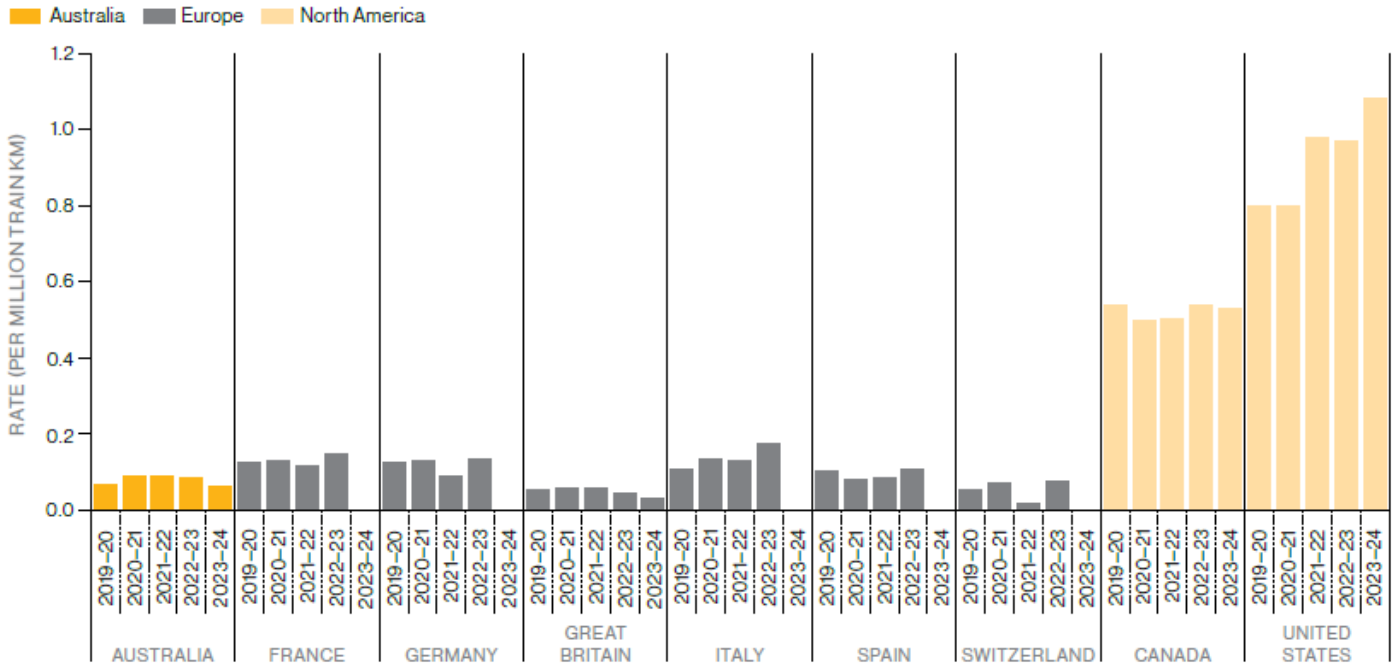
Leverage Natural Language Processing to enhance railway safety analysis

Dr Kevin Wei-Ting Hong
Lecturer in Business Analytics
Management and Organisations
UWA Business School

