Construction occupational accident analysis performed by a large language model using the bow-tie model

May Shayboun[1], Dimosthenis Kifokeris[2], Christian Koch[1,3], Pontus Wärnestål[1] and Christina Claeson-Jonsson[4]

[1]Halmstad University, Halmstad, Sweden; [2]Chalmers University of Technology, Gothenburg, Sweden; [3]Southern Denmark University, Odense, Denmark; [4]NCC, Gothenburg, Sweden

Abstract

Utilizing accident causation models (ACMs) within a construction organization could address challenges in extracting lessons-learned from registered occupational accident data – like learning bottlenecks and inadequate information sharing. Applying text analytics (like large language models, LLMs) to construction site accident data in conjunction with an ACM could furtherly improve the addressing of such challenges. Therefore, in this study we investigate whether we can improve the automation of the accident data analysis (when demanded by a user) in a construction organization by teaching a LLM to perform accident case analysis mapped on the components of the bow-tie model template. In this, we analyse accident report data obtained from a large contracting company in Sweden by combining emergent abilities of LLMs and the bow-tie model through the in-context learning method. We found that the LLM successfully learned to perform accident report analysis based on a format of in-context learning demonstrations, by effectively categorizing and structuring accident data into threats, barriers, and consequences. The use of in-context learning demonstrated a reduction in output hallucination and maintained consistency in aligning with predefined analysis structures. It was confirmed that generating accident analyses without in-context learning resulted in the model's tendency to fabricate information. Nonetheless, shortcomings were also identified, like challenges with data quality and domain complexity, minor inconsistencies in the LLM output, and an inconclusive efficacy of using the conceived model in proactive accident prevention in practice. Nevertheless, this study's contribution is showcasing that an integration of LLMs and ACMs for learning from registered occupational accident data in construction companies is both feasible and potentially advantageous – however, it needs to be furtherly investigated.

Keywords:

Occupational safety, accident reports, large language models, bow-tie model, contractor

1. Introduction

The construction industry is still facing challenges in reducing the frequency of occupational accidents – especially after a levelling out following a long-time and steady frequency decline (Samuelsson, 2022). To tackle those challenges and enhance occupational safety, it is crucial to learn from past experiences (Guan et al., 2024). Sharing information about lessons learned can empower members in organizations to react and protect themselves and others (Lindberg et al., 2010). However, construction organizations (e.g., contractors) face several knowledge sharing issues (Duryan et al., 2020). Those can include the lack of systemic and consistent knowledge transfer across projects, and on-site laborers being usually less informed on safety news and changes compared to the white-collar employees (Duryan et al., 2020). Learning bottlenecks leading to such knowledge sharing challenges are often related to inadequate conduct of cause analyses and planning for actions and interventions (Drupsteen and Hasle, 2014).

Having an accident learning cycle can address those challenges; Silva et al. (2017) described multiple elements in such a cycle, incl. gathering, recording, analysing and coding information, and establishing operational feedback focusing on applying, disseminating and discussing the learned information. However, recording and analysing accident cases are far from perfect in their effectiveness (Gibb et al., 2014). For example, the Swedish Accident Investigation Authority (Statens haverikommission, SHK) found that disseminating lessons learned was one of the weakest points in accident investigation and prevention, and that timely availability of prevention recommendations is important in taking corrective actions (Lindberg et al., 2010). As an analytical accident tool that can be utilized to tackle the aforementioned weak points, the bow-tie model is a type of accident causation model (ACM) and is used for illustrating and visualizing the interactions between hazards, protective and preventive measures (Kuzucuoğlu et al., 2023). The origins of the bow-tie model go back to 1970s and its utilization for hazard and damage process analyses (Fu et al. 2023). Since then, it has been used in the analytic design of different accident scenarios, as well as active and reactive hazard management (Jacinto and Silva, 2010). The bow-tie model can contribute to learning from accidents through simplifying the cause–effect relationships but still retaining an adequate level of detail in communicating barrier or control mechanisms for each failure pathway – while having a form suitable for the understanding and training of process operators (de Ruijter and Guldenmund, 2016).

At the same time, interest in applying text analytics to construction site accident data has grown following advancements in natural language processing (NLP) (Baek et al., 2021; Wu et al., 2022). However, while this development can improve learning from previous accidents, recent NLP-using accident report analyses have faced multiple limitations – such as the need for extensive manual labelling, and NLP being ambiguous and imprecise in understanding natural language expressions (Shayboun, 2022; Wu et al., 2022). There have been some efforts in expanding the application of text analytics in using knowledge-based responses and graph neural networks, but these are limited in their need for domain- and language-specific logical forms (Wang and El-Gohary, 2023). Nevertheless, the rise of large language models (LLMs) in the field of NLP (Zhu et al., 2023) can potentially provide new possibilities in utilizing text analytics for construction site accident analysis. LLMs display emergent abilities, namely abilities not being present in smaller models and not being able to scale up by extrapolating a scaling law (i.e., consistent performance improvements) from small-scale models (Wei et al., 2022). So, due to their enlarged parameter scale, LLMs seem to be able to perform in-context learning, instruction following, and step by step reasoning – unlike smaller language models like BERT (Zhao et al., 2023). Crucial among those abilities, in-context learning is "a paradigm that allows language models to learn tasks given only a few examples in the form of demonstration" (Dong et al., 2022) while not requiring fine-tuning on downstream tasks except for few task-specific demonstrations and appropriate instructions (Zhu et al., 2023).

This background is leading to our research question: Can we improve the automation of the accident data analysis (when demanded by a user) in a construction organization by teaching a LLM to perform accident case analysis based on the bow-tie model template? To address this question, we analyze data on accident reports and corrective measures by combining the emergent abilities of LLMs and the concept of the bow-tie model – thus trying to take advantage of both in an integrated way. Therefore, we map the company's accident reports into the bow-tie components by demonstrating the analysis through multiple examples in an in-context method. The context and content of this paper is a continuation of a user- and safety-related study within a large contracting company in Sweden (Shayboun et al., 2021). In that study, it was found that linking accumulated accident reports to specific work processes added the most value, while one of the company's senior health and safety (H&S) specialists confirmed that gaining knowledge can be challenged by the timely availability of lessons learned (Shayboun et al., 2021).

The paper continues with a literature review, the description of the research methodology, and the results of our analysis. It then concludes with a discussion and final remarks.

2. Literature review

2.1. LLMs and in-context learning

LLMs are the most recent advancement in information retrieval and have demonstrated capabilities in language understanding and generation without needing fine-tuning on downstream tasks (Zhu et al., 2023). Modern information retrieval systems consist of a retrieval and a ranker stages, and the LLMs' high capacity in text semantics make them more suitable in enhancing information retrieval as rerankers, document annotators, or generators of corresponding queries (Zhao et al., 2023). Previously, the Tf-idf and BM25 algorithms have been criticized for their "conceptual flaw: they work only if there is exact overlap of words between the query and document" (Jurafsky and Martin, 2023; Guo et al., 2022). The implication is that the user writing a query or asking a question needs to guess exactly what wording the writer of the answer might have used; this is called the vocabulary mismatch problem (Jurafsky and Martin 2023). Jurafsky and Martin (2023) proposed that it is more successful to use an approach that can handle synonymy (such as dense embedding through Bidirectional Encoder Representations from Transformers (BERT)) rather than using sparse word count vectors. BERT has been recognized with impressive performance in different language understanding, NLP, and information retrieval tasks (Guo et al., 2022; Wang et al., 2024). Kurtz (2022) tested multiple BERT versions on different Swedish tasks and showed that the performance of each version depends on the task and evaluation criteria; the models showing better results were AI-Sweden BERT-large, KB BERT-large 110k, and BERT-base-Swedish-cased-new.

It could be derived from Kurtz's (2022) study that smaller language models depend on language-specific pre-training data, which makes their adaptation to small languages such as Swedish an expensive and resource-intensive task (Holmström et al., 2023). By comparing language models which were trained specifically on Swedish, such as GPT-SW3 (Ekgren et al., 2022), with GPT-3, BLOOM, OPT, and GPT-NEO/J, which were not explicitly trained on

Swedish but do contain a small percentage of Swedish in their training data, it was found that GPT-3 was better than GPT-SW3 in all functional capabilities (Holmström et al., 2023). This indicates that there is probably no need for pre-training LLMs for a specific set of languages, contrary to smaller language models such as BERT, which were shown to perform better in a monolingual setting (Holmström et al., 2023).

As mentioned in the Introduction, the LLMs' emergent abilities (Wei et al., 2022) are especially advantageous for overcoming computationally expensive fine-tuning efforts. Within those, in-context learning works as the LLMs are provided with a few natural language instructions and/or several task demonstrations of input-label pairs (Zhao et al., 2023; Min et al., 2022). In-context learning is different from prompt learning and few-shot-learning because in it, the demonstration is part of the prompt and is applied directly to pre-trained LLMs without requiring parameter update (Dong et al., 2022). Emergent abilities depend on the scaling of the language model; it has been observed in tests on several downstream NLP tasks that after a critical scale threshold is reached, performance increases to substantially above random (Wei et al., 2022). Nevertheless, in addition to the LLM scaling, the performance gain in in-context learning can be enhanced through the independent specification of the input and label spaces using the right demonstration format (Min et al., 2022) as well as the selection of closest neighbors as in-context examples (Dong et al., 2022).

