

Weak signal detection for occupational safety

Skander Ghazzai^{1,2,*}, Daniela Grigori¹ and Raja Rebai²

¹Paris Dauphine - PSL University, CNRS, LAMSADE, Pl. du Maréchal de Lattre de Tassignys, Paris, France

²Air France, 45 rue de Paris, Tremblay-en-France, France

Abstract

In this paper, we address the challenge of detecting weak signals within the working environment using textual data. Our objective is to set up a decision support system to assist occupational safety experts in detecting weak signals. To achieve this, we adopt a unique method that combines portfolio maps and interpretation models. The portfolio maps integrate both structured and unstructured data, providing a holistic view of the potential safety risks in the workplace. The interpretation model further helps in comprehending and categorizing these signals accurately. We leverage the input of human experts on the potential weak signals to populate a dataset that serves as the basis for training a machine learning model. This model is designed to automate and optimize the detection and assessment of weak signals in the future. Our preliminary results demonstrate that this approach not only efficiently identifies weak signals but also offers the potential for continuous improvement in occupational safety management.

Keywords

Weak signals, Decision support system, Portfolio maps,

1. Introduction

In both daily life and the corporate world, weak signals are pervasive, often manifesting as early indicators that we might overlook until their implications become evident. For instance, the onset of illness might be preceded by subtle signs, such as feeling unusually cold, which we might dismiss until more significant symptoms develop. The definition of weak signals varies across different fields, each adapting the concept to its unique context and the specific nuances of the early indicators it seeks to identify and interpret. Originally used to detect preliminary signs of potential threats in the domain of military intelligence, the concept of weak signals has found its application in various domains. Ansoff (1975) [1] described weak signals as the earliest indicators of significant yet unrecognized changes or trends, and, later, Godet (1986)[2] portrayed them as developing changes with potentially significant future implications, necessitating contextual understanding. This paper draws on this fundamental insight to address the critical need for early detection of potential safety issues in occupational settings by introducing a novel approach that leverages text mining and machine learning techniques for weak signal detections. Recognizing their value, many companies are leveraging their untapped textual data to enhance operational efficiency and employee safety in the field of occupational safety. However, the manual process of weak signal detection could be burdensome for safety analysts. To address this, we propose a novel approach combining text mining techniques and machine learning to automate weak signal detection from work-related accidents. This aims to reduce the workload of analysts and support informed decision-making. This study underscores the importance of proactively addressing weak signals to stay competitive and ensure employee safety in today's rapidly evolving environment. In this paper, we will outline the methodology, detailing how we can effectively detect weak signals and enrich their understanding through contextual information.

Published in the Proceedings of the Workshops of the EDBT/ICDT 2024 Joint Conference (March 25-28, 2024), Paestum, Italy

*Corresponding author.

†These authors contributed equally.

✉ skander.ghazzai@dauphine.eu (S. Ghazzai);

daniela.grigori@dauphine.fr (D. Grigori)

ORCID 1234-5678-9012 (S. Ghazzai)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

However, due to data privacy concerns and in compliance with General Data Protection Regulation (GDPR) guidelines, the actual datasets coming from Air France and specific details used in this study will not be disclosed. Instead, we will provide examples and discussions based on hypothetical or suitably anonymized data, ensuring that the presentation of our findings remains as informative as possible while fully respecting the privacy of the data sources.

2. Background and related work

This section begins with a definition of weak signals, followed by a detailed discussion of various methodologies employed for their detection and interpretation. A signal, by definition, is a function that conveys information about a phenomenon. In this context, a signal can be understood as a specific instance or data point within a dataset that embodies meaningful information or patterns indicative of underlying phenomena. Yet, a "weak signal" is often regarded as an early, ambiguous, issue, or opportunity, serving as an initial indication of change. Due to their low intensity, fragmented nature, or potential noise within the data, these precursor signs often prove challenging to detect. A weak signal is by definition not universally recognized. The nature of weak signals varies according to the domain. This detection requires expert supervision. An expert who can navigate the complexities of data and contextualize these signals with the broader of domain-specific landscape. For example, in marketing, it could be identified as a subtle shift in consumer behavior, while in health contexts, it could be a subtle symptom hinting at a more significant medical issue. The concept of weak signals is rooted in the field of strategic foresight and is frequently used to support decision-making and strategic planning by anticipating potential future development. Embracing data-driven decision-making processes enhances the ability to navigate through uncertainty with informed, evidence-based strategies. This alignment with data-driven processes ensures a comprehensive analysis of available data, facilitating more accurate predictions and strategic actions in the face of uncertain future trends. Hiltunen[3] introduces a 3D spatial model for their definition. This model, based on Pierce's model[4], consists of three components:

- "Signal" refers to the frequency or visibility of sig-

nals. In other words, it quantifies how often a signal appears or the visibility of this signal.

