scenarios is complex and expensive.

RIDDOR and press releases corpus

The archive of the Health and Safety Executive (HSE), the UK regulator for health and safety, contains press releases reporting construction accidents and RIDDOR (Reporting of Injuries, Diseases and Dangerous Occurrences Regulations) reports, which contain sufficient information to extract risk scenarios from real world events. Such risk scenarios can be reviewed by construction safety experts to map them to appropriate treatment suggestions to be included into the RL. A risk scenario is characterised by (1) a risk category based on PAS 1192-6 [1]; (2) a construction scope; (3) an associated building element; (4) a location indicating critical and hazardous topological characteristics; (5) an associated activity; and (6) a risk factor indicating the reason behind the risk eventuation.

RIDDOR reports are rich in descriptions of the contextual factors of incidents [6]. A large data set of desensitised RIDDOR reports classified into categories aligned with the risk category, the associated activity, the risk factor, and the construction scope concepts from the RL provided an opportunity to train and test the NLI-based models presented in this report.

The press releases used in this research are public news reports issued by the HSE in a free text format. They report the conditions under which construction accidents occurred. Press releases contain sufficient information to extract risk scenarios from real world events. Characterising risk scenarios from these two types of data sets enables increasing the amount of entries in the RL.

The BIM Risk Library project benefits from characterising risk scenarios from these two types of datasets for automating the creation of incident instances in the Risk Library, which enables designers to identify risk scenarios based on previous sets of events within their BIM models.

Outline of Contributions

In this work, we explore the automatic categorisation of risk scenarios from text using two reference corpora, RIDDOR (Reporting of Injuries, Diseases and Dangerous Occurrences Regulations) reports and HSE press releases which reports construction accidents. With this aim, a text categorisation approach is proposed, exploring contemporary textual embedding methods, which are applied to the domain of discourse of RIDDOR reports and press releases. Considering this classification problem, text data from RIDDOR reports and press releases are used as input, where the target task is to predict the categories of multiple risk scenario elements, such as location, risk factor, activity, building elements, etc.

In summary, we achieved following objectives in this work:

- To support the automation of the creation of risk scenarios in the BIM Risk Library.

Figure 1: Overview of the system
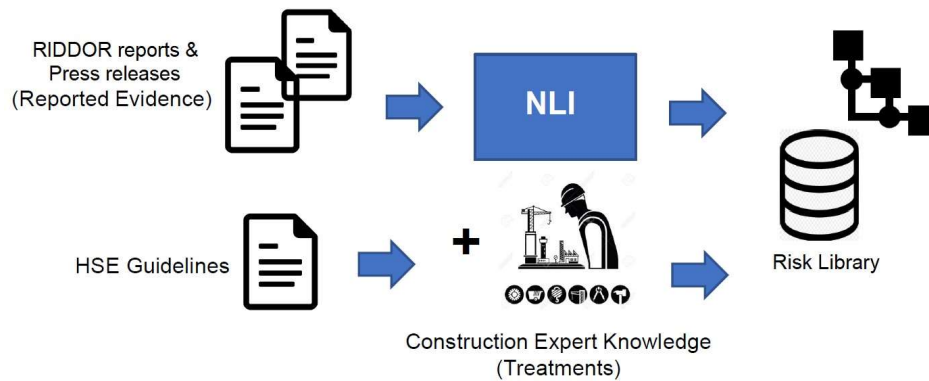
- To determine a categorical system for risk scenario description which balances level of detail and the ability of automated classification.

- To develop classifiers to categorise risk scenarios from text using (press releases and RIDDOR reports).

- To validate the use of contemporary textual embeddings to support text-based inference in the construction domain.

- To validate transfer learning and few-shot learning[1] from RIDDOR reports to press releases.