2.2. LLMs in construction

LLMs are starting to be applied in different areas in the construction domain (Saka et al., 2023; Ghimire et al., 2024). In that vein, it has been claimed that safety (incl. site safety management) as one of the most important and sensitive areas of application, while other potential application areas include automated regulatory compliance in the design phase, as well as risk management in the different construction phases (Saka et al., 2023). Although the spread LLM utilization is still in early stages, the literature has shown a few promising cases of using generative pre-trained transformer (GPT) models for few-shot learning and data augmentation. A relevant example is developing a real-time safety regulation question-answering mechanism by embedding regulation documentation with a small LLM and then retrieving information based on user queries with GPT-4 (Khan et al., 2023). GPT 3.5 was also used for classification, cause identification and summarization of OSHA's highway construction accidents, showing great capabilities in condensing safety knowledge about accident causes (Smetana et al., 2024). Another example concerns construction material lifecycle analysis; based on the ISO 14040 and 14044 standards, academic articles and the CML 2001 assessment method were fed to GPT-3 through feedback-based conversation (Turhan, 2023), while few-shot prompting of GPT was used for an interactive dialogue system for material selection and optimization (Saka et al., 2023). Moreover, GPT-4 was used for construction contract risk assessment based on project contract clauses and an expert assessment knowledge base, which were augmented into an in-context learning for more stable results; few-shot prompting was employed in the form of input-output examples to guide the LLM thinking process (Wong et al., 2024). In this approach, subtle human involvement in accurate risk identification, carefully providing demonstrations, and emulating the experts' thinking pattern through the similarity rate between the case clause and the risk clause, were highlighted (Wong et al., 2024).

Although LLMs seem to be promising when used in construction-related tasks, there are recognizable challenges associated with using GPTs – such as hallucinations (which, in the case of construction safety, could cause accidents if one only relies on GPT-generated information (Saka et al., 2023)), as well as dataset biases or unethical data use. Those might be mitigated by the right domain-specific knowledge being represented and integrated in the GPT models, rather than relying on the general data used for training them (Zhu et al., 2023). There is also a need for tangible demonstration of GPT models through robust validation to foster trust and acceptability (Saka et al., 2023; Ghimire et al., 2024). In addition to the aforementioned concerns, challenges include the availability of structured high-quality data for fine-tuning GPTs in order to understand and integrate construction domain knowledge, confidentiality, and new skillsets required for deployment in the industry (Ashkenazi et al., 2023). There have been some efforts presenting a framework for creating large generative models (LGMs) in the construction industry, including data collection and curation, and extensive evaluation of LGMs by domain experts in terms of semantic coherence, grammar, terminology, and validity of generated outputs (Taiwo et al., 2024). Such case evaluations when using GPT-4 for the enhancement of expert systems' knowledge graphs, showed that assessing the correctness of the curated information in terms of relevance, consistency, and completeness, is challenging (Ashkenazi et al., 2023). In another case, ChatGPT v3.5 was tested in creating a construction project schedule and evaluated by experts in terms of accuracy, efficiency, clarity, coherence, reliability, relevance, consistency, scalability, and adaptability (Prieto et al., 2023). The case study showed that ChatGPT v3.5 provided logical yet linear breakdown of project scheduling tasks, with some errors (like incorrect tasks) that should have been included (Prieto et al., 2023). GPT-4 was then used in information retrieval and user queries of construction documents, and this system was evaluated with experts in terms of answering ability, truthfulness quality, relevance, and reproducibility (Taiwo et al., 2024). The limitations were found to be in the chunking strategies, and the used semantic search techniques were unable to adequately link some complex questions to supporting evidence in the contract document (Taiwo et al., 2024).

3. Research method

The conceptualized system for learning from accidents consists of two components, namely accident case retrieval and bow-tie accident case analysis. This is called a passive reader approach, and it generates answers to the user queries by supplying retrieved documents from information retrieval systems; those are then used as inputs to LLMs for creating passages (Zhu et al., 2023). So, the first system component involves the retrieval of accident cases based on user queries; a detailed description of this component is described in section 3.4. This step provides a collection of accident cases that are related to a certain topic as an input to the LLM; we then use GPT-4o API[1] for the analysis within the second system component (section 3.3). We finally prompt GPT-4o with in-context learning demonstrations that are guiding the output of the generated text. Fig. 1 offers an overview of this process.

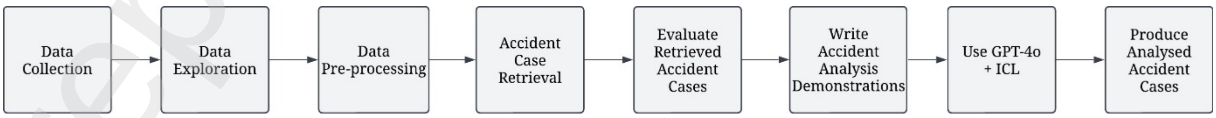


Fig. 1. Process of learning from accidents system

[1] https://platform.openai.com/docs/models/gpt-4o-mini#4ofootnote

3.1. Data collection, exploration and pre-processing

The data was collected in 2020-2021 through a digital accident reporting system used by a large contracting company in Sweden. The collected data comprised of about 3600 accident reports covering the period between 2014 and 2020 and containing different attributes (Table 1).

Table 1. Types and number of existing attributes in the collected dataset

Categorical attributes	Ordinal attributes	Numeric attributes	Textual attributes	Dates
122	15	11	16	9

The data understanding and exploration step is carried out for validating the data quality; missing values, bias, and inconsistent units are targeted (Bruce, 2016; Lau et al., 2023). Specifically, the data, which was organized in a tabular format, was investigated in terms of the number of existing, missing and unique values in each column. Using the results of this investigation, bar charts were created for every feature to assist the analysis. Further manual analysis included documenting a description of the columns' content, taking notes, and making decisions about whether the feature was suitable to be used in the model. The selection criteria for the data were based on quality (e.g., excluding columns with a high percentage of missing values or being empty), their utility for informing causes, circumstances before the accident happened, and prevention measures (excluding non-informative columns that include, e.g., project names), and data sensitivity (excluding personal information). It was found that the most important data quality issues involved the entries offering a general categorization of "Work process", "External factor that influenced the incident", "Work environment" and "Others" – but these features were initially kept in order to re-evaluate their use in the search and analysis of accident cases. Moreover, it was found that the values were not mutually exclusive; after validation by an expert from the company, it was concluded that these values were a result of the reporter assigning more than one value for the same feature and accident case. The data quality issues were validated with two experts in the H&S organization from the contracting company.

3.2. Accident report retrieval

The first step in the case retrieval system is to pre-process the data. The accident cases had multiple free-text entries, incl. the case title and description, cause description and comments, and action description. As such, during pre-processing, we investigated those entries and collected the data parts we decided to use in the model. For some cases the free-text descriptions were repeated in different columns; thus, we checked the textual data for duplication. The duplications were discarded, and the textual data were merged into one continuous text description for the respective case. Then the selected data – as described in the data pre-processing – forms the documents that we use in the retrieval system.

Following the data pre-processing and the formulation of the accident case documents, the accident document retrieval was performed by matching the free text query with the documents' case-text through similarity score and matching the set of filters to their corresponding documents in the dataset (Fig. 2).

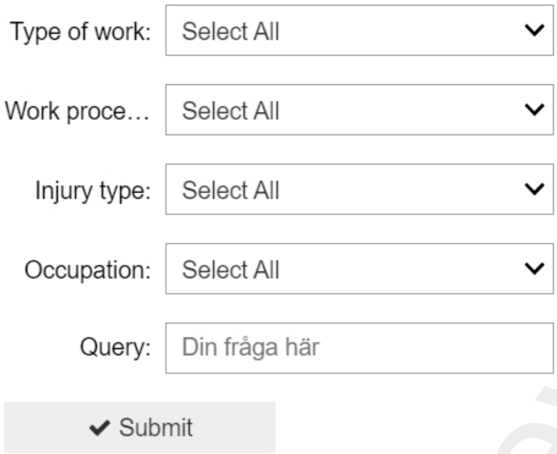


Fig. 2. Data retrieval query and filters.

For this, we used a Swedish BERT model that was trained on approximately 15-20 GB of text (200M sentences, 3000M tokens) from various sources (books, news, government publications, Swedish Wikipedia and internet forums) aiming at providing a representative model for Swedish text[2]. The use of Swedish BERT is motivated by the fact that it is better suited in monolingual settings (in our case, the Swedish language), as mentioned earlier (Holmström et al., 2023). So, we separately tokenized the query and the documents as a first step for processing and followed a bi-encoder approach for the query and each document (see Fig. 3). We then encoded the text into vectors, by using one vector for the query and one for each document. The retrieval is finalized by calculating the cosine similarity score (Jurafsky and Martin, 2024) between the query and document vectors, ranking documents based on their cosine similarity from the highest to lowest, and extracting the documents with the highest score. For this study, we delimit the extraction to the 5 documents with the highest score.

