Leveraging machine learning for analysing text-based data and assessing human factors identified in occupational incident reporting

Maria Tiikkaja, PhD, Research manager et al.

The main research questions were:



Are the workplaces getting enough information on the causes of the accidents to prevent them happening again?



Is it possible to train machine learning models to identify and classify the underlying causes mentioned in incident reports?





Three industrial datasets analysed with Human **Factors Tool[™]**

- 1587 accident and near-miss • reports from four organisations
- 331 569 occupational accident • descriptions from the Workers' **Compensation Center**
- 67 accident investigations • from two organisations

Individual factors and actions

- Competence, conceptual mastery of work 1.
- 2. Situation awareness (perception, memory, decision
- making, response/execution
- 3. Working along norms and procedures 4.
- Understanding the whole/big picture 5. Preconditions, assurance and checking; assumptions
- Underload, overload; coping of these
- 6. 7. Vigilance, alertness, fatigue symptoms
- 8. Phase of life, anxiety, long term stress
- 9. Age; guality and guantity of work experience
- 10. Health, work ability
- 11. Motivation, attitudes
- Emotional state and reactions, mood 12.

Work characteristics 20. Quality and contents of work, work demands 21. Quantity of work, time pressure

- 22. Work distribution, work organisation
- 23. Technology, machines; usability and reliability
- 24. Procedures, regulations, norms; usability, validity
- 25. Possibilities to impact your own work
- Feedback from work. professional 26.
- esteem/appreciation 27. Possibility and ability to analyze and develop one's
- processes 28. Training: contents, effectiveness, possibilities to participate
 - Work hygiene factors, physical environment (noise, illumination, air condition, temperature, layout)

complex, dynamic and uncertain situations

Mastering

29.

Organisational factors

- 40. Management and leadership; structure, styles
- 41. Organisational/safety culture 42.
- levels (units, head office)
- 43.
- 44. Decisions concerning resources, personnel, equipment
- 45.
- 46. Co-operation with the partners; differences in way of actions
- Support from head office 47.





A-M. Teperi 2012

- Co-operation between different units and hierarchy
- Mutual understanding of others' work
- Change management

Group/team factors

- Distributed knowledge of the work, shared mental 30. model/situational awareness
- 31. Knowledge of every member is used
- 32. Misconceptions; misunderstanding; are they corrected
- 33. Structure and cohesion of the group (atmosphere, group dynamics, mutual support)
- 34. Communication in teams and between teams
- 35. Information flow (shift change)
- 36. Decision making (role of supervisor)



Analysing safety incidents with large language model (LLM) using training data

- The LLM was fine-tuned to classify safety incident texts according to HF Tool[™] top levels
- Training data labelled manually according to HF Tool[™] sub-items by sentence
 - Includes accidents, near-misses and safety observations recorded in organisations' systems
 - "Extended" with data from The Workers' Compensation Centre and investigations of accidents led to death
 - Approx. 1100 rows of training data



Modified example of the training data

Description/sentence	Related HF™ Tool item
The team went to the rescue.	33
Misinterpreted.	33
The day had been busy and I was tired.	7
The route was designed for truck traffic, but after changes made in the workplace, it also began to be used by pedestrians.	45
The working pair was tired due to not taking breaks.	7
My colleague and I noticed that the pipe was incorrectly installed.	32
4.10.2024 © Finnish Institute of Occupational Health	

The results

Accident and near-miss reports Accident descriptions from from organisations the Workers' Compensation Center 1%_1% 0%0% E.g: 12 % 18 % Individual: "The forklift collided with the guardrail, no sightings of the driver." 80 % 88 % Work: "The dressing room floor is splashing with water, risk of slipping." ■ Individual ■ Work ■ Group ■ Organisation ■ Individual ■ Work ■ Group ■ Organization © Finnish Institute of Occupational Health 4.10.2024

Classification of accident investigations – **HF-expert vs. LLM**

The LLM correctly classified 73 % of the texts that were categorized by HF-experts

- Individual factors and actions: 77 extracts
- Work characteristics: 137 extracts
- Group factors: 36 extracts
- Organisational factors: 81 extracts



Classification of the Safety and Chemicals Agency investigation – HF-expert vs. LLM

The LLM correctly classified 82 % of the texts that were categorized by HF-experts