Figure 2: System Architecture
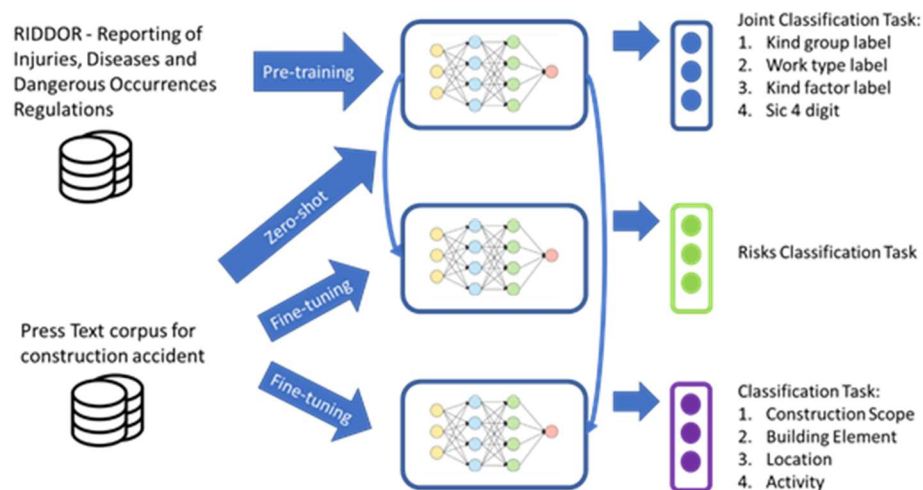
Definition of a risk scenario

Building upon the previous work on the BIM ontology, we defined the core elements of a risk scenario, as the descriptive categories for this work.

A *risk scenario* is characterised by: (1) a construction scope, a concept based on CIRIA C755 CDM 2015 that determines the type of construction work; (2) a building element, a concept that enables classifying the scenario by related building design disciplines; (3) a location relative to the risk, which indicates topological

---

[1] learning from smaller annotation sets

characteristics that can be critical and hazardous; (4) an activity, a concept sub-classified at a high level that determines the stage in the life-cycle of the asset during the occurrence of the risk from preliminary site investigation to material disposal or re-use; (5) a risk category, a concept based on PAS 1192-6 2018 that identifies the type of risk that could occur; and (6) a risk factor that identifies the reason behind the risk eventuation.

The purpose of characterising a *risk scenario* is to inform existing embedded risks in the design phase, based on previous reported incident instances. This identification can enable *treatment suggestions* that can eliminate, reduce, inform or control the risk at relevant design phases. These suggestions can subsequently be prompt to designers using the Risk Treatment Suggestion feature of 3D Repo's SafetiBase to assist them in identifying appropriate treatments to identified risks.

In this document, scenarios are characterised from HSE press releases. Fragments of text leading to the determination of a category for each concept are highlighted using the colour code shown in Table 9. Subsequently, an explanation of the inference of the chosen category is given.

Table 1: Characterisation colour code

Concept	Colour
Construction scope	
Building element	
Location	
Activity	
Risk category	
Risk factor	
Multiple concepts	
Potential false positive	

Where more than one concept can be inferred from the same word or phrase, such phrase will be highlighted as multiple concepts, and a detailed explanation will be provided.

A fragment of text highlighted as potential false positive indicates that, without context, the phrase could lead to incorrectly categorisation. These fragments of text could be used to identify alternative scenarios that are not necessarily present in the referenced press release. An explanation is provided where this situation occurs.

Annotated press release

The text below exemplifies an annotated press release:

'A   based construction company has been fined for safety failings.   s' Court heard how   was working at a site on   which was inspected by the Health and Safety Executive (HSE). The inspector found there was an excavation directly behind the front door, which was the only entrance and exit for the site. The excavation was approximately 3.5 meters deep with no means of shoring to support the sides of the excavation and no edge protection around the top to stop persons, materials or objects falling into it. Of immediate concern was a worker in the bottom of the excavation who was instructed to exit the excavation[...]'

A dataset containing 127 annotated press releases texts was produced.

Transformer-based Architectures

A contemporary Deep Neural Network based architecture was used as the foundation for this analysis: the BERT (Bidirectional Encoder Representations from Transformers) Transformer model [2]. Contemporary textual embeddings allows the unsupervised capture of background knowledge from text as scale, using this encoding of text to support inference. Using deep neural network architectures and with the support of a

multi-head attention mechanism, the BERT model allows for multiple NLP tasks which require inference over textual background knowledge including question answering, textual entailment and text classification.