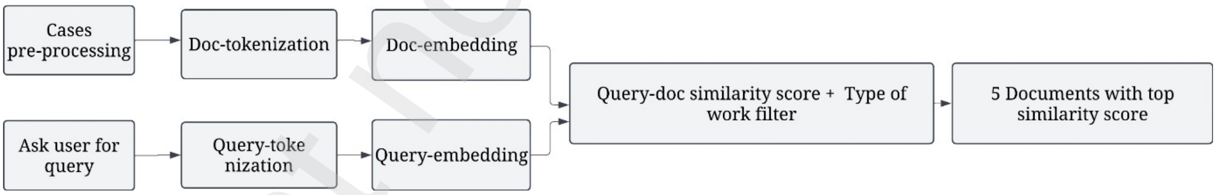


Fig. 3. Accident reports retrieval process

3.3. Formulating in-context learning demonstrations

In this step, we used the bow-tie model (Fig. 4) as a template for analyzing the retrieved accident reports for the in-context learning demonstrations. The demonstrations constitute examples of the analysis template that the LLM should be able to reproduce (Zhu et al., 2023). The bow-tie model itself can be visualized in a structure that resembles a bow and consists of five key elements: the hazard, threats, top events, barriers, and consequences; those elements remain consistent across variations of the bow-tie model (de Ruijter and Guldenmund, 2016; Fu et al., 2020). The top event can be defined as the common node where control was lost and the event

[2] https://huggingface.co/KB/bert-base-swedish-cased

happens just before the various final consequences occur (de Ruijter and Guldenmund, 2016). Linked to the top event is the hazard, which can be described as something in, around, or part of the organization, that has the potential to cause damage (Fu et al., 2020). Threats (the causes of the top event) and consequences (the results of the top event) extend on both sides of the top event (de Ruijter and Guldenmund, 2016; Fu et al., 2020). Barriers can also be included on both sides of the top event; the preventive barriers should stop threats from resulting in the top event or even occurring at all, while the recovery barriers should reduce or completely stop consequences from happening (Fu et al., 2020). However, prevention barriers can still fail; whatever causes a preventive barrier to fail is described as an escalation barrier (Fu et al., 2020).

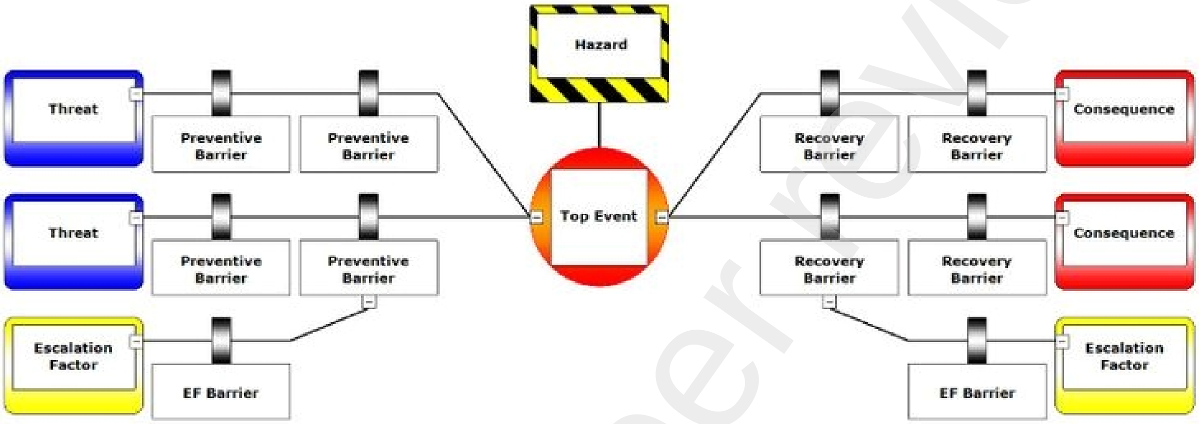


Fig. 4. The bow-tie model (Fu et al., 2020)

Using the aforementioned definitions of the bow-tie components, we formulated two sets of in-context learning demonstrations, as illustrated in Fig. 5 below. We selected two and three relevant cases for queries 1 and 2 respectively, and provided GPT-4o with these queries, their corresponding cases, and their analyses. The accident data could not be mapped on all the bow-tie components; therefore, the respective analyses were reduced to being focused on the components shown in Fig. 5. In writing up the demonstration the hazards were not explicitly mentioned in the text, so we deduced those from the accident case documents. For the rest of the analysis, we were intentional in extracting text parts from the accident descriptions with minimal rephrasing, except for using the profession as a subject, as well as using the categorical causes besides the parts mentioned in the text descriptions. The logic behind choosing this structure of writing the demonstrations was to keep the integrity of the data and simplify the evaluation of the LLM output.

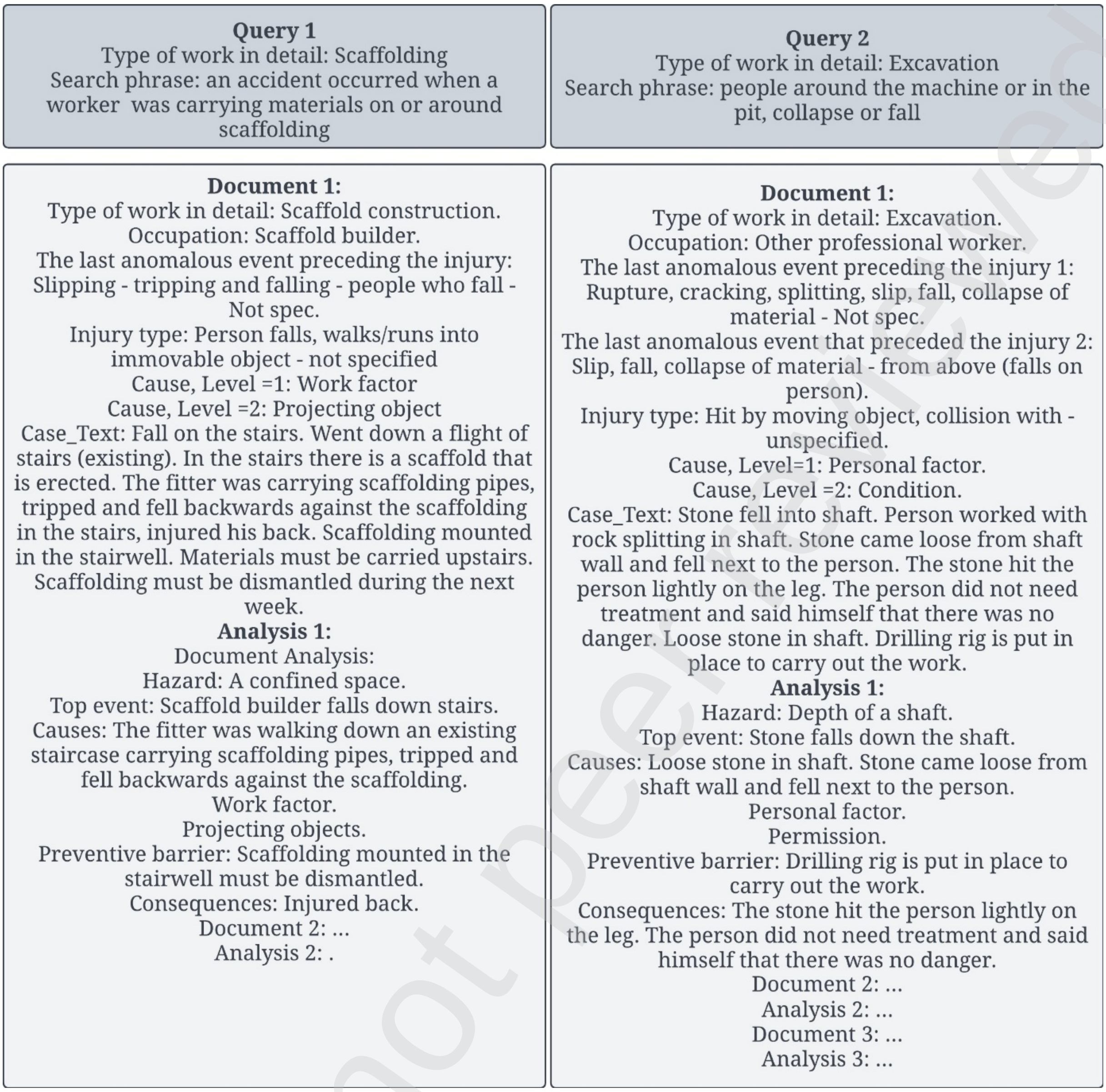


Fig. 5. The utilized in-context learning demonstrations fed to LLM

3.4 Experimental setup

In this step, we used GPT-4o with two system instructions and user prompt settings. For both conditions we instructed the LLM to select the accident cases that were relevant to the query after BERT has retrieved 10 documents. Then, for the first system instruction and user prompt setting, we used the in-context learning demonstrations to train GPT-4o to replicate the analysis of accident reporting documents. For the second setting, the GPT-4o is prompted to analyze the same accident cases without an in-context learning demonstration. We then compared the output in the two conditions and discussed whether there seemed to be potential benefits of using in-context learning in this application area.

System instructions: you are going to be provided with retrieved accident documents and their corresponding analysis examples, and your task is to select the documents relevant to the query and analyze them one by one similarly to the provided document analysis examples.
User prompt: use this analysis example: 'ICL demonstration'. to analyze the following case in the same way: 'Extracted documents').

System instructions: you are going to be provided with retrieved accident documents, and your task is to select the documents relevant to the query and analyze them one by one based on the Bow-tie model.
User prompt: Analyze the following cases: 'Extracted documents'

Fig. 6. Two alternatives for system instructions and user prompt settings: with in-context learning (above), and without in-context learning (below)

We then extracted three sets of 10 documents, each related to a query (see Fig. 7 below), and evaluated the GPT-4o response based on its selection of relevant accident cases to the query, output hallucination, and output validity. Hallucination is the LLM's tendency to generate convincing but false output (Ghimire et al., 2024; Saka et al., 2023), and validity here refers to three interconnected concerns: the factual accuracy of the generated output (Taiwo et al., 2024), whether the GPT responses capture the bow-tie analysis template we provided in the demonstration, and whether the right accident information was assigned to the right bow-tie component in patterns akin to our provided in-context learning demonstrations. For validation, the three extracted three sets of documents were evaluated in terms of LLM output on each of the latter criteria.