- "Issue" represents the number of occurrences or contexts in which the signal appears, indicating the extent to which the signal spreads.
- "Interpretation" signifies the level of understanding of signals by information users, i.e., the degree to which receivers comprehend these signals.

This definition has proven its effectiveness in a lot of studies over the years [5] [6] [7]. Techniques for detecting weak signals can be broadly classified into four categories[8]: statistics-based methods, graph theory, machine learning, and semantics expert knowledge. Each method has its strengths and limitations. In the context of weak signal detection, graph theory is instrumental in analyzing the topology of graphs and structures [9] serving as a method to uncover weak signals. Graph theory, although effective, necessitates a high level of expertise for correct interpretation. Machine learning techniques, on the other hand, can offer powerful predictive capabilities, but the 'black box' nature of some of these models often makes it challenging for the expert to comprehend the reasoning behind specific predictions, especially on sensitive topics[10]. Semantic expert knowledge heavily relies on an individual expert's knowledge and intuition. This dependence makes it less consistent and challenging to scale up due to varying expertise levels. Lastly, statistics-based methods, while often robust and dependable, may fail to capture qualitative nuances. The effectiveness of these methods is heavily reliant on the quality and consistency of the data and the expert knowledge. In this research, we will concentrate on statistics-based methods, specifically portfolio map methods. The choice to concentrate on portfolio maps stemmed from the cyclical and seasonal nature of the work-related accident data available to us, as well as the preferences expressed by experts in the field.

Portfolio maps, a form of statistics-based methodology, are used to detect weak signals and identify emerging trends and issues. They serve as visual tools designed to manage and track multiple weak signals simultaneously.

Key characteristics of portfolio maps include:

- They typically use a two-dimensional matrix to plot signals based on their level of uncertainty/ambiguity and impact/importance. As per Ansoff[1], strong signals are highly impactful and certain, while weak signals are either low in impact or high in uncertainty.
- Portfolio maps are relatively easier to interpret for those who are unfamiliar with advanced data analytics or graph theory, making them accessible to a wider range of decision-makers.
- These maps facilitate proactive measures against potential risks or disruptions by visualizing the relative position of various signals and aiding in prioritization and decision-making.

The foundational work on portfolio maps was initiated within the context of business opportunities for solar cells by Yoon[7]. Yoon proposed a quantitative approach based on keyword text-mining to identify weak signal topics, thereby developing an effective method for measuring "signals" and "issues" as per Hiltunen's approach. Two novel metrics were introduced in this paper, namely the Degree of Visibility (DoV) and the Degree of Diffusion (DoD). The DoV aimed

to evaluate the visibility of a signal, while the DoD focused on the issue-level aspect of a keyword, enabling the identification of emerging weak signals or trends that might not be immediately visible through conventional text-mining approaches. The DoV for keyword i during the period j is calculated as follows:

$$\text{DoV}_{i,j} = \frac{\text{TF}_{i,j}}{\text{NN}_j} \times (1 - \text{tw} \times (n - j))$$

$\text{TF}_{i,j}$ is the number of appearance of the word i in period j , NN is the total number of documents. n is the number of periods and tw is a time weight. This formula suggests that the influence of a keyword is related to its frequency, and recent instances are given greater significance. Similarly, the DoD for a keyword i in period j is calculated as follows:

$$\text{DoD}_{i,j} = \frac{\text{DF}_{i,j}}{\text{NN}_j} \times (1 - \text{tw} \times (n - j))$$

Here $\text{DF}_{i,j}$ represents the document frequency of the keyword i appearing during the period j (the number of documents in which the keyword appears). This formula represents the spread of the signal, with a higher DoD value signifying a wider distribution of the issue. These two metrics are employed to generate two keyword portfolio maps: Keywords Emergence Map (KEM) and Keywords Issue Map (KIM). These maps, featuring a two-dimensional representation, have the horizontal axis indicating the average term or document frequency, and the vertical axis representing the average growth rate of DoV or DoD. These maps assist in identifying weak signals: words with a currently low frequency but high growth rates, suggesting potential swift future escalation in significance. In Figure 1, the median value on the X and Y axes divides the maps into four quadrants, classifying the keywords based on their locations. This division facilitates the automatic and dynamic categorization of keywords. The maps can subsequently be interpreted by experts, and common terms found in the same area are categorized as either a weak signal, a strong signal, or a well-known yet not strong signal. According to [11], this method surpasses human experts when dealing with large textual datasets. Weak signals can be represented in both KIM and KEM maps as keywords situated in the top left corner. However, prior research using portfolio maps has revealed some limitations, such as the difficulty in distinguishing mixed signals on each map and the potential for multiple meanings for a single keyword. The definition of weak signals, as provided earlier, introduces an interpretational dimension. By definition, a weak signal has a low level of interpretation. If it were otherwise, it would be widely understood and thus wouldn't qualify as a 'weak' signal. According to Hiltunen, this dimension includes the context within which individuals anticipate potential future events. However, the fragmented and incomplete nature of weak signals can make it challenging for experts to distinguish between genuine weak signals and noise. To address this issue, researchers have developed various techniques, as outlined below.