Table 2: Characterisation and justification

Concept	Category	Justification
Scope	General excavation	There is no lexical evidence that suggests that another category is stronger
Element	Not applicable	There is no reference in the text that may indicate a building element related to the excavation where risks were identified
Location	Site logistics-Excavation area	There was a worker inside the excavation
Activity	Install construction	A reference to a site in this context usually means there are construction activities going on
Risk I	Trapped-Crushed by excavation	The lack of shoring suggest that there is risk of being crushed if a collapse occurs
Risk II	Fall-From open edge	The lack of edge protection around the top indicates that there is risk of falling into the excavation
Risk III	Struck-By falling object	Also a resulting risk of the lack of edge protection around the top
Risk factor	Task-Excavation	In this case, the 3 identified risks can be associated to the same risk factor, since they are a result of this type of work. This is also indicated by the dimensions of the excavation highlighted in the text. Alternative risk factors can be found in Table 11

Table 3: Alternative risk factors

Alternative characterisation	Justification
Risk factor I: Physical-Collapse Risk factor II: Physical-Opening	These risk factors are associated to the identified risks as well

The BERT model can achieve remarkable versatility based on a two step parameter optimization process: first pre-training the BERT model on large-scale general natural language corpus, then fine-tuning the model on specific natural language inference tasks. The pre-training process allows the model to induce a representation from large-scale corpora, which can be later specialised (fine-tuned for a specific task). In this project, the BERT model is used as a text encoder component of the proposed risk scenario classification model.

Methods and Experiments

As shown in Fig. 2, we propose three transformer-based models to deliver the objectives if this project. These three settings are:

- 1) Training a model using a common set of risk scenario categories which are present at RIDDOR reports, aiming for transferring this knowledge to a general risk scenario classifier.

- 2) Leveraging the embedded textual knowledge from RIDDOR reports to support the target risk scenarios classification on press release text, by fine-tuning a pre-trained model in the press releases.

- 3) Evaluate zero-shot learning[2] on the press-release corpus by using the classifiers built in (1).

Joint Classification on RIDDOR Reports

For this first model, named RIDDOR-NLI, four classifiers sharing a mutual core natural language encoder are deployed, each classifier is trained to predict one of the report's categorised features (work kind group, work process, main factor, and sic4 digit).

Given one of four classification pipelines a, as shown in Eq. 1, using as an input a RIDDOR text case seq, a class distribution can be determined. The model uses RoBERTa-base [4] parameters (θ) and an output layer with parameters W_a and β_a. Because each categorised feature in each pipeline consists of multiple classes (more than 2 classes), in the second line of the Eq. 1, we choose to maximize the cross entropy of outputs and tagged labels, where the parameters W,B,θ can be optimized into $\mathcal{W}^R, \mathcal{B}^R, \theta^R$.

$$y_a = W_a f_\theta^{BERT}(seq)[0] + \beta_a$$
$$\max_{\mathcal{W},\mathcal{B},\theta} \sum_{W_a,\beta_a \in \mathcal{W},\mathcal{B}} \log \frac{e^{y_a[l]}}{\sum_j e^{y_a[j]}} \to (\mathcal{W}^R, \mathcal{B}^R, \theta^R)$$

(1)

We implemented the RIDDOR-NLI model, then test it on RIDDOR-51k, a RIDDOR reports dataset containing over 51,000 reports. The RIDDOR-NLI model is trained on 20% reports randomly selected from all the reports in RIDDORs-51k, and the model's performance is evaluated by testing the trained model on the rest 80% reports.

For each input RIDDOR report, the RIDDOR-NLI model outputs a probability for each candidate category. The higher probability indicates a likelihood that this candidate category should be assigned to the text.

Two major measurements are used for performance evaluation, mean accuracy and mean ranking. To calculate mean accuracy, only the candidate category with highest output probability for each feature is considered as predicted category, a correct score is tagged as 1 if the predicted category is correct, otherwise tagged as 0, and the mean value of correct score is calculate as mean accuracy. For the mean ranking evaluation, for each report, all candidate categories are sorted by their probabilities, where the value of correct categories' ranking position is used as an evaluation metric (the lower mean ranking indicates more accurate result).

Table 4: Test on RIDDORs-51k

	Cat#	Mean Accuracy	Mean Accuracy * Cat#	Mean Ranking	Mean Ranking / Cat#
kindgrplabel	16	0.713	11.408	1.597	0.100
work process label	18	0.447	8.046	2.569	0.143
main factor label	15	0.634	9.51	1.936	0.129
sic4 digit label	22	0.335	7.37	3.753	0.171

In Tab. 4, we report the results of testing the RIDDOR-NLI model on the RIDDORs-51k dataset. The second column "Cat#" is the number of candidate values for each category. The accuracy value should be put under the perspective that all categories have at least 15 candidate values (the more categories, the more challenging the classification task). For "sic4 digit label", with 22 candidate values, the mean accuracy is 33.5% and the mean ranking is 3.753 (out of 22 candidate values). This is the worst/case category. On the other side of the spectrum, for "kindgrp label", the mean accuracy reaches 71.3%, and the mean ranking is 1.597 (out of 16 candidate categories).