Query 1
Type of work in detail: Demolition work
Search phrase: Injuries due to material hazardous to health such as dust, quartz and asbestos

Document 1:
Type of work in detail: Demolition work.
Position: Land/Road O Construction Worker.
Cause, Level=1: Unfocused/distracted.
Cause, Level =2: Other Action (MUST BE DESCRIBED IN A TEXT FIELD).
Case_Text: A worker got something (glass fiber, according to the doctor) in the eye during the demolition of an old district heating culvert. Used safety glasses. The safety glasses not tight fitting. When tearing down a district heating line, existing insulating material dusted so that the worker probably got fragments from this in the eye. Tight-fitting glasses. Tighter safety glasses.
GPT response:
Hazard: Material hazardous to health (glass fiber in this case).
Top event: Worker got fiberglass in the eye during demolition work.
Causes:
- Used safety glasses, but they were not tight-fitting.
- Existing insulation material dusted during demolition of district heating pipe, which probably caused fragments in the eye.
Preventive barrier: Tighter safety glasses.
Consequences: Fragments of fiberglass in the eye, which led to possible eye irritation or damage.

Query 2
Type of work in detail: Earthwork and foundation
Search phrase: Accidents caused by a vehicle hitting someone

Document 1:
Type of work in detail: Earthwork and foundation
Occupation: Other professional worker
Injury type: Person falls, walks/drives into immovable object - not specified
Cause, Level=1 : Lack of barriers/signs
Cause, Level=2: Other personal factor (MUST BE DESCRIBED IN A TEXT FIELD)
Case_Text: Parked car without handbrake. Car was parked on the road behind shutdown. Professionals did not apply the handbrake and see from a distance that the car starts to roll. Running into the car, he opens the driver's door and stumbles. Lands with the leg in front of the rear wheel and slides with the vehicle. The vehicle and professional employee roll out of the shutdown and into a busy lane before the car stops. Truck stopped traffic on the trail so nothing else happened. Carelessness. No handbrake engaged, skidded when trying to stop the car, unlucky that there was an opportunity to go into traffic, great help from a helpful truck driver who stopped the traffic. Ensure handbrake is engaged.
GPT response:
Hazard: Parked vehicle without handbrake.
Top event: Professional employee tripped and ended up in front of a moving car.
Causes:
Professionals did not apply the handbrake.
Carelessness.
No handbrake engaged when parking.
Insufficient barrier.
Other personal factor.
Preventive barrier: Ensure handbrake is engaged.
Consequences: Professionals roll the vehicle into a busy lane, but no further damage thanks to the assistance of the truck driver.

Query 3
Type of work in detail: Other prefabricated articles
Search phrase: accidents or injuries when using lifting devices or lifting equipment.

Document 1:
Type of work in detail: Other prefabricated articles
Occupation: Blacksmith
Injury type: Crushing, crushing - middle
Cause, Level =1 : Improper lifting
Cause, Level =2: Other personal factor (MUST BE DESCRIBED IN A TEXT FIELD)
Case_Text: Was pinched when lifting. Lifted steel into place and beam tipped on forks and a worker pinched his finger. Accidents at work simply. Beam lifted and finger pinched. The worker got pinched during assembly of steel and broke his finger. Talked to the worker and went over the situation, they now know that they should not have their hand between the beam and the edge of the joist.
GPT response:
Hazard: Lifting device.
Top event: Balk tipped on forks and a worker pinched his finger.
Causes:
Improper lift.
Other personal factor (details must be described in the text field).
Preventive barrier: Inform and train personnel not to have their hand between the beam and the edge of the joist when lifting.
Consequences: A worker pinched while mounting steel and broke finger.

Fig. 7. GPT response with and without in-context learning.

Query 1
Type of work in detail: Demolition work
Search phrase: Injuries due to material hazardous to health such as dust, quartz and asbestos

Document 1:
Type of work in detail: Demolition work.
Position: Land/Road O Construction Worker.
Cause, Level=1: Unfocused/distracted.
Cause, Level =2: Other Action (MUST BE DESCRIBED IN A TEXT FIELD).
Case_Text: A worker got something (glass fiber, according to the doctor) in the eye during the demolition of an old district heating culvert. Used safety glasses. The safety glasses not tight fitting. When tearing down a district heating line, existing insulating material dusted so that the worker probably got fragments from this in the eye. Tight-fitting glasses. Tighter safety glasses.
GPT response without ICL
Hazard: Exposure to dust and fiberglass during demolition work.
Top Event: Fiberglass entering the worker's eye.
Threats:
- Improper sealing of protective goggles.
- Inadequate containment of dust.

Preventive Barriers:
- Use of tightly sealed protective goggles.
- Implementation of dust control measures such as wetting materials or using dust extractors.

Consequences:
- Eye injury requiring medical attention.

Mitigative Barriers:
- Immediate medical intervention.
- Enhanced training on proper use of personal protective equipment (PPE).

Query 2
Type of work in detail: Earthwork and foundation
Search phrase: Accidents caused by a vehicle hitting someone

Document 1:
Type of work in detail: Earthwork and foundation
Occupation: Other professional worker
Injury type: Person falls, walks/drives into immovable object - not specified
Cause, Level=1 : Lack of barriers/signs
Cause, Level=2: Other personal factor (MUST BE DESCRIBED IN A TEXT FIELD)
Case_Text: Parked car without handbrake. Car was parked on the road behind shutdown. Professionals did not apply the handbrake and see from a distance that the car starts to roll. Running into the car, he opens the driver's door and stumbles. Lands with the leg in front of the rear wheel and slides with the vehicle. The vehicle and professional employee roll out of the shutdown and into a busy lane before the car stops. Truck stopped traffic on the trail so nothing else happened. Carelessness. No handbrake engaged, skidded when trying to stop the car, unlucky that there was an opportunity to go into traffic, great help from a helpful truck driver who stopped the traffic. Ensure handbrake is engaged.
GPT response without ICL
Hazard: Construction site activities involving parked vehicles.
Top Event: Worker falls and vehicle rolls into traffic.
Threats:
1. Improper securing of parked vehicles (handbrake not engaged).
2. Site layout allowing parked vehicles to potentially roll into traffic areas.

Preventive Barriers:
1. Strict protocols for vehicle parking, including mandatory brake engagement.
2. Regular checks by supervisors to ensure compliance with parking protocols.

Escalation Factors:
1. Forgetfulness or negligence from workers.
2. Slippery conditions increasing the risk of losing control while attempting to stop the vehicle.

Consequences:
1. Worker injury from falling and potential crushing.
2. Vehicle rolling into active traffic lanes, increasing the risk of broader traffic accidents and related injuries.

Mitigation Barriers:
1. On-site staff intervention to stop traffic and prevent further harm.
2. Training and reinforcement to the workers on the importance of securing vehicles.

Recovery Measures:
1. Immediate medical assistance for the injured worker.
2. Traffic management to clear any risk caused by the rolling vehicle.
3. Reviewing and reinforcing parking and vehicle securing procedures among site workers.

Query 3
Type of work in detail: Other prefabricated articles
Search phrase: accidents or injuries when using lifting devices or lifting equipment.

Document 1:
Type of work in detail: Other prefabricated articles
Occupation: Blacksmith
Injury type: Crushing, crushing - middle
Cause, Level =1 : Improper lifting
Cause, Level =2: Other personal factor (MUST BE DESCRIBED IN A TEXT FIELD)
Case_Text: Was pinched when lifting. Lifted steel into place and beam tipped on forks and a worker pinched his finger. Accidents at work simply. Beam lifted and finger pinched. The worker got pinched during assembly of steel and broke his finger. Talked to the worker and went over the situation, they now know that they should not have their hand between the beam and the edge of the joist.
GPT response without ICL
Hazard: Lifting and positioning steel beams.
Threats:
- Improper lifting technique.
- Insufficient communication of safety procedures.

Preventive Measures:
- Proper training on lifting techniques.
- Clear communication and enforcement of safety protocols.

Top Event: Beam tipping during lifting.
Mitigation Actions:
- Immediate medical attention to the injured.
- Review and reinforce safety training and procedures to avoid placing hands between heavy objects and surfaces.
- Monitoring adherence to proper lifting protocols.

Fig. 8. GPT response without in-context learning.

4. Results

In this section we indicatively showcase three queries and select one document analysis for each of the queries to demonstrate samples of the GPT-4o output.

4.1 Accident case retrieval

The first step of the system process, namely the retrieval of accident cases, showed interesting results when it comes to the relevance of the retrieved cases and the utility of the search filters. When using all the filters and a query, the retrieval system was likely to not extract any cases. This indicates that the filters can cancel each other and is consistent with our data pre-processing observations of skewed tails of the features as they are often populated with only a few instances of the rest of the unique values of the same feature. There are also missing values which hinder the use of the filters effectively. Gradually eliminating filters showed better retrieval. Moreover, it was mostly possible to determine the case relevance by reading the case description rather than through the case categorical attributes due to their latter's generalized nature. By using BERT's information retrieval, it was quickly noticed that not all retrieved cases were relevant to the query; three and six out of 10 retrieved cases were relevant to first and second query, respectively.

The results of the retrieved cases were structured in the format shown in Fig. 5. After looking at those, we found that missing values can compromise the data's specificity and consistency – therefore potentially making it less useful for information retrieval. This finding resulted in the decision to exclude more data. On the other hand, parts of the categorical data that were sourced from the ready drop-down categorization were found to be especially helpful in formulating the in-context demonstrations. Those parts were "The last aberrant event that preceded the injury", "Injury type", "Cause", and "Occupation". Those were selected in an iterative process of examining the retrieved cases while formulating the in-context learning demonstration for use in the bow-tie model; we then assigned the respective parts of the accident description text to the corresponding components of the bow-tie model.

