

UNIVERSITY OF PRETORIA



The Intersection of Transportation Engineering and Data Science

Rozina Myoya

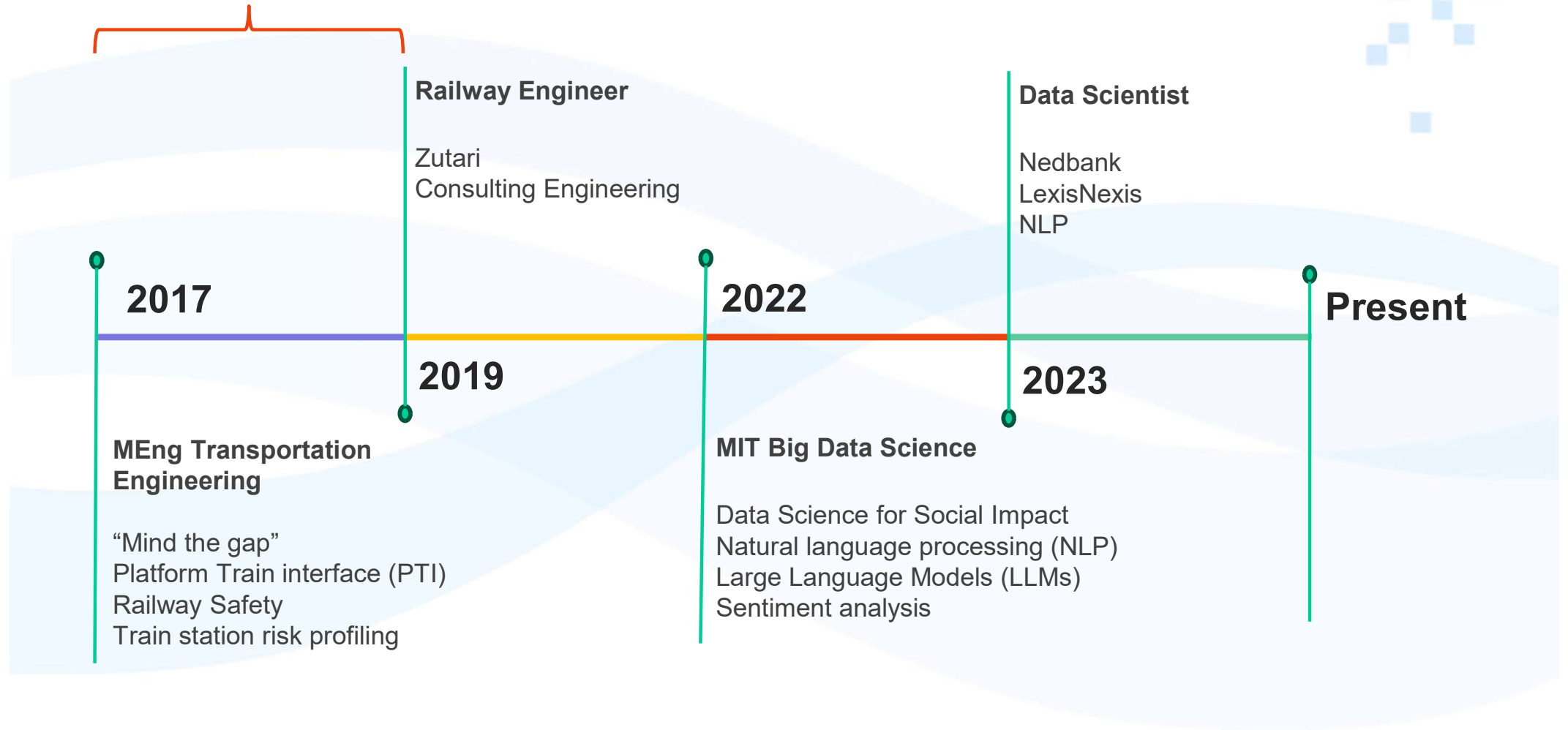


UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

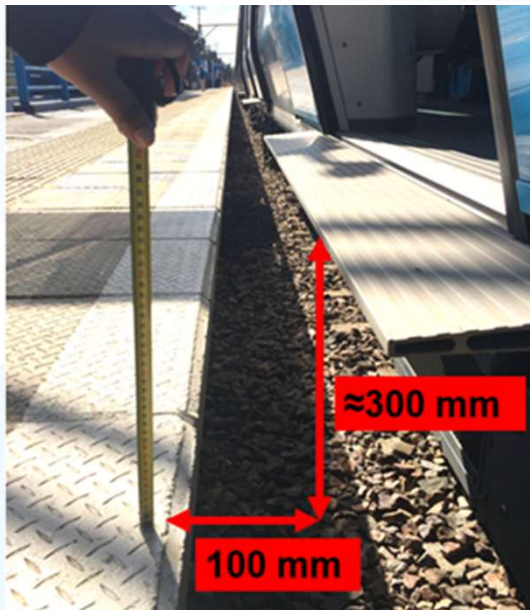
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Research Timeline

Phase 1



Research (2017 – 2019)



Unreliable Train
Schedule

Overcrowding

Train surfing
(Staff riding)

