Natural Language Tools for Human Reliability Analysis

University of Strathclyde Glasgow

COSSAN Software

Virtual Raphael

Human Reliability Analysis (HRA) involves the use qualitative and quantitive methods to identify and analyse the human contribution to risk, which can be incorporated into overall system safety analysis.

Why is HRA important?

Reliability is a major requirement in safety-critical systems such as nuclear power plants. Evidence from past accidents shows that human errors are common contributors to accidents in such systems.
HRA can be applied to systematically identify and analyse human error – related events. It offers opportunities to improve human-machine interfaces, reliability, and safety.

The problem of data

Data are scarce and with different level of accuracy and confidence

Multi-attribute Technological Accidents Dataset (MATA-D)

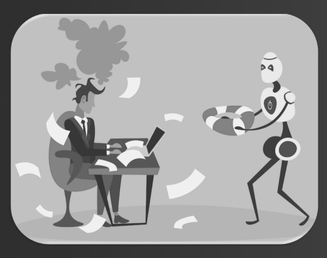
- 238 accident reports classified by risk experts
- Use a framework of organisational, tecnological and human errors, based on Hollnagel's Cognitive Reliability and Error Analysis Method (CREAM)
- Across 31 industrial sectors including oil & gas, aviation, chemical, nuclear, etc.)
- 23 major incidents from the Nuclear industry, such as Fukushima 2011, Kyshtym Disaster 1957 and more

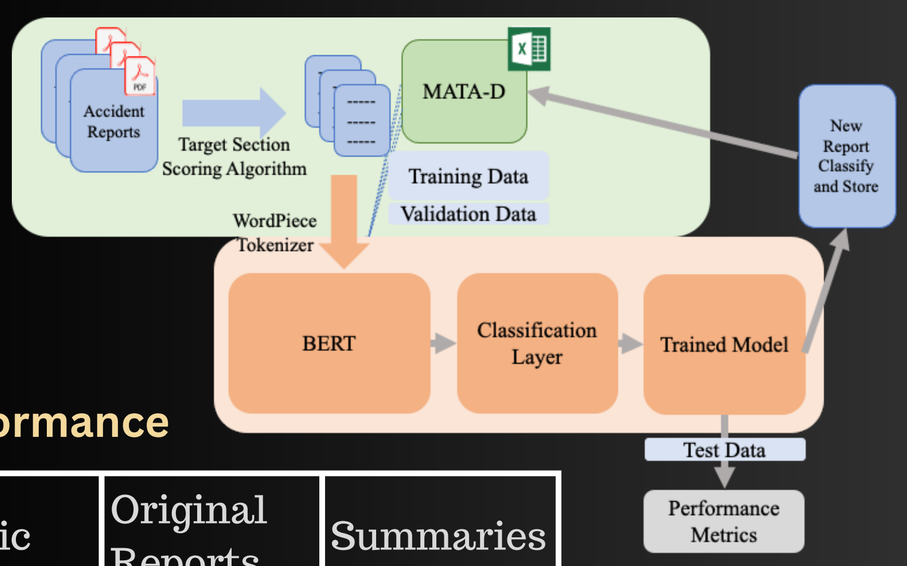
AI Human Factors Classificator

To read and classify one accident report, it takes an experienced expert around 3 days. This takes time away from value added mitigation activities; reduces the rate at which the databases (MATA-D) could expand to improve learning opportunities.

Solution

Train machines to help an expert classify reports and detect possible triggers for human errors. Using modern Natural Language Processing models, such as BERT (Bidirectional Encoder Representations from Transformers), fine tuned on MATA-D.