Nagoya U / U Tokyo - UWA Workshop
05/02/2026

THE UNIVERSITY OF WESTERN AUSTRALIA

Railway fatalities – ONRSR



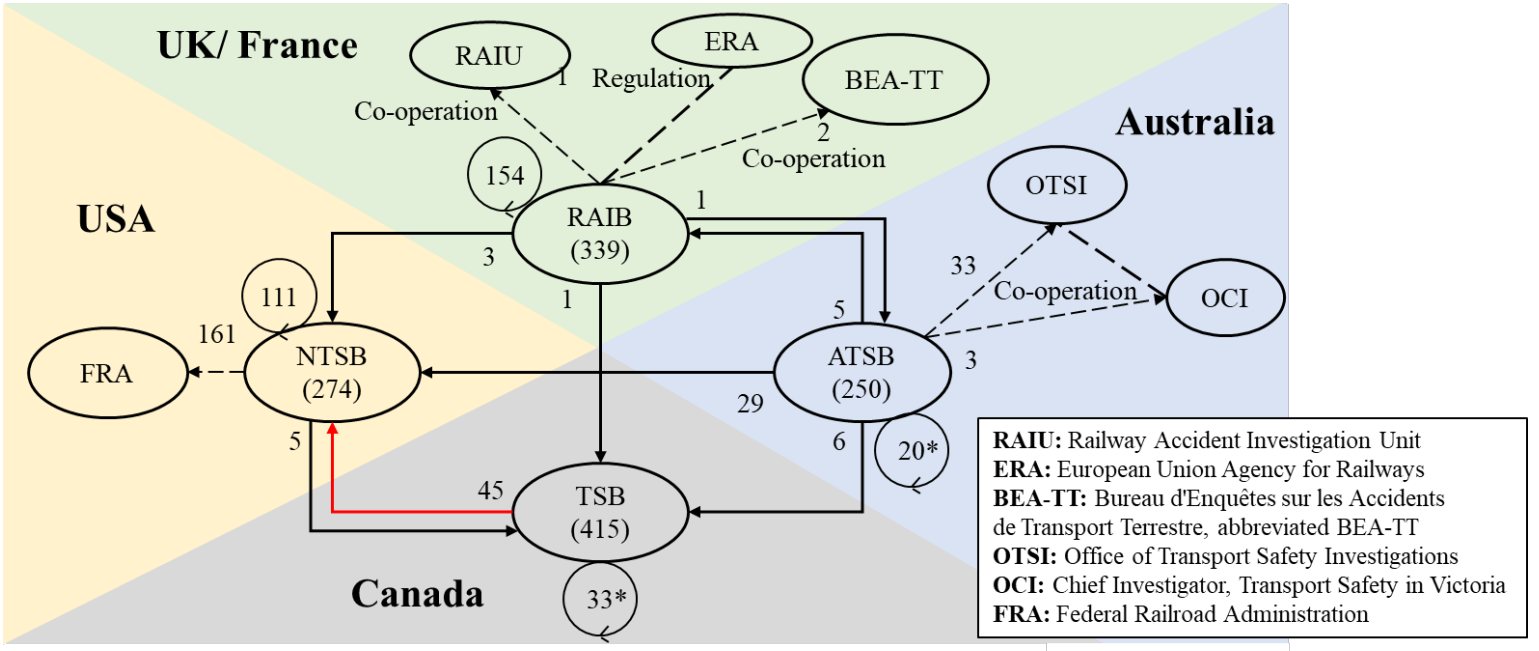
https://www.onrsr.com.au/publications/corporate-publications/rail-safety-report

Current challenges in railway safety

- Experience-based apprenticeship for knowledge transition across generations
- Potential blame culture for reliable systems
- Immediate action-oriented solutions
- Jurisdictions as the boundary of learning

Hong, W. T., Clifton, G., & Nelson, J. D. (2023). Railway accident causation analysis: Current approaches, challenges and potential solutions. Accident Analysis & Prevention, 186, 107049.

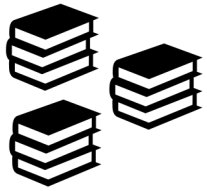
Current challenges in railway safety



* The occurrences of self-reference might be underestimated because some reports published by TSB and ATSB do not provide reference list or the self-reference is not listed in the reference.

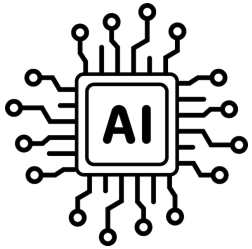
Hong, W. T., Clifton, G., & Nelson, J. D. (2025). Understanding industry learning behaviours in the context of safety: A railways perspective. Transportation Research Interdisciplinary Perspectives, 34, 101694.

Proposed solution



Existing safety knowledge

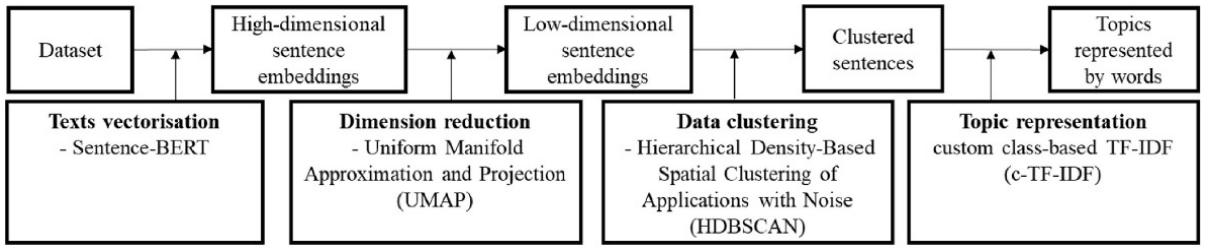
New railway accidents



AI models (local/online)