There are 16 candidate values for the "kindgrp label" category, 18 for "work process label", 15 for "main factor label" and 22 for "sic4 digit label". We normalised the mean accuracy and mean ranking results for each category in order to make them comparable. As shown in the columns "Mean Accuracy * Cat#" and "Mean Ranking / Cat#" in Tab. 4, the model performs best on "kindgrp label" highest normalised mean

---

[2] the ability to generalise over a new dataset

accuracy, 11.408, and lowest normalised mean ranking, 0.100followed by "main factor label", "work process label", and "sic4 digit label".

The score ranges demonstrate that the baseline classification reached an operational level of use, and could be used to scale/up the annotation of new resources. Together with the classes, the classifier delivers a probability distribution which can be used to estimate uncertainty, which cna be used for human validation (when uncertainty is low).

Zero-shot classification for specific attributes on press text

As we have the trained RIDDOR-NLI model parameters $\mathcal{W}^R, \mathcal{B}^R, \theta^R$, we can apply the model directly to predict two categories on the press releases namely, activity and sic4 digit. Pre-processing for mapping the attributes "work kind group" to "activity" are required, as the candidate values in these two categories are not identical.

Table 5: Zero-shot on Press Release

	Cat#	Mean Accuracy	Mean Accuracy * Cat#	Mean Ranking	Mean Ranking / Cat#
activity	9	0.357	3.213	2.690	0.299
sic4 digit-label	22	0.273	6.006	4.221	0.192

We use the same evaluation measurements as in RIDDOR-NLI test on RIDDORs-51k dataset, mean accuracy and mean ranking. As shown in Tab. 5, the "activity" category has 9 candidate values, the zero-shot RIDDOR-NLI model can achieve 35.7% mean accuracy and 2.690 mean ranking (out of 9 candidate values). For "sic4 digit label", with 22 candidate values, mean accuracy is 27.3%, and mean ranking is 4.221.

Similar to the analysis in RIDDOR-NLI test on RIDDORs-51k dataset, we also normalise the mean accuracy and mean ranking results with number of candidate categories. According to the results shown in Tab. 5, the zero-shot model performs better on "sic4 digit label" category with higher normalised mean accuracy, 6.006, and lower normalised mean ranking, 0.192.

This experiment demonstrated that it was possible to transfer the domain knowledge from RIDDORs to another related domain of discourse, press-release, in a zero-shot fashion.

Fine-tuning based classification on press text for risk scenarios extraction

In this part of the work, we set up two risk tasks on press release text, risk and scenario (Construction scope, Building element, Location, Activity) classification.

Risk prediction. For each report, one or more risks should be tagged. In this case, we can form the risk prediction problem as N binary classification problems, N is the number of all possible risk classes. As shown in Eq. 2, we first normalise the output of BERT with the application of sigmoid function (which will output a probability distribution in a range from 0 to 1), then we use binary cross entropy loss function to train the model.

$$y_c = \frac{1}{1 + e^{-(W_c f_{\theta^R}^{BERT}(seq)[0] + \beta_c)}}$$
$$\mathcal{L} = t \ln y_c + (1 - t) \ln (1 - y)$$
$$\max_{W_c, \beta_c, \theta^R} \mathcal{L} \rightarrow (W_c^P, \beta_c^P, \theta^P)$$

(2)

For each input text, the model can output a probability for each risk, which indicates the likelihood that the risk is described in this report. To evaluate the classification model's performance on the risk prediction task, we use as measurement average precision, as the classification task is binary. Average precision aggregates a precision-recall curve as the weighted mean of precision values achieved at each threshold, with the increase in recall from the previous threshold used as a weight.

In "Occ" columns of Tab. 6, we list the occurrence of risks. We also provide average precision results based on the model's two different parameter initialization settings, pure RoBERTa-base and RIDDORs pre-trained parameters.

We report the results of the model with the pure RoBERTa-base initialization in the "RoBERTa" column in Tab. 6. We can also take the natural language encoder of RIDDOR-NLI model pre-trained on RIDDORs51k dataset as an initialization, in this case, we provide the results in the "RIDDORs" columns.

We only report results for three risks categories due to limited number of annotated data, the remaining set of risks occurring in the dataset for less than 14 times.