Multiwords Analysis This technique, enhanced by natural language processing, determines the frequency of occurrence of a specific word or phrase (potential weak signals and related keywords) in the presence of other specific words or phrases. This analysis, also known as co-occurrence analysis, provides insights into the relationships

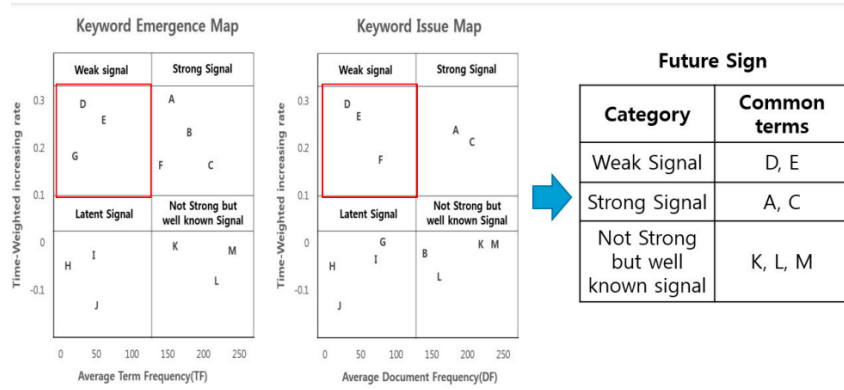


Figure 1: Future sign classification[11]

and contexts between different signals. The measurement of the association strength between terms or words in a given context is quantified[12]. This technique, augmented by natural language processing, determines the frequency of words co-occurring in conjunction with potential weak signals and related keywords. Previous studies [13] [14] have indicated that during the analysis of the portfolio map categorization, interpretability often suffers because keywords related to weak signals tend to be isolated terms. Consequently, the lack of context and relationships among these keywords restricts the scope and depth of the information that can be obtained. Hence, techniques like multiword analysis can play a crucial role in mitigating this interpretability issue by providing additional context.

Degree of Transmission This concept has been introduced to incorporate the third dimension of Hiltunen’s future sign model [3], known as the interpretation. The approach in [15] proposed a novel metric, the Degree of Transmission (DoT), specifically designed to assess the significance of terms within various sources, such as ScienceDirect, New York Times, and Twitter, from which keywords are automatically extracted. The Degree of Transmission is calculated as follows :

$$DoT_i = \sum Hindex_{journal}$$

Here, DoT_i represents the degree of transmission for term *i*, while *Hindex_{journal}* refers to the H-index of the journal in which the term appears. The H-index, which measures the impact and productivity of scholarly works, is summed across all the journals where a term is present, providing an overall measure of the term’s influence in the scientific literature. This approach has demonstrated promising results when both the Degree of Diffusion (DoD) and the Degree of Visibility (DoV) are multiplied by their corresponding DoT to enhance interpretability in the portfolio maps.

Topic Modeling [16] proposes a methodology for weak signals detection using LDA and Word2Vec. The approach is based on clustering topics at multi-level documents and extracting significant descriptors (weighted list of words). This model has the advantage of proposing a method for detecting weak signals based on tree-multi-clustering, unlike other works that are essentially based on the portfolio maps method, but don’t consider the temporal aspect.

Limitations of existing methods Many current methodologies rely on expert opinions or manual evaluations, presenting significant challenges for scalability for large datasets or real-time analysis. The adaptability of these methods to different domains and datasets is essential. By addressing these limitations, we aim to develop a more robust and versatile method for weak signal detection and analysis within the realm of safety occupational management. Our approach integrates the strengths of portfolio maps and machine learning algorithms, offering an improved solution for detecting and interpreting over time.

3. Methodology

In the following, we describe the approach adopted for developing a novel model for detecting weak signals in occupational safety management. This model aims to overcome the limitations of previous work, which lacks context, requires time-consuming analysis, and remains static over time (i.e., does not leverage user input for improvement). Our model combines text mining with a portfolio map approach, and a machine-learning algorithm, to identify and interpret weak signals more efficiently. The portfolio map approach visualizes the current state of an organization or system, analyzing the relationships between different elements to spot potential risks or opportunities. We utilized the concept of a weak signal consistent with Hiltunen’s definition, using a portfolio map that includes two axes: Visibility and Issue. To facilitate interpretation, we developed a user interface that aids users in confirming or discarding potential weak signals, enabling them to restrict the analysis to a given context. To further leverage expert input, each time a weak signal is confirmed, we populate a database with that specific weak signal, along with all factors contributing to its identification.