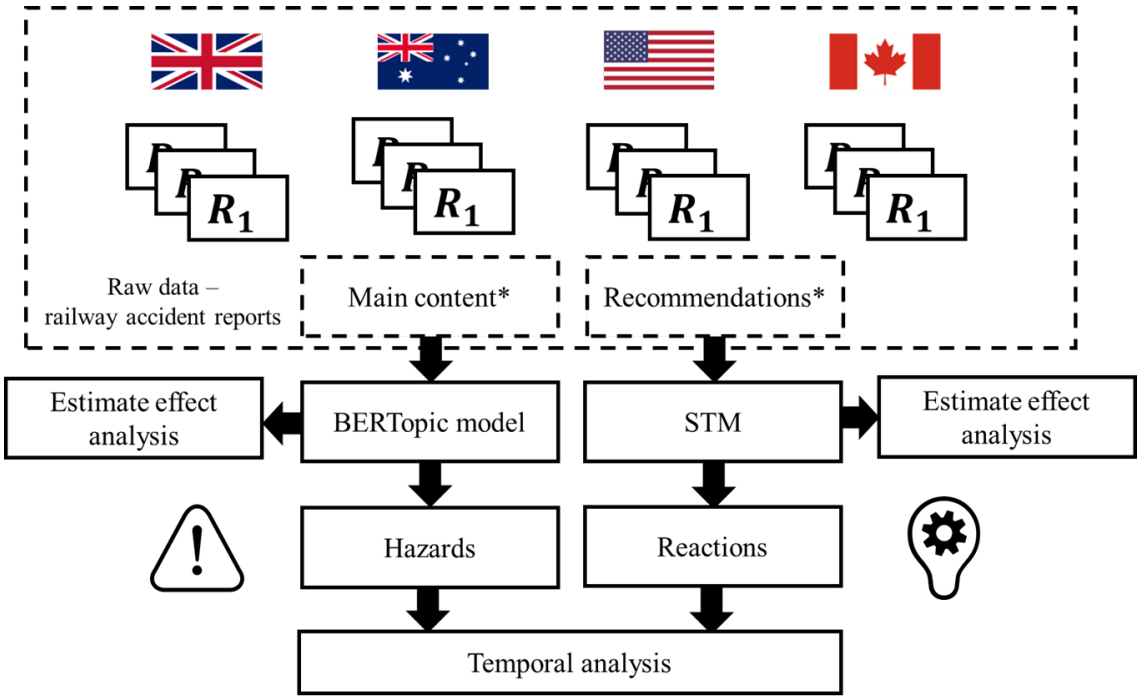
Risk analysis
Hazard identification
Auto-decision making
Safety recommendation

Aim: leveraging AI models as knowledge management systems for efficient decision-making and policymaking in the railway safety context



Hong, W. T., Clifton, G., & Nelson, J. D. (2024). A data-driven conceptual framework for understanding the nature of hazards in railway accidents. Transport policy, 152, 102-117.

Methodological framework



	No. of reports	No. of sentences	Period	Note
RAIB	339	124,990	2005–2019	Review reports are removed.
ATSB	250	84,679	1999–2021	Reports are retrieved from websites directly.
NTSB	274	92,406	1996–2021	Reports earlier than 1996 are scanned files.
TSB	415	104,720	1993–2021	Reports are retrieved from websites directly.

How much time will it take to digest this data before making a decision?