According to the results in Tab. 6, for all three risks, the average precision result in "RIDDORs" column is significantly better than results in column "RoBERTa". The model with RIDDOR-NLI natural language encoder as initialization outperforms the model initialized by pure RoBERTa, the average precision reach 0.882, 0.902, and 0.571 on three given risks. We can conclude that using natural language model pre-trained on RIDDORs-51k improves the prediction accuracy on selected risks.

Table 6: Prediction for Risk

		Average Precision	
	Occ	RoBERTa	RIDDORs
"Fall-through fragile material"	16	0.663	0.882
"Fall-From open edge"	46	0.693	0.902
"Struck-By falling object"	14	0.211	0.571

Scenarios prediction. Each report in the press release corpus contains for scenario categories: Construction scope, Building element, Location, and Activity, in a multi-class setting.

In this case, as shown in Eq. 3, we can use a similar training strategy as in the RIDDOR-NLI model. There is one distinct classification pipeline for each of four target categories (Construction scope, Building element, Location, and Activity). For each given pipeline b, we train the model with an entropy based loss function to optimize these parameters.

Similarly, we initialised the natural language encoder parameters with trained RIDDOR-NLI model's natural language encoder parameters θ^R instead of pure RoBERTa-base parameters.

$$y_b = W_b f_{\theta^R}^{BERT}(seq)[0] + \beta_b$$
$$\max_{W_b, \beta_b, \theta^R} \log \frac{e^{y_b[l]}}{\sum_j e^{y_b[j]}} \to (W_b^S, \beta_b^S, \theta_b^S)$$

(3)

We show the occurrence of different scenarios in Tab. 7. For "Construction scope" and "Building element", the occurrence distribution of values is balanced. However, for "location", the values "Site logisticsExcavation area", "High Level-Near Opening", and "High Level-Near Edge" have a high occurrence frequency, summing them up, these 3 categories occur 76 times out of 122.

Further, for "Activity", most of values are of "Install construction" type (88 out of 127).

In Tab. 8, we report the results for each category.

Similar to the risk prediction task results, we initialize the BERT based natural language encoder of NLI model with two settings: pure RoBERTa-base and the encoder of the pre-trained RIDDOR-NLI model reported as "RoBERTa" and "RIDDORs" respectively. We also report the normalised score values (columns "MA * Cat#" and "MR / Cat#").

Observing the "Mean Accuracy (MA)" and "Mean Ranking (MR)" columns in Tab. 8, the mean accuracy results of two different initialization settings are practically equivalent, the mean ranking of pre-trained RIDDOR-NLI initialization setting is slightly better than pure RoBERTa-base initialization results. The model performs best on "Construction Scope" category followed by "Activity", "Building Element" and "Location".

Conclusions

Through this study, we explore several models for characterising risk scenarios from HSE data archive. Specifically, we propose three detailed NLP solutions to achieve our major objectives, four experiments are

conducted to determine the effectiveness of proposed solutions. According to the experiment results, we have several conclusions:

Table 7: Occurrence in Press Release Data

Construction scope	115
"D2-Roof coverings and finishes"	17
"C2-In situ concrete"	11
"C9-Refurbishment of existing buildings"	15
"B1-General excavation"	14
"B3-Trenches for foundations and services"	13
Building element	85
"Flat roof"	11
"Slab"	15
"Mechanical equipment"	17
"Temporary structure"	11
Location	122
"High Level-Near Opening"	26
"Site logistics-Confined area"	11
"Site logistics-Excavation area"	31
"Site logistics-Exposed area"	13
"High Level-Near Edge"	19
Activity	127
"Install construction"	88
"Demolition, removal"	14

Table 8: Prediction for Scenarios

	Cat#	Mean Accuracy (MA)		MA * Cat#	Mean Ranking (MR)		MR / Cat#
		RoBERTa	RIDDORs	RIDDORs	RoBERTa	RIDDORs	RIDDORs
Construction Scope	39	0.514	0.514	20.046	1.943	1.857	0.048
Building Element	18	0.538	0.538	9.684	1.962	1.923	0.107
Location	12	0.568	0.568	6.816	2.136	2.054	0.171
Activity	17	0.744	0.769	13.073	1.333	1.282	0.075

- A model trained on RIDDORs-51k effectively performs on the RIDDOR report classification task, with at least over 33.5% mean accuracy for all four categories. The model is most effective on the category "kindgrp label", where the mean accuracy reaches 71.3%, and the mean ranking is 1.597 (out of 16 candidate categories).