In the subsequent sections, we outline the systematic steps we adopted for our weak signal detection process. We start by preparing the documents, and ensuring they are ready for further analysis. Following this, we calculate new metrics to construct portfolio maps, an essential tool that aids in signal identification. After creating these maps, we perform a future sign classification. This step not only assists experts in interpreting the data but also provides a structured framework for identifying potential weak signals. The next step is the interpretation stage where we provide the expert with an interface that helps him discard weak signals from

noise and strong signals. Finally, using the labeled data, we develop and train a machine-learning model. This model, designed to automatically detect and assess weak signals, introduces an element of automation and efficiency to the process. Through this comprehensive approach, we aim to enhance both the effectiveness and precision of weak signal detection.

3.1. Data Preparation

In this section, we describe the steps taken to prepare the dataset for the study including data cleaning and preprocessing. The dataset used in this research was collected directly from the company (Air France) history containing a structured part related to the user and an unstructured part which is the "Verbatim" textual data that describes the event.

3.1.1. Data Cleaning

For this study, we work with a French data corpus and follow a standard preprocessing workflow to ensure data quality. Firstly, we remove and correct misspelled keywords, punctuation marks, and special characters from the text. Additionally, we convert all the text to lowercase to ensure consistency in our analysis. These data-cleaning steps help to improve the accuracy and reliability of our subsequent analyses.

3.1.2. Data preprocessing

To maximize the value of the information extracted from the text and reduce its complexity, we tokenize the text, splitting it into individual words. We use the word tokenizer NLTK tool for tokenization, a widely used tool for natural language processing. Once the text is tokenized, we perform lemmatization and part-of-speech tagging to refine the data further. This involves identifying the root form of each word, as well as its grammatical function in the sentence. We use the FrenchLefffLemmatizer and Camabert-ner NER model that was fine-tuned from camemBERT [17] on a wiki ner-fr dataset. Finally, we remove stop words, common words that do not add significant meaning to the text. In French, common stop words include "le," "la," "de," and "des." Removing these words helps to reduce noise and improve the accuracy of our analysis. Overall, our data preprocessing workflow ensures that the data is clean, consistent, and ready for our model, minimizing the number of keywords.

3.2. Adding contextual information

The process of detecting weak signals is a meticulous task, as each piece of information can potentially be valuable. To add contextual data to each event, we enrich the descriptive information with knowledge extracted from the enterprise database. Each sentence describing the event is represented as a combination of keywords, i.e., $sentence = \{keyword_1, \dots, keyword_n\}$. We augment these keywords with contextual information such as metadata (related to the place and time) and user-related data (role, department, number of past events, etc.). $Keyword_i \rightarrow \{Metadata_i, User-related_i\}$. This methodology allows us to capture the contextual details associated with each keyword, enabling us to use all the available data efficiently. As an example, consider the structure of this sentence: "

Keyword₁ event place Keyword₂" After performing our data preparation, each keyword would carry specific contextual information. For instance, the keywords could carry Metadata such as ["Location", "Time"] and User-related information such as ["Role", "Experience", "Age"]. By integrating this information, we can calculate metrics like the degree of Transmission, which allows us to identify potentially hazardous events more effectively. Moreover, this methodology can enhance our data-driven decision-making and a more robust detection of critical events.

3.2.1. Feature Engineering: Topic Modeling for Event Typologies

To extract and utilize as much information as possible from the textual data at hand, we collaborated with domain experts to fine-tune a Camembert model - a transformer-based language model specifically optimized for French language processing. This fine-tuning process involved training the model to categorize accident descriptions based on a pre-established set of typologies denoted as T. Crafted and validated by field experts, these typologies span a wide array of accident categories. The resulting model acts as a classifier, capable of processing unstructured text descriptions of accidents and categorizing each into its corresponding typology. Additionally, an "Unknown" category is provisioned for instances where the model cannot confidently assign a description to any of the known typologies.

3.3. Contextualized portfolio maps

In this part, we are going to introduce how are we going to construct the portfolio maps. The first step is the data representation which gives a numerical value to the issue and visibility to each keyword.

3.3.1. Data representation

We want to represent numerically the visibility and issue for every keyword while staying consistent with the definition of Hilutmen. We propose a new approach that represents the Visibility of every keyword defined as being the number of different populations (categories of users) that used the keyword with the degree of visibility (DoV). To capture the uncertainty of a given signal we propose the Issue notion, represented by the degree of diffusion with the number of different event typologies the keyword has been used in, trying to capture the level of uncertainty. The new formulas that we introduced with the contextualization are as follows:

$$\text{Visibility}_{i,j} = \text{DoV}_{i,j} \times \text{nP}_{i,j}$$

$$\text{Issue}_{i,j} = \text{DoD}_{i,j} \times \text{nT}_{i,j}$$

Where: $\text{DoV}_{i,j}$ is the degree of visibility of the keyword i in the period j , $\text{DoD}_{i,j}$ is the degree of diffusion of the keyword i in the period j , $\text{nT}_{i,j}$ is the number of different typologies that the keyword i is found in the period j , $\text{nP}_{i,j}$ is the number of different populations that used the keyword i in the period j , i is the index for the keyword, j is the index for the period. The data derived from the textual documents we analyze offers a variety of information. Details such as the frequency of accident reports made by an individual and the experience of the employee contribute to this rich context. Inspired by the H-index we created new metrics Degree of Transmission (DoT) taking into account

this personalized, user-related data. By associating such additional context with all keywords, we can significantly enhance the performance of our weak signal detection. By combining the old definition of Yoon with the new contextualization, we try to have a more precise representation of the impact/importance and the uncertainty/ambiguity of every keyword.

3.3.2. Portfolio maps categorization

After representing our signals with the metrics previously designed, we construct two portfolio maps:

- Keyword Emergence Map KEM where the X-axis is mapped with the geometric mean of the term frequency, and the increasing rate of visibility for each keyword is mapped into the Y-axis.
- Keyword Issue Map KIM has the X-axis mapped with the geometric mean of the term appearance, and the increasing rate of the issue for each keyword is mapped into the Y-axis.

Using these portfolio maps, we are equipped to identify potential weak signals in terms of both visibility and issue. Typically, keywords classified as potential weak signals exhibit below-average term frequency and document appearance, coupled with an appearance in different topics and used by different populations which results in above-average increases in issue and visibility rates, indicative of emerging trends. We then identify commonalities between the weak signal candidates derived from each keyword map, categorizing these as our prime weak signal contenders. In the subsequent phase, these candidates are presented to domain experts for interpretation and validation.

3.4. Interpretation

Interpretation refers to the ability to understand the significance of the information extracted from the portfolio maps categorization. As a result of the whole system, experts will have access to two outputs that will help them in the decision-making process :

- A list of potential weak signals represented in the Keyword Issue Map, depending on their Degree of Diffusion and Degree of Transmission.
- A list of potential weak signals represented in the Keyword Emergence Map, depending on their Degree of Visibility and Degree of Transmission.

Our goal in this part is to minimize the risk of misinterpreting a weak signal by maximizing the level of interpretability of every signal.

3.4.1. Domain-specific

Having to work with all the potential weak signals listed in the portfolio maps can be challenging and can be time-consuming. To be more efficient in how we are treating the keywords we tried to come up with a way to rank the list of potential weak signals. with the help of experts, we developed a set of ranking rules representing the emergency of which we have to treat the results as follows:

1. The list of potential weak signals represented in both the Keyword Emergence Map and the Keyword Issue Map

2. The keyword with the highest increasing visibility rate
3. The keyword present in the KIM but not in the KEM
4. The keyword present in the KEM but not in the KIM

By applying those rules, we try to focus our attention on the most relevant keywords thus reducing the time and effort required to identify and analyze weak signals.

3.4.2. Multi-word analysis

After ranking, the output words from the Portfolio Maps emerge as potential weak signals or terms related to weak signals. This stage presents a challenge due to the various interpretations a single keyword might encompass, a problem highlighted in previous studies [13] [14]. To mitigate this issue, we implement multi-word expression analysis, a technique designed to refine our results. This method examines the words that appear immediately before and after the identified term in every instance, excluding common stopwords. Consequently, we generate a co-occurrence relationship related to a keyword, ranking it based on its increasing rate of visibility and issues.

3.4.3. Predictive modeling

Identifying weak signals in portfolio maps and interpreting the results is a multi-step complex task that not only requires a signal to be found in the intersection of the KIM and KEM but also the validation throughout the interpretation and confirmation of the expert. To use knowledge from the past, we aim to leverage every expert input to create a model that predicts the probability of a keyword being related to a weak signal using this procedure :

1. The expert selects a keyword from the portfolio map found in the intersection of the KIM and KEM.
2. The expert using the interpretation model labels the signal (related to a weak signal or not)
3. Add the newly labeled signal to the dataset

After every use of our model, we gather the expert's feedback (label), incorporate it into our existing dataset, and once a substantial amount of new information has been accumulated, we recalibrate our model in light of this freshly acquired data. After creating the model we try to implement the model output in the portfolio maps. This model predicts the probability of other keywords being a weak signal without taking into consideration the position of the keyword in any of the portfolio maps and is thus complementary to the portfolio map. This approach is supported by previous research in the field of predictive modeling, which has shown that combining expert input with machine learning can lead to more accurate predictions and better decision-making [18][19].

4. Experimental Evaluation and Results

In this section, we present a preliminary evaluation of our model's effectiveness in detecting potential weak signals and preventing future accidents, using real-world occupational safety data. We then show the results of our approach applied to our dataset.