4.2 Relevant data and case selection

The cases selected by GPT-4o almost matched our selection of the cases that are relevant to the query. In other words, the LLM never missed a case which was previously labeled as relevant by the researcher. However, the GPT also selected an extra document that we did not consider relevant. The reason behind this is unclear, as the case involved moving baskets and a pole being detached from its mount, but the case description itself is not clearly mentioning lifting devices or lifting equipment (see Query 3, Fig 7). Thus, GPT-4o showcased an overall good performance; however there still were challenges in this step of the process. It is observed that there is some room for subjectivity in the selection process. For example, in the second query, there were two cases of vehicles colliding with passing animals; those were eventually not chosen to be analyzed, because both ours and the LLM's selection approach followed the query precisely. These cases could be argued to be relevant to the query, but the GPT-4o did not select them because they did not involve hitting a *person* with a vehicle. Similarly, in the third query, it is not straightforward to determine which cases are relevant. For example, one document mentioned the use of forks for carrying element support beams in the future as a preventive action – however, the accident had been not caused by lifting tools, but rather the lack of them.

4.2 Hallucinations

Overall, we very rarely observed hallucinations within these three experiments. The LLM performed very satisfactorily when it came to adhering to the accident case data. There was no alarming fabrication of information that did not exist in the accident descriptions, or the documents provided to GPT-4o to analyze. However, a few deviations were indeed observed. Specifically, the LLM made up a word to describe the type of worker involved in the accident, specifically, instead of using the case description's term "Blue collar or civil engineering worker" (in Swedish: "Yrkesarbetare eller Mark/Väg o Anl. Arbetare"), it instead invented a non-existent word that could be described as "working worker" (in Swedish: "Arbetsarbetare"). Another instance involved rephrasing the text, but without any fabrication of false or different information in the causes of the first query (see Fig. 7). Again, in query 2, it can be observed that there is rephrasing of the mentioned cause as "Insufficient barrier" instead of "Lack of barriers/signs" – but we deemed that the difference between those two was not semantically significant. In the responses to query 3, there was also very little hallucination to be observed, except for the term "Inform and train personnel" in the preventive barrier, while training was not mentioned in the accident case text.

Nevertheless, by looking at the generated analysis without the in-context learning setting (see highlighted text in Fig. 7), it can be observed that the plausible but incorrect analyses increase consistently for the three accident reports compared to the generated analysis with in-context learning. We could even detect some pattern of generating recommendations related to training in using proper personal protective equipment or proper procedures, as well as immediate medical attention for the injured worker. Moreover, by looking the generated text content, we find that there is a probability that the LLM response is making assumptions about the threats/causes and is generating output such as the "site layout allowing parked vehicles to potentially roll into traffic areas" and "insufficient communication of safety procedures." Overall, the analysis shows that the in-context learning demonstrations helped GPT-4o reduce hallucinations and mostly use the accident descriptions for the analysis without fabrication.

4.3 Validity of output

By looking at the accident documents analyzed by GPT-4o and in-context learning, we found that the output follows the same bow-tie concepts and structure designed in the demonstrations. We also found that the output follows a very similar pattern of copying the text part corresponding to the respective bow-tie component. Nevertheless, minor inconsistencies were spotted. By looking at the causes in Fig. 7, we can notice that the LLM ignored causes mentioned in the first and second query documents – specifically "Other Action (MUST BE DESCRIBED IN A TEXT FIELD)" and "Other personal factor (MUST BE DESCRIBED IN A TEXT FIELD)" – but decided to add them in the third query. This observation showcases a general issue with using GPT, namely its tendency to be inconsistent.

It is also observed that although the LLM response adheres to the data provided in the document, it sometimes excludes potentially important text parts. In query 3 the analysis does not include that it was a blacksmith who had the accident. In query 2 we also can see that the details of how the worker came to roll with the vehicle are excluded from the analysis. Similarly in another document, the case text mentions that the prevention measure was discussed in the weekly meeting, but this information was omitted in the analysis which might be important indicator for the importance of the incident. This observation might be related to the overall observation that the GPT response seems to be less likely to include the event description under the causes category as we provided in the demonstrations, and this might be influenced by that the causes are explicitly mentioned in the documents, and it looks that their presence override what the GPT learns from the demonstrations. This shows that the model can be very sensitive to the fed data and the demonstrations.

Comparing the GPT responses with and without the ICL demonstrations, we can see that the output tends to be inconsistent with one bow-tie categorization. In query 3, the model generates an output with escalation factors while it does not in the other two queries. The response also shows that the model use "mitigative barriers", "mitigation action", and "recovery barrier" interchangeably. While in query 3, the GPT response does not include any consequences. This might be because the bow-tie model has multiple variations. This also shows that the ICL demonstrations help the GPT response to be more consistent to follow a certain format.

5. Discussion

The findings from our study demonstrate that LLMs can be effectively combined with bow-tie analysis. By leveraging In-context learning capabilities of GPT-4o, we observed that the LLMs provided sufficient accident analysis and adhered to the structure of bow-tie components in the hand-crafted demonstrations. This showcase potentials for multiple use cases where the same approach could be applied for different purposes that support safety processes and safety personnel in the construction industry. This approach could be used with other accident analysis models or other use cases for improving data quality or automatic fill of accident reports into predefined categories. The experiments also showcased that LLM-based systems could be successfully used in enhancing the selection process in IR systems. We have seen that that the GPT is able to do very similar choices compared to the researcher selection. However, there are limitations related to subjectivity and lack of sufficient data to include or exclude cases when compared to the query.

For validation, we focused on the metrics of hallucinations and validity of the generated output. Compared to the generated output without the ICL demonstrations we have seen the ICL combination with the bow-tie format reduce hallucinations and maintain the validity of the output. We find that the ICL demonstrations were contributing to guiding the model to generate accurate and contextually appropriate accident analysis. This is consistent with the literature on the helpful conditions for ICL to work effectively which are the formatting of demonstrations (Min et al., 2022), and selecting closest neighbors as in-context examples (Dong et al. 2022). Without ICL, the model exhibited a higher tendency of output fabrication and inconsistency, reaffirming the importance of demonstrating structured examples for a more reliable GPT output.

With this concept of a system, we aim to assist in the improvement of occupational safety in the context of the contracting company, by making the search and analysis process of accident reports more consistent. Nevertheless, we cannot make claims about the utility of the proposed system in terms of recommending prevention and/or control actions – the data used in this study showed that that prevention-related information was mentioned only briefly, while recommendations of potential implementation were almost completely lacking. This is especially critical, because it follows an already documented pattern in which accident investigation models are lacking in designing and implementing recommendations (Lundberg et al., 2009). While this may imply that recommendations and their implementation could be derived from the analysis (Lundberg et al., 2009), it also makes it hard to establish the usefulness of providing accident analysis information as a form of a proactive prevention tool. Therefore, future research should investigate whether providing safety professionals with accident information can influence accident prevention proactively.

While the results are promising, several limitations must be acknowledged. Firstly, the data quality issues (e.g., data entries with missing values and inconsistencies), significantly impacted the model's performance. Although pre-processing improved some issues, complete reliance on textual data could still limit the model's efficacy when faced with ambiguous or poorly reported accidents. The downside of the pre-processing is that large parts of the data were excluded and deemed not useful for the proposed system. Secondly, complex realities of construction site incidents might not always be fully captured by NLP algorithms, which can occasionally lead to marginally relevant or overlooked cases in the retrieval process. This was clearly shown by the need to re-select relevant cases after employing BERT for retrieving accident cases. Moreover, the textual accident descriptions summarized background information without enough details about how and why the events occurred. This needs to be corrected; a potential measure would be to incorporate more comprehensive and detailed data. Additionally, the observed minor inconsistencies in the bow-tie analysis, such as occasional omission of important event details, indicate that there is room for improving the model's sensitivity and context awareness.

Regardless of the limitations, the system shows efficacy in categorizing and structuring accident data into threats, barriers, and consequences which can simplify the complexity often associated with accident causation analysis. This indicates potential scalability and adaptability of LLMs in learning new types of accident scenarios and suggests that such systems can evolve with increasing data – thus making them a long-term asset for construction firms focusing on health and safety. Future research should focus on expanding the dataset and improving its quality through rigorous validation and inclusion of more detailed and diverse accident reports, which can potentially provide a broader foundation for the LLM's learning. Furthermore, the integration of feedback loops where the model's analysis is continuously reviewed by human experts, should also be investigated. This could ensure that the LLMs remain updated with user preference.

6. Conclusions

In summary, this study highlights the potential of combining LLMs (and specifically, a customization of GPT-4o) with bow-tie analysis for processing and learning from past accident reports in the context of a contracting company in Sweden. The LLM successfully learned to perform accident report analysis based on a format shown to the model through in-context learning demonstrations. The use of in-context learning demonstrated a reduction in output hallucination and maintained consistency in aligning with predefined analysis structures. The experiment also confirmed that generating accident analyses without the in-context learning demonstrations resulted in the model's tendency to fabricate information that, while not being necessarily false, did not exist in the data. The use of GPT also showed to enhance the retrieval of cases from the reported accidents' database through adding it as an extra step after the similarity score BERT retriever.

This research sets a promising direction for the use of large language models (incl. generative pre-trained transformers) in conjunction with accident causation models (specifically, the bow-tie model) in improving occupational safety within the construction industry. Nonetheless, challenges remain, particularly related to data quality and domain complexity. Moreover, it is not possible to make conclusions about the efficacy of using the developed concept of a system in proactive accident prevention in practice, which indicates future research needs. Data quality issues, such as missing values, inconsistent entries and the format of reporting, resulted in excluding large parts of the data – which hindered the enhancement of textual data with details about accidents. Moreover, by looking critically at the GPT output, minor inconsistencies were detected, indicating that there is room for improving the model's performance.