- Zero-shot "activity" and "sic4 digit label" predictions on the press release text with the model was demonstrated and proved to be possible across the different texts. Final results for the target categories are 35.7% and 27.3% for "activity" and "sic4 digit label" respectively.

- Pre-training the model encoder component on the RIDDORs-51k dataset can improve the prediction accuracy of a target model on the press release risk prediction task. The model initialised with RIDDORs-51k pre-trained settigns largely outperforms the pure RoBERTa-base initialization. The best performing model achieves average precision of 0.882, 0.902, and 0.571 predicting three given risks, "Fall-through fragile material", "Fall-From open edge", and "Struck-By falling object".

- Finally, we find that pre-training the model on RIDDORs-51k dataset can't achieve major improvements on the scenarios classification dataset. Comparing the accuracy results across different features, the

NLI model performs most effectively on the "Construction Scope"category considering it has 39 candidate values. For this category, the mean accuracy is 51.4%, and the mean ranking is 1.857 (out of 39 candidate categories).

In summary, contemporary ML-based architectures, such as transformer-based models, provide an feasible infrastructure to populate the BIM risk library with HSE textual background knowledge. These can be adapted to similar scenarios within HSE. Performance can significantly vary across different categories, as these are differently expressed in text. The probability scores associated with the classification can provide a practical method to support post-hoc human revision.

References

[1] British Standards Institution. *PAS 1192-6:2018 Specification for collaborative sharing and use of structured Health and Safety information using BIM*. BSI Standards Limited 2018, 2018. ISBN 978 0 580 95567 9.

[2] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *NAACL-HLT*, 2019.

[3] M. A. Hossain, E. L. Abbott, D. K. Chua, T. Q. Nguyen, and Y. M. Goh. Design-forSafety knowledge library for BIM-integrated safety risk reviews. *Automation in Construction*, 94(July):290–302, 2018. ISSN 09265805. doi: 10.1016/j.autcon.2018.07.010. URL https://linkinghub.elsevier.com/retrieve/pii/S0926580518300840.

[4] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov. Roberta: A robustly optimized bert pretraining approach. *ArXiv*, abs/1907.11692, 2019.

[5] NBS. NBS' 10th National BIM Report—NBS, 2020. URL https://www.thenbs.com/knowledge/national-bim-report-2020.

[6] P. Thompson, T. Yates, E. Inan, and S. Ananiadou. Semantic Annotation for Improved Safety in Construction Work. *Proceedings of the 12th Conference on Language Resources and Evaluation*, (May): 1990–1999, 2020.

[7] J. Yuan, X. Li, X. Xiahou, N. Tymvios, Z. Zhou, and Q. Li. Accident prevention through design (PtD): Integration of building information modeling and PtD knowledge base. *Automation in Construction*, 102:86–104, 2019. ISSN 09265805. doi: 10.1016/j.autcon.2019.02.015.

Appendix F   Risk scenarios in the Risk Library (v 11/06/2021)