4.1. Data Representation

Safety occupational management event reporting follows a strict protocol, generating a dynamic database that is continuously updated with new reports. These reports contain both structured and unstructured data relating to the event and the user. Due to GDPR constraints, we will not delve into the structured data, which primarily includes user-related information such as employee profiles, work details, and accident history. The primary focus of our study lies in the unstructured data, including the event label and description, which represent rich sources of textual information. This unstructured segment can provide valuable insights into weak signals when analyzed effectively.

4.2. Portfolio maps applied to our dataset

The first step of our application is the construction of the portfolio maps (Figures 2 and 3). We represent the growth rate of every signal keyword with our designed metrics and then employ portfolio map categorization. This pro-

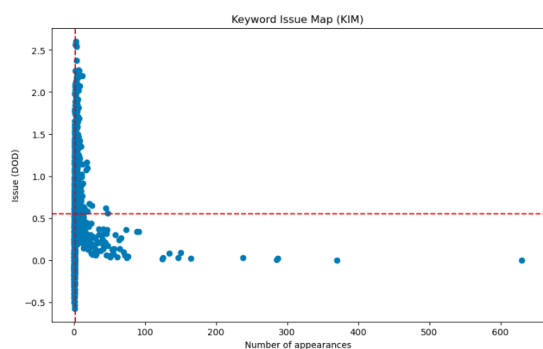


Figure 2: Keyword Issued Map (KIM)

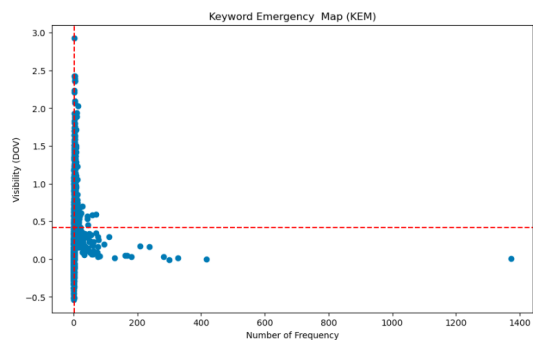


Figure 3: Keyword Emergence Map (KEM)

cess allows us to identify a set of potential weak signals. Given the sensitive nature of the data, continuous monitoring and updating are vital. Each new accident report plays a pivotal role in this ecosystem, substantially impacting both the composition of our portfolio map and the list of potential weak signals. Each keyword has the potential to shift the dynamics of the portfolio, leading to new and meaningful discoveries. In the following, we analyze the difference between our newly developed metrics and the classical ones, distinguishing between strong signals and potentially weak signals. Our approach here is instance-by-instance, and we aim to discern the comparative rates of increase for these metrics. We focus our analysis on the keyword 'salarie' due to its importance as it consistently appears

at the start of each report, thereby providing a common reference point. To maximize the impact of our analysis, we conduct a time-series analysis, closely examining the metrics associated with the keyword 'salarie' throughout June (Figures 4 and 5). Each metric calculation only encompasses