1278 hours vs 20 minutes… or even less

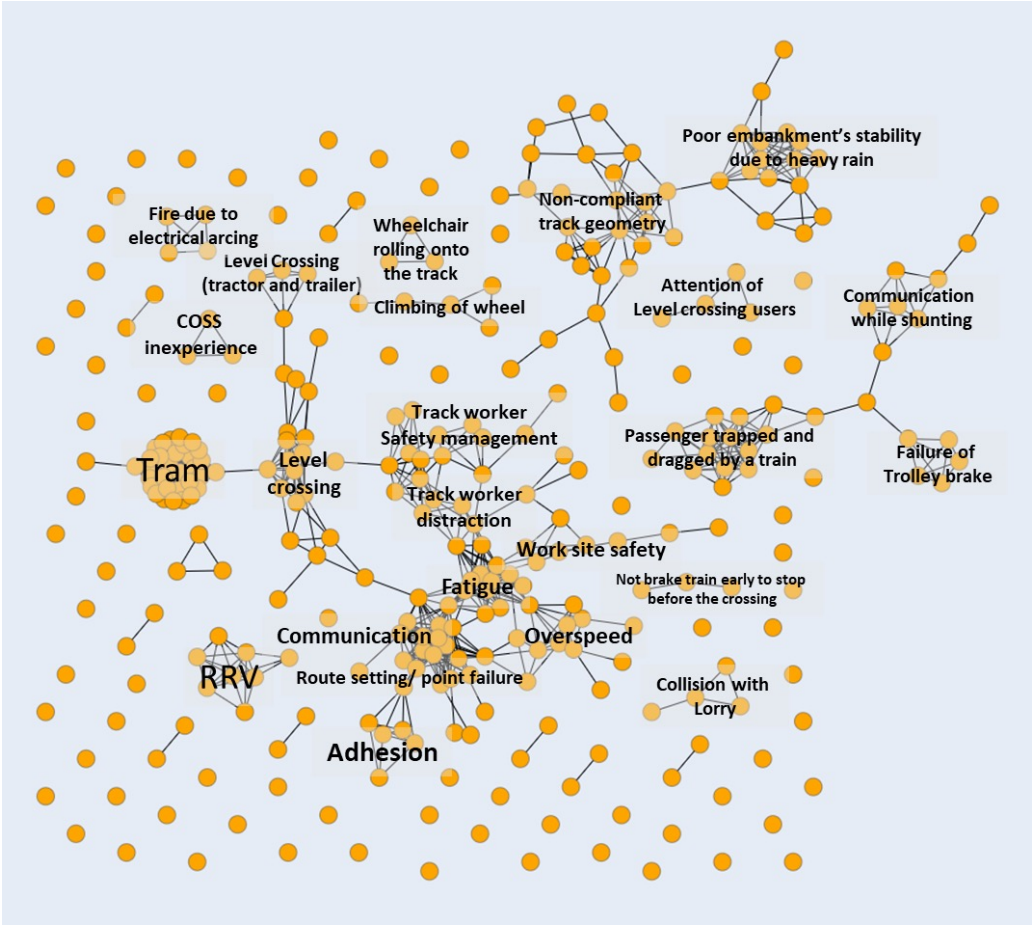
Demonstration – the UK

Data: Rail Accident Investigation Branch (RAIB)