- Organisational factors: 17 extracts
- Group factors: 0 extracts
- Work characteristics: 13 extracts
- Individual factors and actions: 8 extracts



4.10.2024

HF-expert analysis of accident investigations

- Text data on accident investigations (67 in total)
- Theory-oriented classification of the causes identified in the reports -> which HF Tool[™] sub-items can be seen

Individual factors and actions

Competence, conceptual mastery of work Situation awareness (perception, memory, decision

- making, response/execution
- Working along norms and procedures
- 4. Understanding the whole/big picture
- 5. Preconditions, assurance and checking; assumptions
- 6. Underload, overload; coping of these
- 7. Vigilance, alertness, fatigue symptoms
- 8. Phase of life, anxiety, long term stress
- 9. Age; quality and quantity of work experience
- 10. Health, work ability
- 11. Motivation, attitudes
- 12. Emotional state and reactions, mood

Work characteristics

- 20. Quality and contents of work, work demands
- 21. Quantity of work, time pressure
- 22. Work distribution, work organisation
- 23. Technology, machines; usability and reliability
- 24. Procedures, regulations, norms; usability, validity
- 25. Possibilities to impact your own work
- 26. Feedback from work. professional esteem/appreciation
- 27. Possibility and ability to analyze and develop one's processes
- Training: contents, effectiveness, possibilities to participate
 Work hygiene factors, physical environment (noise
 - Work hygiene factors, physical environment (noise, illumination, air condition, temperature, layout)

Mastering complex, dynamic and uncertain situations

Organisational factors

- 40. Management and leadership; structure, styles
- 41. Organisational/safety culture
- 42. Co-operation between different units and hierarchy levels (units, head office)
- 43. Mutual understanding of others' work
- 44. Decisions concerning resources, personnel, equipment
- 45. Change management
- 46. Co-operation with the partners; differences in way of actions
- 47. Support from head office

Group/team factors

- 30. Distributed knowledge of the work, shared mental model/situational awareness
- 31. Knowledge of every member is used
- 32. Misconceptions; misunderstanding; are they corrected
- 33. Structure and cohesion of the group (atmosphere, group dynamics, mutual support)
- 34. Communication in teams and between teams
- 35. Information flow (shift change)
- 36. Decision making (role of supervisor)

A-M. Teperi 2012

The results

- 46% of the identified factors were related to work characteristics
 - The most identified were working methods and instructions (HF24), functionality and availability of equipment systems and technology (HF23), and physical working environment and working conditions (HF29)

• 36% were related to individual factors and actions

• The most common were situational awareness (HF2), anticipating situations (assumptions and verification) (HF5) and following instructions and agreed procedures (HF3)

• 11% were related to organisational factors

- Most commonly related to overall management and mutual understanding of each other's work (HF43), decisions made at the organisational level (HF44) and cooperation with partners (HF46)
- 7% were related to group factors

© Finnish Institute of Occupational Health

4.10.2024

Conclusions

- The fine-tuned model works better than the data would allow us to expect
 - Due to the pre-training, which has 'encapsulated' information about Finnish language and high-level concepts such as those of HF
 - GPT-4 trials underway
- We can utilize the LLM in analysing the HF scope of any Finnish text
 - Gives interesting insight to the safety culture of an organisation: Is the training and reporting focused simply on an individual, or perhaps on other HF levels too
- The analysis reflects inadequate reporting -> Most texts do not have enough features that could be identified reliably with the HF Tool[™]
- Data quality and collection can be improved significantly (e.g. reducing noise)
 -> for more information, please contact me later!

4.10.2024

Based on the work done in and after "AI Safety project"

- 9/2020 9/2023
- Four large industrial companies participated
- Funded by the Finnish work environment fund, Finnish Institute of Occupational Health (FIOH) and participating organisations
- More info: <u>www.ttl.fi/en/aisafety</u>



Työsuojelurahasto Arbetarskyddsfonden The Finnish Work Environment Fund

© Finnish Institute of Occupational Health





Maria Tiikkaja





Akseli Nurmi



Vuokko Puro



Henriikka Kannisto



Tarja Heikkilä

4.10.2024



Thank You!

For more information and discussion, please contact: maria.tiikkaja@ttl.fi

LOOKING FOR MORE INSIGHTS INTO WORKING LIFE? FOLLOW US ON SOCIAL MEDIA AND VISIT OUR WEBSITE TTL.FI/EN