Despite limitations, the conceptualized system managed to effectively categorize and structure accident data into threats, barriers, and consequences. Future research can expand and improve the pre-processing of the dataset, incorporate detailed accident reports, and investigate feedback loops with human experts to ensure continuous improvement and relevancy of the LLMs.

7. Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used [Chat GPT] to help editing the language of the paper to improve readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

8. References

Ashkenazi, O., Isaac, S., Giretti, A., Carbonari, A., & Durmus, D. (2023). Chapter Transforming Building Industry Knowledge Management: A Study on the Role of Large Language Models in Fire Safety Planning.

Baek, S., Jung, W., & Han, S. H. (2021). A critical review of text-based research in construction: Data source, analysis method, and implications. Automation in Construction, 132, 103915.

Bruce, P. C. (2016). *Data mining for business analytics : Concepts, techniques, and applications with xlminer*. John Wiley & Sons, Incorporated.

de Ruijter, A., & Guldenmund, F. (2016). The bowtie method: A review. *Safety science*, *88*, 211-218.

Dong, Q., Li, L., Dai, D., Zheng, C., Wu, Z., Chang, B., ... & Sui, Z. (2022). A survey on in-context learning. arXiv preprint arXiv:2301.00234.

Drupsteen, L., & Hasle, P. (2014). Why do organizations not learn from incidents? Bottlenecks, causes and conditions for a failure to effectively learn. *Accident Analysis & Prevention*, *72*, 351-358.

Duryan, M., Smyth, H., Roberts, A., Rowlinson, S., & Sherratt, F. (2020). Knowledge transfer for occupational health and safety: Cultivating health and safety learning culture in construction firms. *Accident Analysis & Prevention*, *139*, 105496.

Ekgren, A., Gyllensten, A. C., Gogoulou, E., Heiman, A., Verlinden, S., Öhman, J., ... & Sahlgren, M. (2022, June). Lessons Learned from GPT-SW3: Building the First Large-Scale Generative Language Model for Swedish. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference* (pp. 3509-3518).

Fu, G., Xie, X., Jia, Q., Li, Z., Chen, P., & Ge, Y. (2020). The development history of accident causation models in the past 100 years: 24Model, a more modern accident causation model. Process Safety and Environmental Protection, 134, 47-82.

Ghimire, P., Kim, K., & Acharya, M. (2024). Opportunities and Challenges of Generative AI in Construction Industry: Focusing on Adoption of Text-Based Models. *Buildings*, *14*(1), 220.

Gibb, A., Lingard, H., Behm, M., & Cooke, T. (2014). Construction accident causality: learning from different countries and differing consequences. *Construction Management and Economics*, *32*(5), 446–459. https://doi-org.ezproxy.bib.hh.se/10.1080/01446193.2014.907498

Guan, J., Zixuan, Y., Chan, A. P., Choi, T., & Yang, Y. (2024). Factors affecting learning from incidents: A cross-industry review. Journal of Loss Prevention in the Process Industries, 105297.

Guo, J., Cai, Y., Fan, Y., Sun, F., Zhang, R., & Cheng, X. (2022). Semantic models for the first-stage retrieval: A comprehensive review. ACM Transactions on Information Systems (TOIS), 40(4), 1-42.

Holmström, O., Kunz, J., & Kuhlmann, M. (2023, May). Bridging the Resource Gap: Exploring the Efficacy of English and Multilingual LLMs for Swedish. In *Proceedings of the Second Workshop on Resources and Representations for Under-Resourced Languages and Domains (RESOURCEFUL-2023)* (pp. 92-110).

Jacinto, C., & Silva, C. (2010). A semi-quantitative assessment of occupational risks using bow-tie representation. Safety Science, 48(8), 973-979.

Jurafsky, D., & Martin, J. H. (2024). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (3rd ed., draft).

Khan, N., Kimito, E. C., Tran, S., Pedro, A., Soltani, M., Hussain, R., ... & Park, C. (2023). Chapter Extracting Information from Construction Safety Requirements Using Large Language Model.

Kurtz (2022, March 16). The KBLab Blog: Evaluating Swedish Language Models. Retrieved from https://kb-labb.github.io/posts/2022-03-16-evaluating-swedish-language-models/

Kuzucuoğlu, D., Koc, K., Kazar, G., & Tokdemir, O. B. (2023). Prioritization of risk mitigation strategies for contact with sharp object accidents using hybrid bow-tie approach. Safety science, 166, 106248.

Lau, S., Gonzalez, J., & Nolan, D. (2023). *Learning data science*. O'Reilly Media, Incorporated.

Lindberg, A. K., Hansson, S. O., & Rollenhagen, C. (2010). Learning from accidents–what more do we need to know?. *Safety Science*, *48*(6), 714-721.

Lundberg, J., Rollenhagen, C., & Hollnagel, E. (2009). What-You-Look-For-Is-What-You-Find–The consequences of underlying accident models in eight accident investigation manuals. *Safety science*, *47*(10), 1297-1311.

Min, S., Lyu, X., Holtzman, A., Artetxe, M., Lewis, M., Hajishirzi, H., & Zettlemoyer, L. (2022). Rethinking the role of demonstrations: What makes in-context learning work?. *arXiv preprint arXiv:2202.12837*.

Prieto, S. A., Mengiste, E. T., & García de Soto, B. (2023). Investigating the use of ChatGPT for the scheduling of construction projects. *Buildings*, *13*(4), 857.

Saka, A., Taiwo, R., Saka, N., Salami, B. A., Ajayi, S., Akande, K., & Kazemi, H. (2023). GPT models in construction industry: Opportunities, limitations, and a use case validation. *Developments in the Built Environment*, 100300.

Samuelsson, B. (2022). Arbetsskador i byggverksamhet 2021: Privat och offentlig verksamhet.

SHAYBOUN, M. (2022). Toward Accident Prevention Through Machine Learning Analysis of Accident Reports.

Shayboun, M., Koch, C., & Kifokeris, D. (2021). Learning from accidents: Machine learning prototype development based on the CRISP-DM business understanding. In *Proceedings of the Joint CIB W099 & W123 International Conference 2021: Changes and innovations for improved wellbeing in construction* (p. 43).

Silva, S. A., Carvalho, H., Oliveira, M. J., Fialho, T., Soares, C. G., & Jacinto, C. (2017). Organizational practices for learning with work accidents throughout their information cycle. *Safety science*, *99*, 102-114.

Smetana, M., Salles de Salles, L., Sukharev, I., & Khazanovich, L. (2024). Highway Construction Safety Analysis Using Large Language Models. *Applied Sciences*, *14*(4), 1352.

Taiwo, R., Bello, I. T., Abdulai, S. F., Yussif, A. M., Salami, B. A., Saka, A., & Zayed, T. (2024). Generative AI in the Construction Industry: A State-of-the-art Analysis. *arXiv preprint arXiv:2402.09939*.

Turhan, G. D. (2023). Life Cycle Assessment for the Unconventional Construction Materials in Collaboration with a Large Language Model. In *Proceedings of the International Conference on Education and Research in Computer Aided Architectural Design in Europe*. Education and research in Computer Aided Architectural Design in Europe.

Wang, X., & El-Gohary, N. (2023). Deep learning-based relation extraction and knowledge graph-based representation of construction safety requirements. Automation in Construction, 147, 104696.

Wang, J., Huang, J. X., Tu, X., Wang, J., Huang, A. J., Laskar, M. T. R., & Bhuiyan, A. (2024). Utilizing BERT for Information Retrieval: Survey, Applications, Resources, and Challenges. *ACM Computing Surveys*, *56*(7), 1-33.

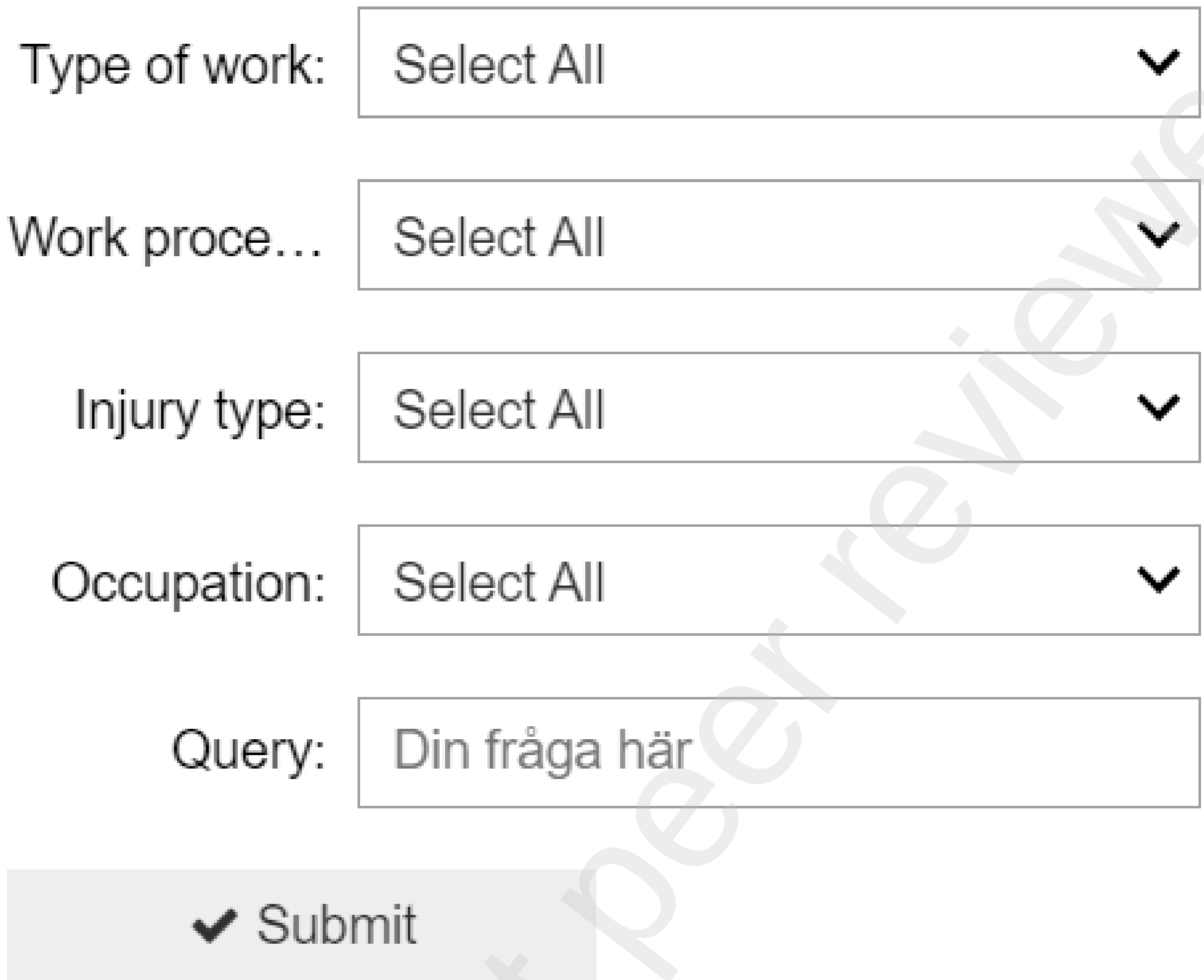
Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., ... & Fedus, W. (2022). Emergent abilities of large language models. *arXiv preprint arXiv:2206.07682*.