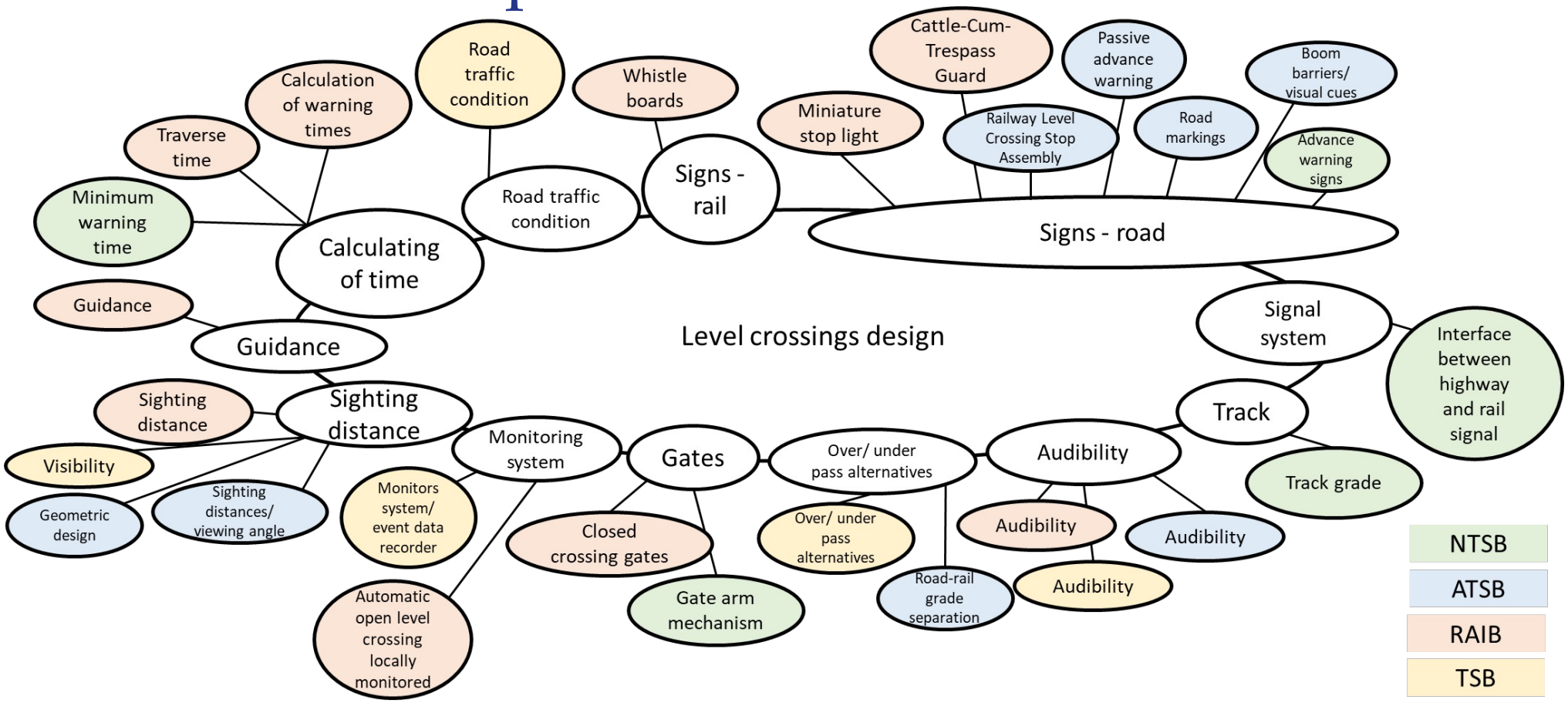
Covered period: 2005-2019

Model: BERTopic

Method: topic-topic similarity matrix



HazardMap


Level crossings design
Calculating of time
Traverse time
Calculation of warning times
Minimum warning time
Road traffic condition
Road traffic condition
Whistle boards
Signs - rail
Miniature stop light
Cattle-Cum-Trespass Guard
Passive advance warning
Railway Level Crossing Stop Assembly
Road markings
Boom barriers/ visual cues
Advance warning signs
Signs - road
Signal system
Interface between highway and rail signal
Track
Track grade
Audibility
Audibility
Audibility
Audibility
Over/ under pass alternatives
Over/ under pass alternatives
Road-rail grade separation
Gates
Closed crossing gates
Gate arm mechanism
Monitoring system
Monitors system/ event data recorder
Automatic open level crossing locally monitored
Guidance
Guidance
Sighting distance
Sighting distance
Visibility
Geometric design
Sighting distances/ viewing angle
NTSB
ATSB
RAIB
TSB

HazardMap

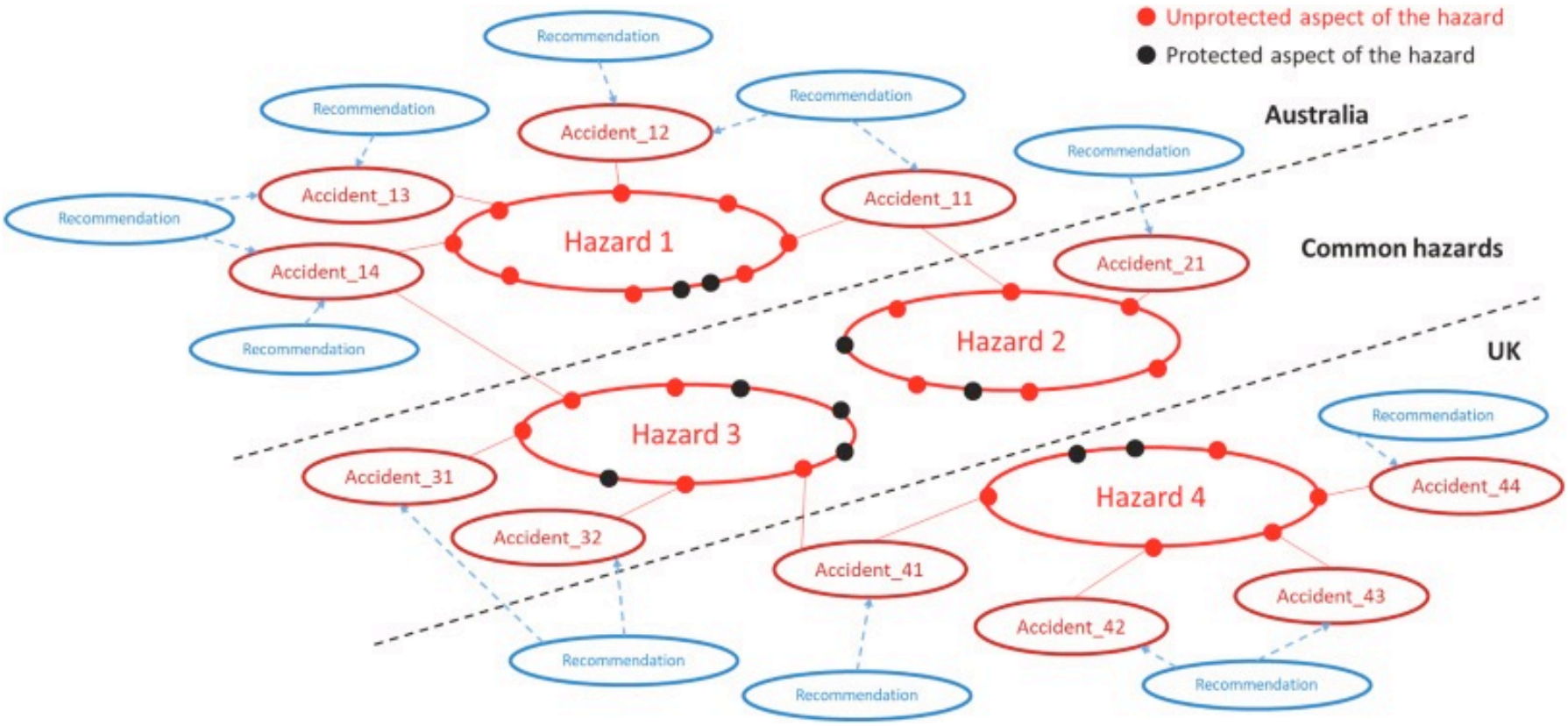


Figure 1: The illustration of HazardMap (Hong et al., 2024)