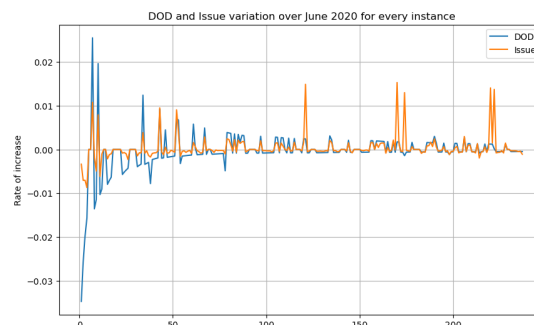


Figure 4: Metrics variation over June 2020 for every instance: DOD Vs issue

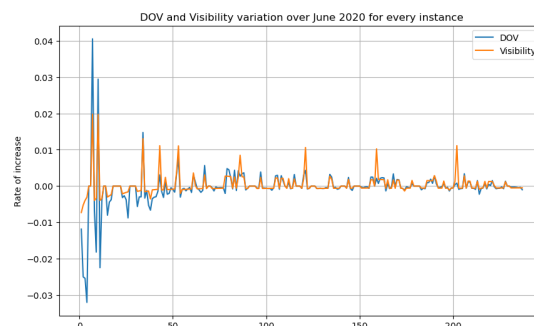


Figure 5: Metrics variation over June 2020 for every instance: DOV vs Visibility

past and present instances, intentionally leaving out future ones. This methodology allows us to conduct an exhaustive comparison of traditional metrics (DOV and DOD) and the newly developed ones ('Visibility' and 'Issue'). By doing so, we gain a clear understanding of the rate at which these metrics increase over time, offering vital insights for the construction of our portfolio maps. In our comparison of the 'Degree of Visibility' (DOV) and 'Visibility' metrics, we initially noticed that DOV exhibits a sharper peak. This suggests a higher increase rate for DOV compared to the 'Visibility' metric. Essentially, the keyword 'salarie' appears more frequently in reports over time, indicating its growing visibility. However, this rise in visibility does not correspond to equal growth in the 'Visibility' metric. As 'Visibility' is a product of DOV and the diversity of typologies in which the keyword appears, a rapid increase in DOV won't match the 'Visibility' increase rate if the keyword 'salarie' doesn't appear in a variety of typologies there being in the same typologies there being a well-known signal. This divergence highlights the more nuanced understanding offered by the 'Visibility' metric, which takes into account not just keyword frequency but also the diversity of contexts (typologies) where the keyword is present. On the other hand, there are instances when 'Visibility' peaks significantly higher than DOV. This occurs when the keyword, despite being used less frequently, is present in a wider range of contexts, leading to a sharper rise in 'Visibility' compared to DOV. The 'Issue' and DOD demonstrate a similar behavior to 'Visibility' and DOV, respectively. We have implemented our

system as an auxiliary tool for a proactive approach. It automatically updates whenever the list of potential weak signals changes, or when the model predicts that a keyword might be a weak signal. This proactive approach facilitates early detection and prevention efforts, reinforcing the system's contribution to accident prevention.

4.3. Interpretation

Despite our efforts to include contextualized metrics, we acknowledged the need for an approach that offers a broader and global view of the system. To address this, we developed a dynamic dashboard that empowers the expert to select specific data from the structured dataset, including user-related and event-related information. We transformed all available data into selectable variables, thereby augmenting the depth and relevance of the contextualization process. The dynamic dashboard facilitates the confirmation of weak signals enabling experts to filter and choose data from a comprehensive set of structured variables. The dynamic dashboard facilitates

1. Type of periods
2. Sources: Population
3. Affiliation: where the accident has been submitted
4. Topic: typologies of the event
5. Type of Keyword (Named entity recognition(NER))

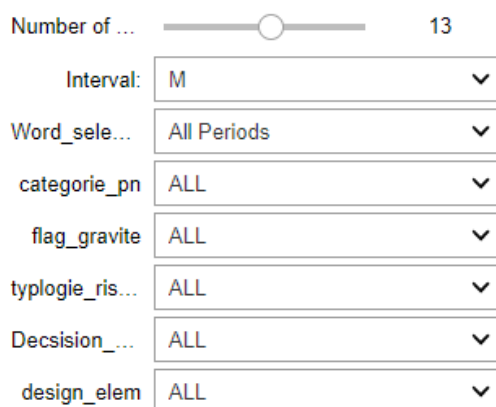


Figure 6: Partial View of the Dashboard Interface

the confirmation of weak signals enabling experts to filter and choose data from a comprehensive set of structured variables. In the interest of illustrating our approach to enhancing data contextualization and facilitating the identification of weak signals, Figure 6 presents a partial view of the dynamic dashboard interface. This segment showcases a subset of the configurable options available to analysts, including 'Type of periods', 'Sources: Population', 'Affiliation' indicating the submission source of the accident report, 'Topic' relating to the typologies of the event, and 'Type of Keyword', which leverages Named Entity Recognition (NER) for deeper analysis. It's important to note that this figure represents only a fraction of the full interface, specifically chosen to demonstrate the flexibility and depth of analysis without disclosing sensitive variable details. Through these selectable options, the dashboard empowers analysts to tailor their examination of the data, thus enhancing the precision and relevance of their findings. This not only provides a more in-depth and targeted analysis but also accelerates the confirmation

of weak signals. The ability to combine the strengths of both textual and structured data significantly enhances the detection and interpretation of weak signals, ensuring a more thorough and accurate understanding of underlying trends and relationships.

Multi-word analysis To show more accurate and interesting information about the detected term of interest than an analysis based only on single words, we perform a multi-word analysis. This enhances the overall interpretability, by showing the co-occurrence of keywords that have been used with the selected signal. Figure 7 represents the Multi-word analysis for a potentially weak signal. Because of RGPD issues, we had to anonymize the keywords and the events associated with them.

	Word_anonyme	event_anonyme	Value
204	keyword_204	event_1	13.005272
92	keyword_92	event_41	4.745167
116	keyword_116	event_1	3.514938
45	keyword_45	event_24	2.987698
221	keyword_221	event_1	2.987698

Figure 7: Multi-word analysis example

Machine learning We start by estimating the probability of a keyword's occurrence based on the available metadata. This process involves the use of CamemBERT[17] tokenization and embedding in combination with the structured data. The resulting information is then presented to the expert for portfolio map interpretation. Each time a keyword is selected by the expert, we enrich a dataset that will, in the future, enable us to develop a more effective machine-learning model. However, due to the low absolute frequency of weak signals within our dataset, the effectiveness of the current approach could not be assessed. Nonetheless, it allows continual refinement and improvement of the model.