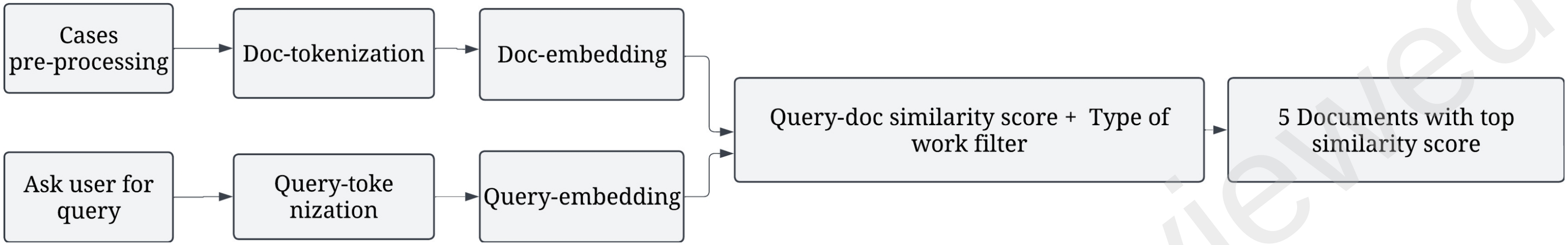
Wong, S., Zheng, C., Su, X., & Tang, Y. (2024). Construction contract risk identification based on knowledge-augmented language models. *Computers in Industry*, *157*, 104082.

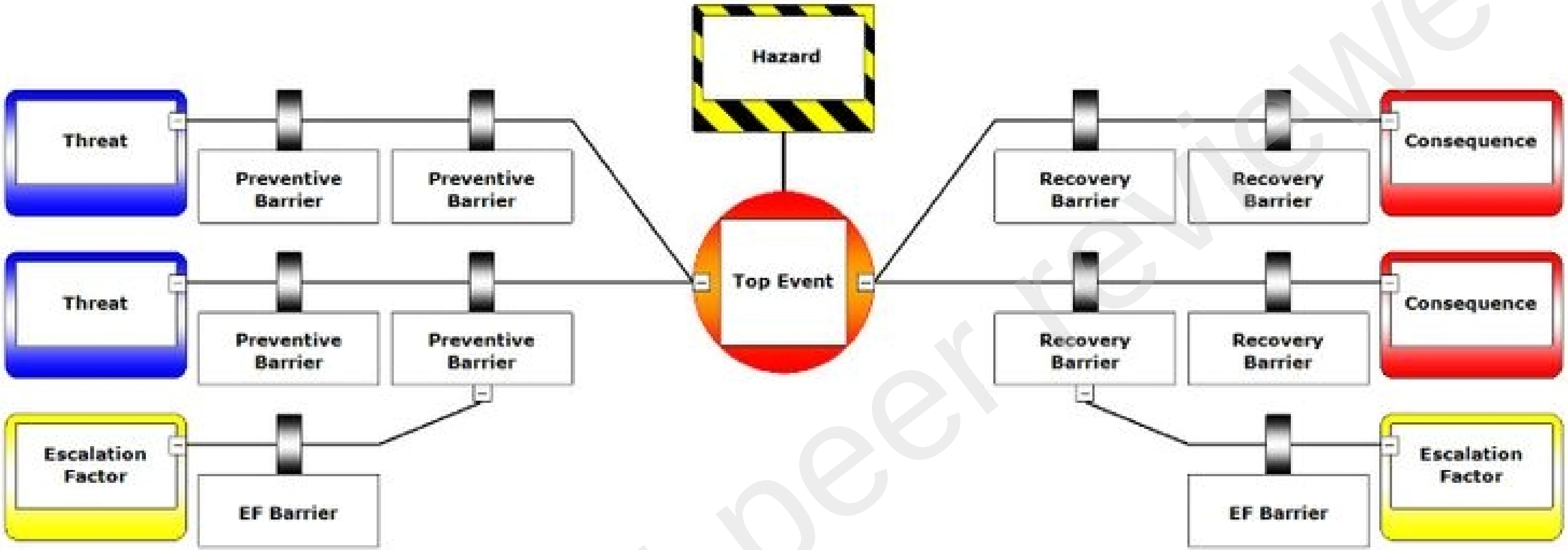
Wu, C., Li, X., Guo, Y., Wang, J., Ren, Z., Wang, M., & Yang, Z. (2022). Natural language processing for smart construction: Current status and future directions. Automation in Construction, 134, 104059.

Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., ... & Wen, J. R. (2023). A survey of large language models. arXiv preprint arXiv:2303.18223.

Zhu, Y., Yuan, H., Wang, S., Liu, J., Liu, W., Deng, C., ... & Wen, J. R. (2023). Large language models for information retrieval: A survey. *arXiv preprint arXiv:2308.07107*.


Type of work:
Select All
Work proce…
Select All
Injury type:
Select All
Occupation:
Select All
Query:
Din fråga här
Submit


Cases pre-processing
Doc-tokenization
Doc-embedding
Ask user for query
Query-toke nization
Query-embedding
Query-doc similarity score + Type of work filter
5 Documents with top similarity score


Hazard
Threat
Preventive Barrier
Preventive Barrier
Top Event
Recovery Barrier
Recovery Barrier
Consequence
Threat
Preventive Barrier
Preventive Barrier
Recovery Barrier
Recovery Barrier
Consequence
Escalation Factor
EF Barrier
EF Barrier
Escalation Factor

Query 1

Type of work in detail: Scaffolding

Search phrase: an accident occurred when a worker was carrying materials on or around scaffolding

Document 1:

Type of work in detail: Scaffold construction.

Occupation: Scaffold builder.

The last anomalous event preceding the injury: Slipping - tripping and falling - people who fall - Not spec.

Injury type: Person falls, walks/runs into immovable object - not specified

Cause, Level =1: Work factor

Cause, Level =2: Projecting object

Case_Text: Fall on the stairs. Went down a flight of stairs (existing). In the stairs there is a scaffold that is erected. The fitter was carrying scaffolding pipes, tripped and fell backwards against the scaffolding in the stairs, injured his back. Scaffolding mounted in the stairwell. Materials must be carried upstairs. Scaffolding must be dismantled during the next week.

Analysis 1:

Document Analysis:

Hazard: A confined space.

Top event: Scaffold builder falls down stairs.

Causes: The fitter was walking down an existing staircase carrying scaffolding pipes, tripped and fell backwards against the scaffolding.

Work factor.

Projecting objects.

Preventive barrier: Scaffolding mounted in the stairwell must be dismantled.

Consequences: Injured back.

Document 2: ...

Analysis 2: .

Query 2

Type of work in detail: Excavation

Search phrase: people around the machine or in the pit, collapse or fall

Document 1:

Type of work in detail: Excavation.

Occupation: Other professional worker.

The last anomalous event preceding the injury 1: Rupture, cracking, splitting, slip, fall, collapse of material - Not spec.

The last anomalous event that preceded the injury 2: Slip, fall, collapse of material - from above (falls on person).

Injury type: Hit by moving object, collision with - unspecified.

Cause, Level=1: Personal factor.

Cause, Level =2: Condition.

Case_Text: Stone fell into shaft. Person worked with rock splitting in shaft. Stone came loose from shaft wall and fell next to the person. The stone hit the person lightly on the leg. The person did not need treatment and said himself that there was no danger. Loose stone in shaft. Drilling rig is put in place to carry out the work.

Analysis 1:

Hazard: Depth of a shaft.

Top event: Stone falls down the shaft.

Causes: Loose stone in shaft. Stone came loose from shaft wall and fell next to the person.

Personal factor.

Permission.

Preventive barrier: Drilling rig is put in place to carry out the work.

Consequences: The stone hit the person lightly on the leg. The person did not need treatment and said himself that there was no danger.

Document 2: ...

Analysis 2: ...

Document 3: ...

Analysis 3: ...