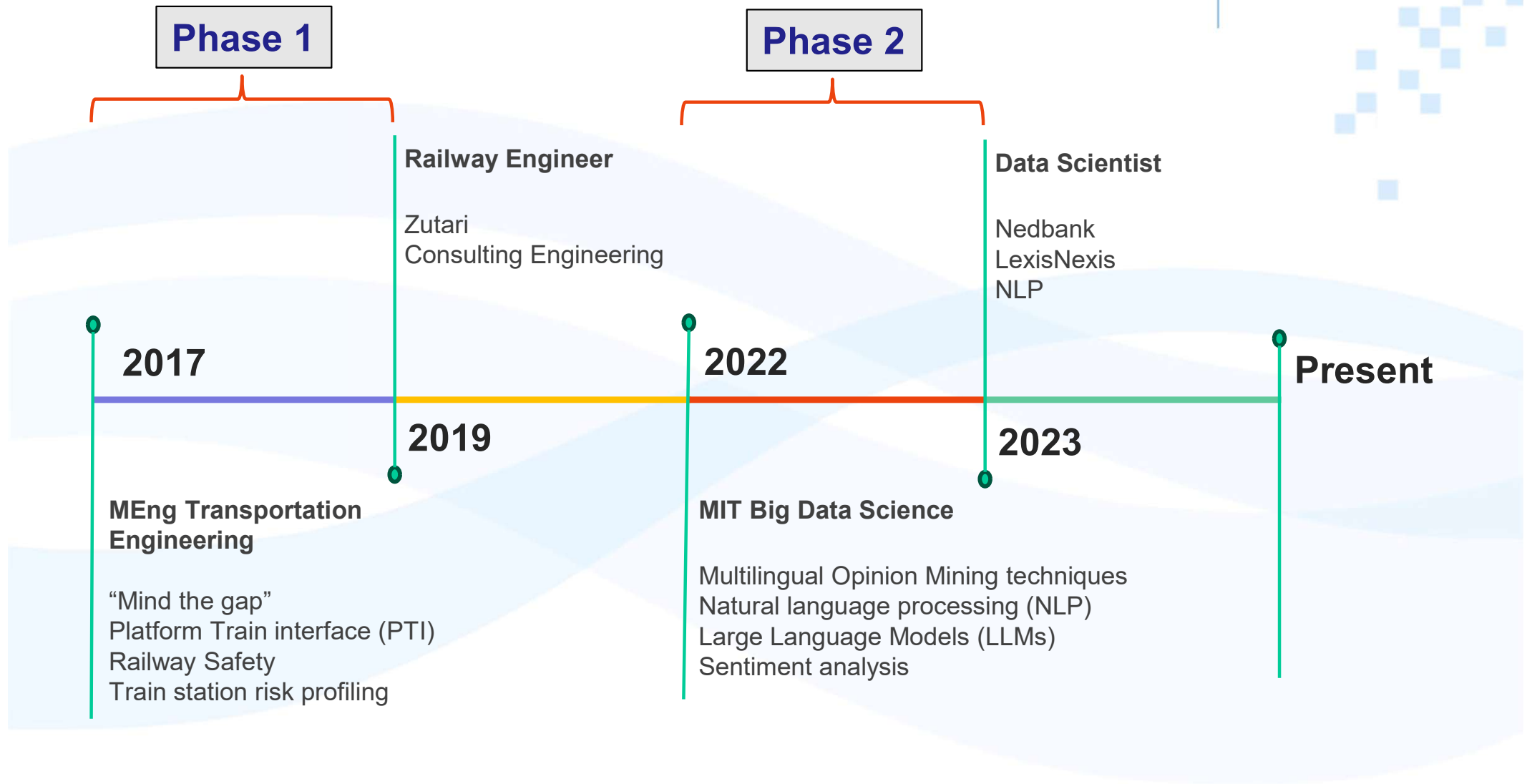
Harassment

Availability of
Toilets