HazardMap

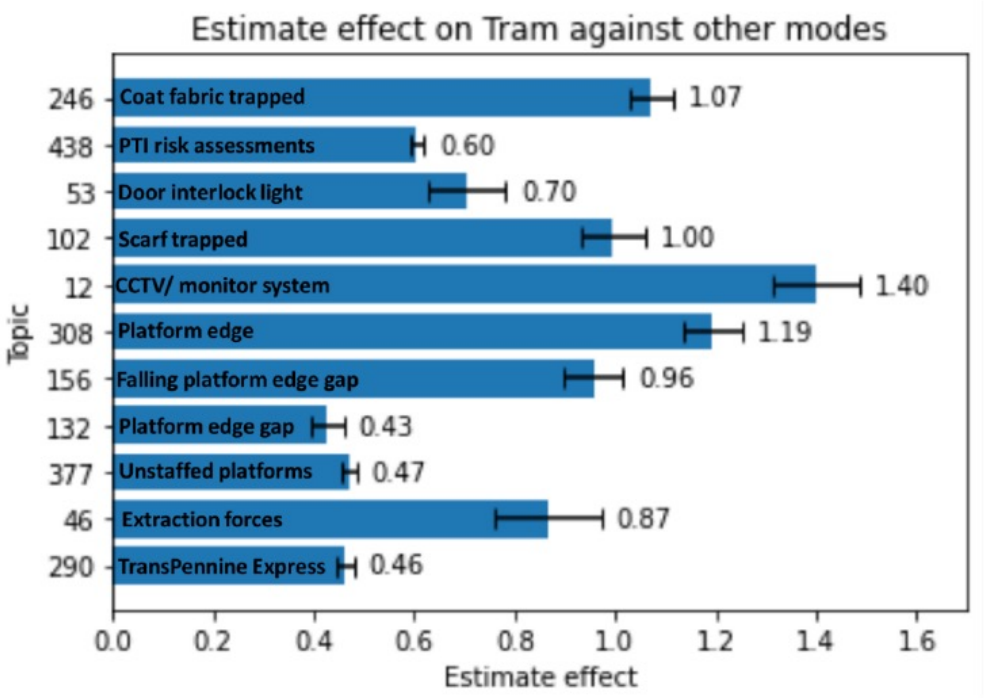


Figure 6-12: The estimate effect of core PTI hazards on trams against other modes from the RAIB dataset