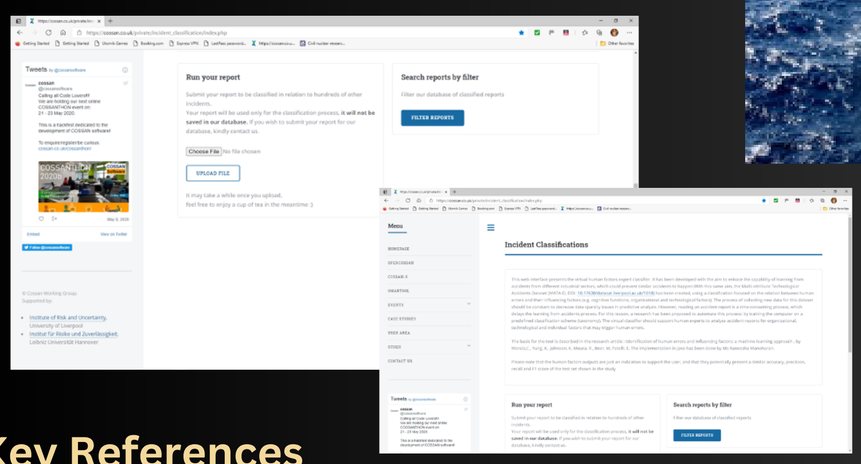


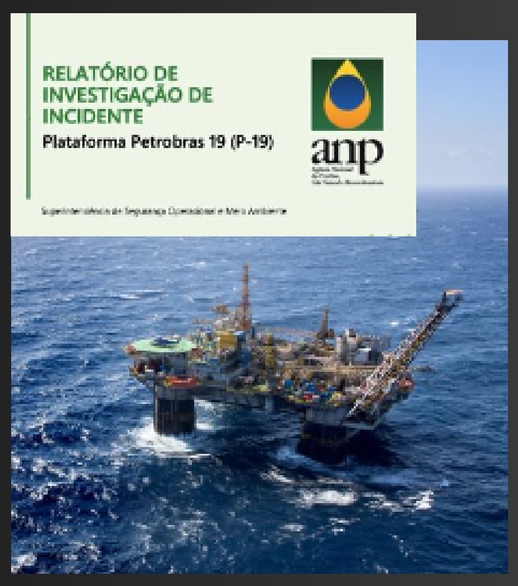


Performance

Metric	Original Reports	Summaries
Accuracy	91%	83%
Precision	83%	72%
Recall	70%	61%
F1 score	76%	66%

Web interface



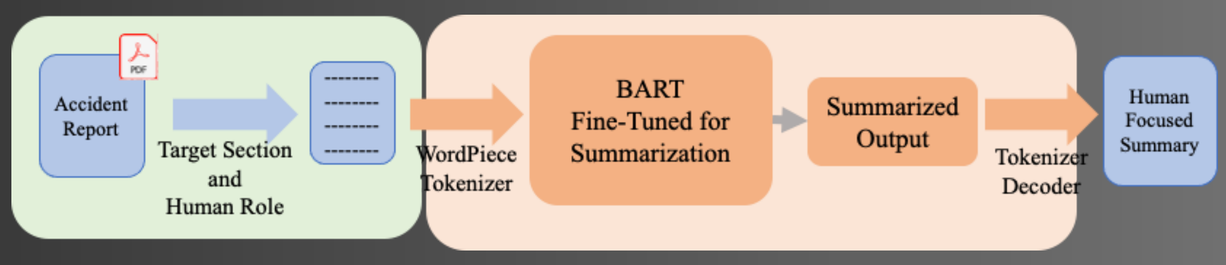


Accident Report Summarizer

It is challenging to explain and identify the computers decision making process with information extracted automatically, models based on 'black-box' machine learning, and large complicated reports/data.

Solution

Identifying the key data used by the computer, and in turn providing a summarized report of the human role that supports the model, provides decision makers with the understanding and confidence necessary.
Key target sections are identified based on a confidence scoring system, along with any sentences containing pronouns, or human related nouns. These are then input to BART, which performs abstractive summarization and outputs an overview of the human role in the incident.



Case Study

Fatality on Platform P-19 in August 2022, when CO_2 firefighting system was spuriously activated

Summarised report

- Firefighting system designed to safeguard operators, introduced additional risks due to its failure
- Safety technicians assigned to test and visually inspect the systems (not supervising maintenance)
- Rescue team unfamiliar with the plant
- Lack of communication from the issuer to the executors in the planning meetings
- Teams failure in maintenance caused leakage from generators
- ....

Additional information and software available at:

Key References

- Moura, Raphael, Michael Beer, Edoardo Patelli, John Lewis, and Franz Knoll. 'Learning from Past Accidents to Improve System Design'. Safety Science 84 (April 2016): 37–45.
- Morais, Caroline, Ka Lai Yung, Karl Johnson, Raphael Moura, Michael Beer, and Edoardo Patelli. 'Identification of Human Errors and Influencing Factors: A Machine Learning Approach'. Safety Science 146 (2022):
- Johnson, Karl, Caroline Morais and Edoardo Patelli. 'A Data Driven Approach to Elicit Causal Links between Performance Shaping Factors and Human Failure Events'. 32nd European Safety and Reliability Conference, Dublin, Ireland. (2022).
- Johnson, Karl, Caroline Morais and Edoardo Patelli. 'AI Tools for Human Reliability Analysis'. 5th ECCOMAS Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering, Athens, Greece. (2023).

Contact: Karl Johnson (karl.johnson@strath.ac.uk), Caroline Morais (cmorais@anp.gov.br). Edoardo Patelli (edoardo.patelli@strath.ac.uk)

Acknowledgment: EPSRC grant EP/T517938