5. Conclusion

This research paper has introduced innovative metrics and processes tailored to detect weak signals in the domain of occupational safety management. By integrating a portfolio map approach and machine learning algorithms, the proposed model aims to enhance the efficiency of the detection process. We have adopted a contextualized approach designed to support analysts and facilitate informed decision-making. This approach effectively captures contextual aspects and optimizes the utilization of available data. Through the synergy of expert knowledge with machine learning techniques, our model adeptly identifies weak signals and potential emergent trends. Future research should focus on further refining the model, incorporating additional contextual information, and utilizing more advanced machine-learning techniques to enhance the detection of weak signals. Regarding the model's interpretability, we anticipate that integrating highly sophisticated large language

models, such as ChatGPT, could make a significant contribution to this objective. Such integration could provide richer, more nuanced interpretations of the detected weak signals, thereby enabling safety management professionals to make more informed decisions. Furthermore, evaluating our model's adaptability across a variety of domains and industries could ascertain its universal applicability and versatility in diverse occupational safety management contexts.

References

- [1] H. I. Ansoff, Managing strategic surprise by response to weak signals, *California Management Review* 18 (1975) 21–33. doi:10.2307/41164635.
- [2] M. Godet, *From Anticipation to Action: A Handbook of Strategic Prospective*, UNESCO Publishing, 1994.
- [3] E. Hiltunen, The future sign and its three dimensions, *Futures* 40 (2007) 247–260. doi:10.1016/j.futures.2007.08.021.
- [4] C. S. Peirce, Some consequences of four incapacities, *Journal of Speculative Philosophy* 2 (1868) 140–157.
- [5] S. Roh, J. Choi, Exploring signals for a nuclear future using social big data, *Sustainability* 12 (2020) 5563. doi:10.3390/su12145563.
- [6] P. Krigsholm, K. Riekkinen, Applying text mining for identifying future signals of land administration, *Land* 8 (2019). doi:10.3390/land8120181.
- [7] J. Yoon, Detecting weak signals for long-term business opportunities using text mining of web news, *Expert Systems with Applications* 39 (2012) 12543–12550. doi:10.1016/j.eswa.2012.04.059.
- [8] P. Rousseau, D. Camara, D. Kotzinos, Weak signal detection and identification in large data sets: A review of methods and applications, *ResearchGate*, 2021. doi:10.13140/RG.2.2.20808.24327/1.
- [9] H. A. Jamra, M. Savonnet, E. Leclercq, Beam: A network topology framework to detect weak signals, *International Journal of Advanced Computer Science and Applications* (2022). doi:10.14569/ijacsa.2022.0130402.
- [10] C. Rudin, Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead, *Nature Machine Intelligence* 1 (2018) 206 – 215. doi:10.1038/s42256-019-0048-x.
- [11] S. Roh, J. Choi, Exploring signals for a nuclear future using social big data, *Sustainability* 12 (2020) 5563. doi:10.3390/su12145563.
- [12] L. Vaughan, E. Romero-Frías, Exploring web keyword analysis as an alternative to link analysis: a multi-industry case, *Scientometrics* (2012). doi:10.1007/S11192-012-0640-X.
- [13] Y.-J. Lee, J.-Y. Park, Identification of future signal based on the quantitative and qualitative text mining: A case study on ethical issues in artificial intelligence, *Quality and Quantity: International Journal of Methodology* 52 (2018) 653–667.
- [14] C. Park, H. j. Kim, A Study on the Development Direction of the New Energy Industry Through the Internet of Things - Searching for Future Signals Using Text Mining, Technical Report, Korea Energy Economics Institute, 2015.
- [15] I. Griol-Barres, S. Milla, A. Cebrián, H. Fan, J. Millet, Detecting weak signals of the future: A system implementation based on text mining and natural language processing, *Sustainability* 12 (2020) 1–1. doi:10.3390/su12198141.
- [16] J. Maitre, M. Menard, G. Chiron, A. Bouju, Détection de signaux faibles dans des masses de données faiblement structurées, *RIDoWS* 3 (2019) null. doi:10.21494/ISTE.OP.2020.0463.
- [17] L. Martin, B. Muller, P. J. O. Suárez, Y. Dupont, L. Romary, Éric Villemonte de la Clergerie, D. Seddah, B. Sagot, Camembert: A tasty french language model, in: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020, pp. 7203–7219.
- [18] H. Park, A. Megahed, P. Yin, Y. Ong, P. D. Mahajan, P. Guo, Incorporating experts' judgment into machine learning models, *Expert Systems With Applications* (2023). doi:10.1016/j.eswa.2023.120118.
- [19] F. d. V. Tamer Boyacı, Caner Canyakmaz, Human and machine: The impact of machine input on decision making under cognitive limitations, *Management Science* (2023). doi:10.1287/mnsc.2023.4744.