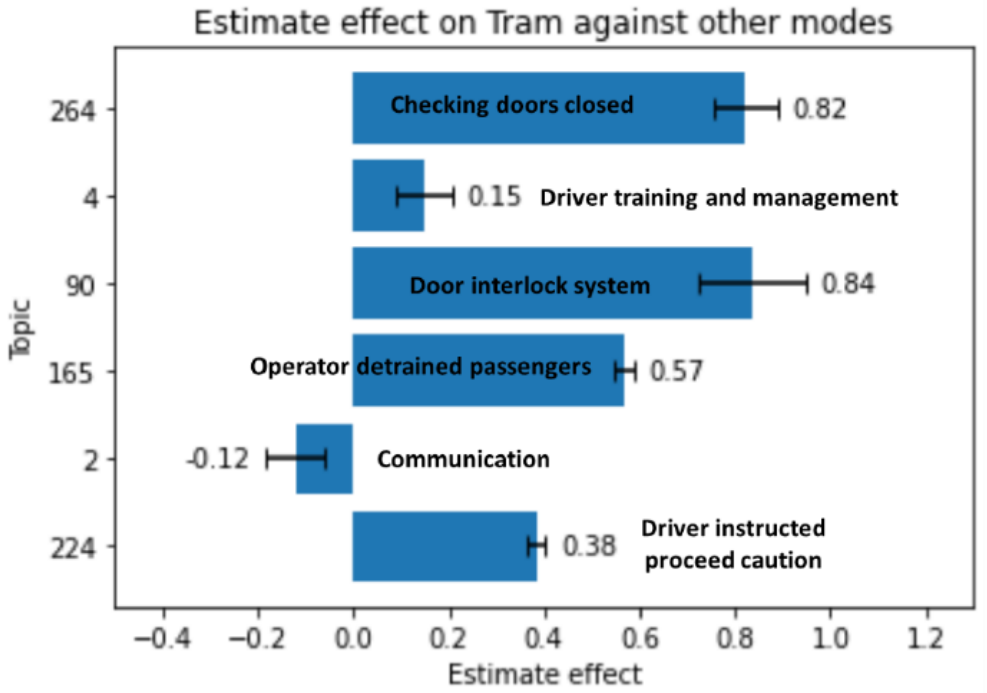
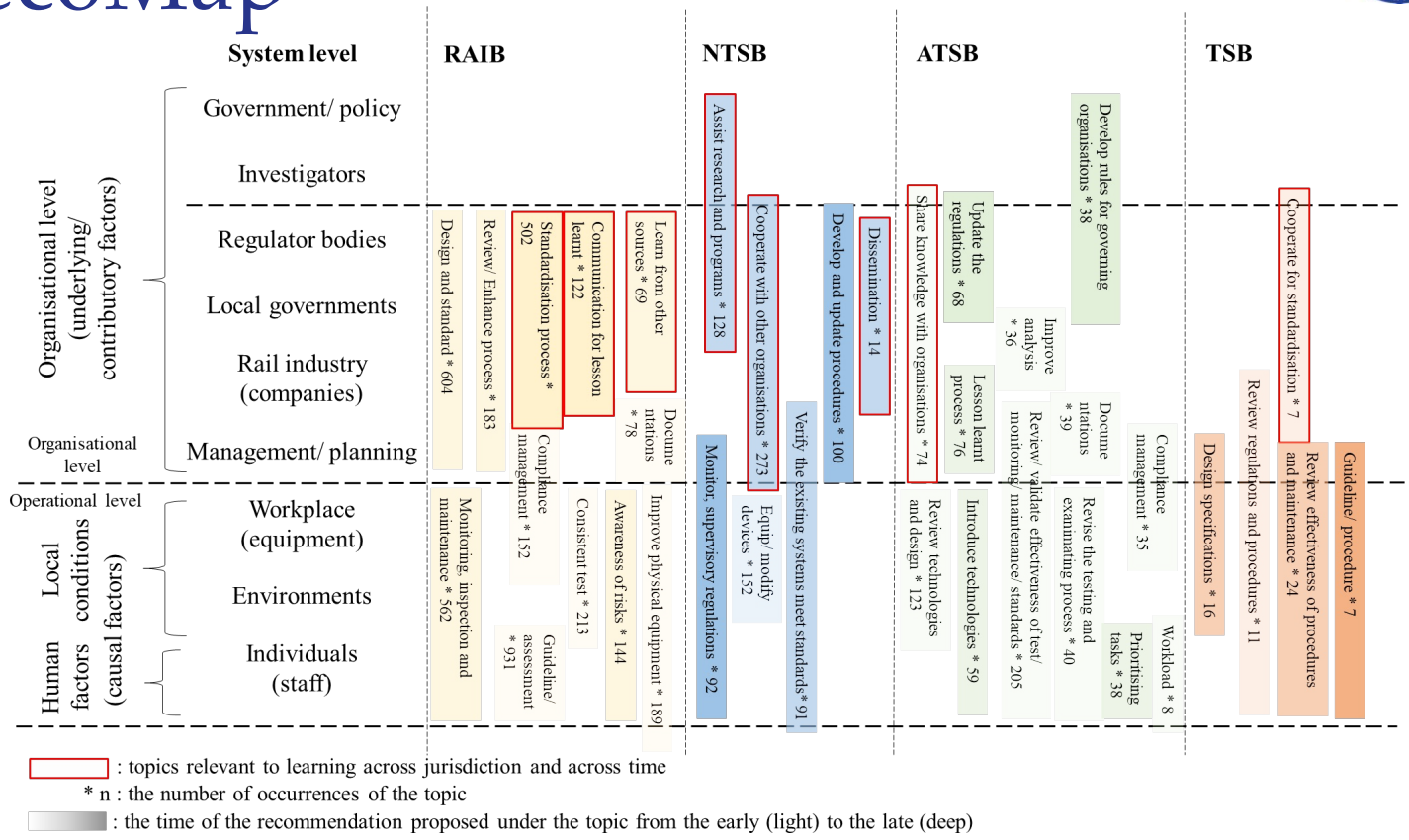


Figure 6-13: The estimate effect of supplementary PTI hazards on trams against other modes from the RAIB dataset