Risk Category	Risk Location	Element Type	Risk Factor	Construction Scope	Associated Activity	Treatments
Fall-From open edge	High Level-Near Opening	Slab	Physical-Opening	In situ concrete	Install construction	45
Fall-From open edge	High Level-Near Edge	Slab	Physical-Edge	In situ concrete	Install construction	30
Fall-From open edge	High Level-Near Edge	Flat roof	Physical-Edge	In situ concrete	Install construction	17
Fall-From open edge	High Level-Between Joist	Frame/beam	Physical-Spacing	In situ concrete	Install construction	14
Fall-From open edge	High Level-Near Opening - Stairwell	Stair	Physical-Opening	In situ concrete	Install construction	14
Fall-From open edge	High Level-Near Opening - Shaft	Slab	Physical-Opening	In situ concrete	Install construction	8
Fall-From open edge	High Level-Near Edge	Cladding	Task-Lifting	In situ concrete	Install construction	15
Fall-From open edge	High Level-Near Edge	Pitched roof	Physical-Edge	In situ concrete	Install construction	9
Fall-From open edge	High Level-Near Edge	Temporary structure	Task-Temporary structure	In situ concrete	Install construction	10
Trapped-Unintended collapse	Site logistics-Excavation area	Raft foundation	Task-Excavation	General excavation	Install construction	20
Struck-By falling object	Site logistics-Excavation area	Temporary structure	Physical-Collapse	Deep basements and shafts	Install construction	16
Struck-By machinery or part	Site logistics-Excavation area	N/A	Task-Excavation	Deep basements and shafts	Install construction	13
Handling-Mechanical lifting operation	Site logistics-Crane area	Temporary structure	Task-Lifting	General civil engineering, including small works	Install construction	17
Trapped-Crushed by excavation	Site logistics-Excavation area	N/A	Physical-Collapse	General excavation	Install construction	16
Trapped-Crushed by excavation	Site logistics-Excavation area	Mechanical equipment	Physical-Collapse	Trenches for foundations and services	Install construction	20
Event-Fire or explosion	Site logistics-Welding area	Stair	Task-Welding	General steelwork	Install construction	8
Struck-By falling object	High Level-Near Edge	Cladding	Physical-Edge	External Cladding	Install construction	8
Fall-From open edge	High Level-Near Opening - Shaft	Lift	Physical-Opening	Lifts, escalators and auto walks	Install construction	4
Mechanical Effect-Noise	Site work area	Internal Wall	Physical-Noise	Masonry	Install construction	6
Fall-From open edge	High Level-Near Edge	Column	Physical-Edge	In situ concrete	Install construction	8
Trapped-Unintended collapse	Site work area	Slab	Physical-Collapse	In situ concrete	Install construction	11
Trapped-Unintended collapse	High Level-Near Edge	Temporary structure	Physical-Collapse	General civil engineering, including small works	Install construction	11
Struck-By moving vehicle	Site logistics-Traffic route	Temporary structure	Physical-Contact with moving vehicle	Site Layout	Install construction	9

Event-Fire or explosion	Site logistics-Traffic route	UG gas pipe	Physical-Gas release and ignition	Access (onto and within site)	Install construction	5
Fall-From MEWP (Mobile Elevated Working Platform)	Mobile plant-MEWP (Mobile Elevated Working Platform)	Ceiling	Task-MEWP (Mobile Elevated Working Platform)	Electrical services	Maintenance	11
Struck-By falling object	Mobile plant-MEWP (Mobile Elevated Working Platform)	Ceiling	Task-MEWP (Mobile Elevated Working Platform)	Electrical services	Maintenance	10
Struck-By falling object	Site logistics-Crane area	Mechanical equipment	Physical-Weight	Mechanical services	Maintenance	9
Struck-By falling object	Mobile plant-BMU (Building Maintenance Unit)	Window	Task-Cleaning glazing/window	Windows/glazing including windows cleaning	Maintenance	10
Fall-Collapse of BMU (Building Maintenance Unit)	Mobile plant-BMU (Building Maintenance Unit)	Window	Task-Cleaning glazing/window	Windows/glazing including windows cleaning	Maintenance	9
Mechanical Effect-Noise	Site work area-Plant room	Mechanical equipment	Task-Cleaning machinery	Mechanical services	Maintenance	9
Fall-From open edge	High Level-Near Edge	Guard rail	Physical-Edge	External Cladding	Install construction	9

Discovering Safety is a large and ambitious scientific endeavour, funded by the Lloyd's Register Foundation. The programme is being delivered by the Health and Safety Executive (HSE) and their key partners.

This long-term programme of work will seek to better understand the global health and safety landscape, its challenges and how best to access and use data to develop solutions for maximum benefit. It will apply expertise in data science, data analytics, artificial intelligence and machine learning to health and safety contexts, advancing areas such as text mining and natural language processing, which will have important spin-off benefits; and it will seek to make a real difference from the insights, through education, practical improvement initiatives and commercial tools and services.

HSE, through its Science Division, is one of the world's leading providers of health and safety solutions to industry, government and professional bodies.

The main focus of our work is on understanding and reducing health and safety risks. We provide health and safety consultancy, research, specialist training and products to our customers worldwide.

Our long history developing health and safety solutions means that we're well placed to understand the changing industrial, regulatory and societal landscape, and to anticipate future issues.

We employ over 360 scientific, medical and technical specialists, drawing on their wealth of knowledge and experience to deliver evidence-based solutions to our clients. Our work is supported by accredited management systems.

Discovering Safety
HSE Science Division
HSE Science and Research Centre
Harpur Hill
Buxton
Derbyshire
SK17 9JN
UK

W: www.discoveringsafety.com

T: +44 (0)20 3028 2000
E: discoveringsafety@hse.gov.uk