System instructions: you are going to be provided with retrieved accident documents and their corresponding analysis examples, and your task is to select the documents relevant to the query and analyze them one by one similarly to the provided document analysis examples.
User prompt: use this analysis example: 'ICL demonstration'. to analyze the following case in the same way: 'Extracted documents').

System instructions: you are going to be provided with retrieved accident documents, and your task is to select the documents relevant to the query and analyze them one by one based on the Bow-tie model.
User prompt: Analyze the following cases: 'Extracted documents'

Query 1
Type of work in detail: Demolition work
Search phrase: Injuries due to material hazardous to health such as dust, quartz and asbestos

Document 1:
Type of work in detail: Demolition work.
Position: Land/Road O Construction Worker.
Cause, Level=1: Unfocused/distracted.
Cause, Level =2: Other Action (MUST BE DESCRIBED IN A TEXT FIELD).
Case_Text: A worker got something (glass fiber, according to the doctor) in the eye during the demolition of an old district heating culvert. Used safety glasses. The safety glasses not tight fitting. When tearing down a district heating line, existing insulating material dusted so that the worker probably got fragments from this in the eye. Tight-fitting glasses. Tighter safety glasses.
GPT response:
Hazard: Material hazardous to health (glass fiber in this case).
Top event: Worker got fiberglass in the eye during demolition work.
Causes:
- Used safety glasses, but they were not tight-fitting.
- Existing insulation material dusted during demolition of district heating pipe, which probably caused fragments in the eye.
Preventive barrier: Tighter safety glasses.
Consequences: Fragments of fiberglass in the eye, which led to possible eye irritation or damage.

Query 2
Type of work in detail: Earthwork and foundation
Search phrase: Accidents caused by a vehicle hitting someone

Document 1:
Type of work in detail: Earthwork and foundation
Occupation: Other professional worker
Injury type: Person falls, walks/drives into immovable object - not specified
Cause, Level=1 : Lack of barriers/signs
Cause, Level=2: Other personal factor (MUST BE DESCRIBED IN A TEXT FIELD)
Case_Text: Parked car without handbrake. Car was parked on the road behind shutdown. Professionals did not apply the handbrake and see from a distance that the car starts to roll. Running into the car, he opens the driver's door and stumbles. Lands with the leg in front of the rear wheel and slides with the vehicle. The vehicle and professional employee roll out of the shutdown and into a busy lane before the car stops. Truck stopped traffic on the trail so nothing else happened. Carelessness. No handbrake engaged, skidded when trying to stop the car, unlucky that there was an opportunity to go into traffic, great help from a helpful truck driver who stopped the traffic. Ensure handbrake is engaged.
GPT response:
Hazard: Parked vehicle without handbrake.
Top event: Professional employee tripped and ended up in front of a moving car.
Causes:
Professionals did not apply the handbrake.
Carelessness.
No handbrake engaged when parking.
Insufficient barrier.
Other personal factor.
Preventive barrier: Ensure handbrake is engaged.
Consequences: Professionals roll the vehicle into a busy lane, but no further damage thanks to the assistance of the truck driver.

Query 3
Type of work in detail: Other prefabricated articles
Search phrase: accidents or injuries when using lifting devices or lifting equipment.

Document 1:
Type of work in detail: Other prefabricated articles
Occupation: Blacksmith
Injury type: Crushing, crushing - middle
Cause, Level =1 : Improper lifting
Cause, Level =2: Other personal factor (MUST BE DESCRIBED IN A TEXT FIELD)
Case_Text: Was pinched when lifting. Lifted steel into place and beam tipped on forks and a worker pinched his finger. Accidents at work simply. Beam lifted and finger pinched. The worker got pinched during assembly of steel and broke his finger. Talked to the worker and went over the situation, they now know that they should not have their hand between the beam and the edge of the joist.
GPT response:
Hazard: Lifting device.
Top event: Balk tipped on forks and a worker pinched his finger.
Causes:
Improper lift.
Other personal factor (details must be described in the text field).
Preventive barrier: Inform and train personnel not to have their hand between the beam and the edge of the joist when lifting.
Consequences: A worker pinched while mounting steel and broke finger.

Query 1
Type of work in detail: Demolition work
Search phrase: Injuries due to material hazardous to health such as dust, quartz and asbestos

Document 1:
Type of work in detail: Demolition work.
Position: Land/Road O Construction Worker.
Cause, Level=1: Unfocused/distracted.
Cause, Level =2: Other Action (MUST BE DESCRIBED IN A TEXT FIELD).
Case_Text: A worker got something (glass fiber, according to the doctor) in the eye during the demolition of an old district heating culvert. Used safety glasses. The safety glasses not tight fitting. When tearing down a district heating line, existing insulating material dusted so that the worker probably got fragments from this in the eye. Tight-fitting glasses. Tighter safety glasses.
GPT response without ICL
Hazard: Exposure to dust and fiberglass during demolition work.
Top Event: Fiberglass entering the worker's eye.
Threats:
- Improper sealing of protective goggles.
- Inadequate containment of dust.

Preventive Barriers:
- Use of tightly sealed protective goggles.
- Implementation of dust control measures such as wetting materials or using dust extractors.

Consequences:
- Eye injury requiring medical attention.

Mitigative Barriers:
- Immediate medical intervention.
- Enhanced training on proper use of personal protective equipment (PPE).

Query 2
Type of work in detail: Earthwork and foundation
Search phrase: Accidents caused by a vehicle hitting someone

Document 1:
Type of work in detail: Earthwork and foundation
Occupation: Other professional worker
Injury type: Person falls, walks/drives into immovable object - not specified
Cause, Level=1 : Lack of barriers/signs
Cause, Level=2: Other personal factor (MUST BE DESCRIBED IN A TEXT FIELD)
Case_Text: Parked car without handbrake. Car was parked on the road behind shutdown. Professionals did not apply the handbrake and see from a distance that the car starts to roll. Running into the car, he opens the driver's door and stumbles. Lands with the leg in front of the rear wheel and slides with the vehicle. The vehicle and professional employee roll out of the shutdown and into a busy lane before the car stops. Truck stopped traffic on the trail so nothing else happened. Carelessness. No handbrake engaged, skidded when trying to stop the car, unlucky that there was an opportunity to go into traffic, great help from a helpful truck driver who stopped the traffic. Ensure handbrake is engaged.
GPT response without ICL
Hazard: Construction site activities involving parked vehicles.
Top Event: Worker falls and vehicle rolls into traffic.
Threats:
1. Improper securing of parked vehicles (handbrake not engaged).
2. Site layout allowing parked vehicles to potentially roll into traffic areas.

Preventive Barriers:
1. Strict protocols for vehicle parking, including mandatory brake engagement.
2. Regular checks by supervisors to ensure compliance with parking protocols.

Escalation Factors:
1. Forgetfulness or negligence from workers.
2. Slippery conditions increasing the risk of losing control while attempting to stop the vehicle.

Consequences:
1. Worker injury from falling and potential crushing.
2. Vehicle rolling into active traffic lanes, increasing the risk of broader traffic accidents and related injuries.

Mitigation Barriers:
1. On-site staff intervention to stop traffic and prevent further harm.
2. Training and reinforcement to the workers on the importance of securing vehicles.

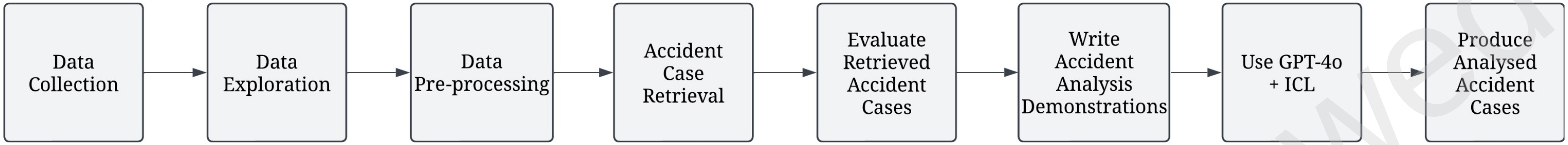
Recovery Measures:
1. Immediate medical assistance for the injured worker.
2. Traffic management to clear any risk caused by the rolling vehicle.
3. Reviewing and reinforcing parking and vehicle securing procedures among site workers.

Query 3
Type of work in detail: Other prefabricated articles
Search phrase: accidents or injuries when using lifting devices or lifting equipment.

Document 1:
Type of work in detail: Other prefabricated articles
Occupation: Blacksmith
Injury type: Crushing, crushing - middle
Cause, Level =1 : Improper lifting
Cause, Level =2: Other personal factor (MUST BE DESCRIBED IN A TEXT FIELD)
Case_Text: Was pinched when lifting. Lifted steel into place and beam tipped on forks and a worker pinched his finger. Accidents at work simply. Beam lifted and finger pinched. The worker got pinched during assembly of steel and broke his finger. Talked to the worker and went over the situation, they now know that they should not have their hand between the beam and the edge of the joist.
GPT response without ICL
Hazard: Lifting and positioning steel beams.
Threats:
- Improper lifting technique.
- Insufficient communication of safety procedures.

Preventive Measures:
- Proper training on lifting techniques.
- Clear communication and enforcement of safety protocols.

Top Event: Beam tipping during lifting.
Mitigation Actions:
- Immediate medical attention to the injured.
- Review and reinforce safety training and procedures to avoid placing hands between heavy objects and surfaces.
- Monitoring adherence to proper lifting protocols.


Data Collection
Data Exploration
Data Pre-processing
Accident Case Retrieval
Evaluate Retrieved Accident Cases
Write Accident Analysis Demonstrations
Use GPT-4o + ICL
Produce Analysed Accident Cases