RecoMap



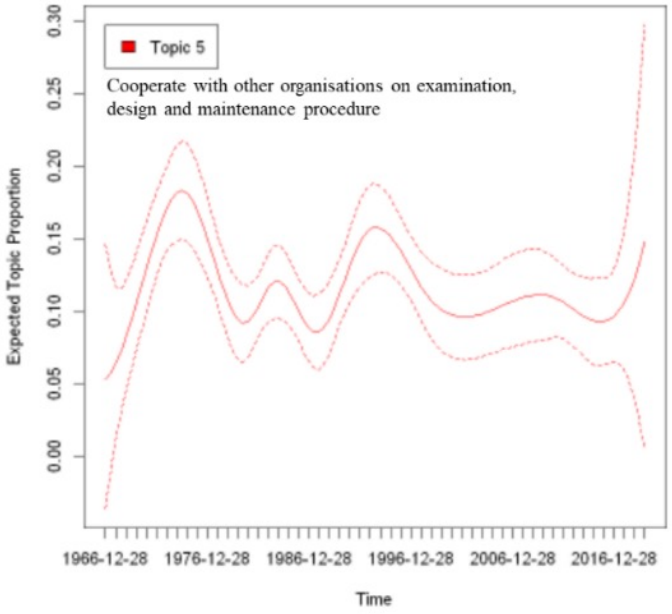
Hong, W. T., Clifton, G., & Nelson, J. D. (2025). RecoMap–a semi-automated tool for analysing railway accident recommendations across jurisdictions and over time. Transportation Planning and Technology, 1-21.

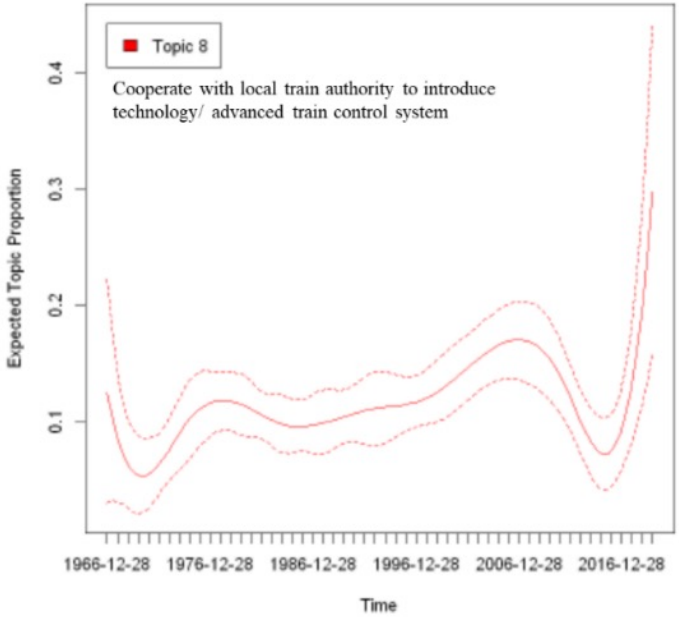
RecoMap

Recommendation type	Description	Example	Role
Assignment	Assign an objective for organisations to resolve identified hazards	Network Rail should identify and implement suitable measures to mitigate the risk of a runaway [train].	Supportive
Action	Assign specific methods to address identified hazards	Network Rail should amend its National Hazard Directory to include the access point alongside South Hampstead station.	Interfering
Reminder	Remind the compliance of existing rules or procedures	Federal Railroad Administration should increase monitoring of their employees for compliance with existing applicable rules and procedures	Supportive

	Interfering	Neutral	Supportive
Operational	ATSB (early years)	RAIB (early years) TSB (early years)	TSB (current)
Neutral Organisational	NTSB (early years) NTSB (current)	RAIB (current)	ATSB (current)

RecoMap


Topic 5
Cooperate with other organisations on examination, design and maintenance procedure
Expected Topic Proportion
0.00
0.05
0.10
0.15
0.20
0.25
0.30
1966-12-28
1976-12-28
1986-12-28
1996-12-28
2006-12-28
2016-12-28
Time


Topic 8
Cooperate with local train authority to introduce technology/ advanced train control system
Expected Topic Proportion
0.1
0.2
0.3
0.4
1966-12-28
1976-12-28
1986-12-28
1996-12-28
2006-12-28
2016-12-28
Time

Current limitations and future direction

Limitation

- Validating the outputs from NLP-based models remained challenging
- Collaborating with AI in the high-reliability industry decision-making requires more evidence and empirical experiments
- Data sensitivity and regime

Future work

- Leverage statistical approaches to design evaluation methods
- More small and medium-scale implementations across different industries and government agencies
- Build a knowledge management system to support expert-AI collaboration framework in railway safety decision making

THE UNIVERSITY OF WESTERN AUSTRALIA

Thank you!

Dr Kevin Wei-Ting Hong

wei-ting.hong@uwa.edu.au