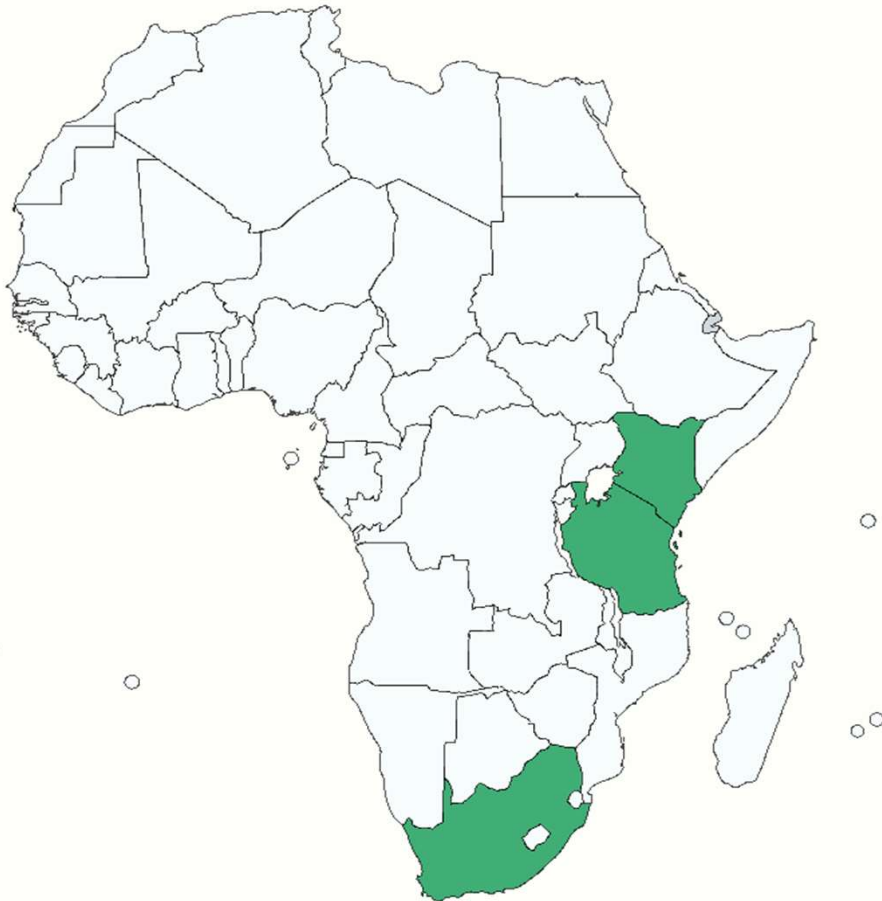
User Experience



Research Timeline



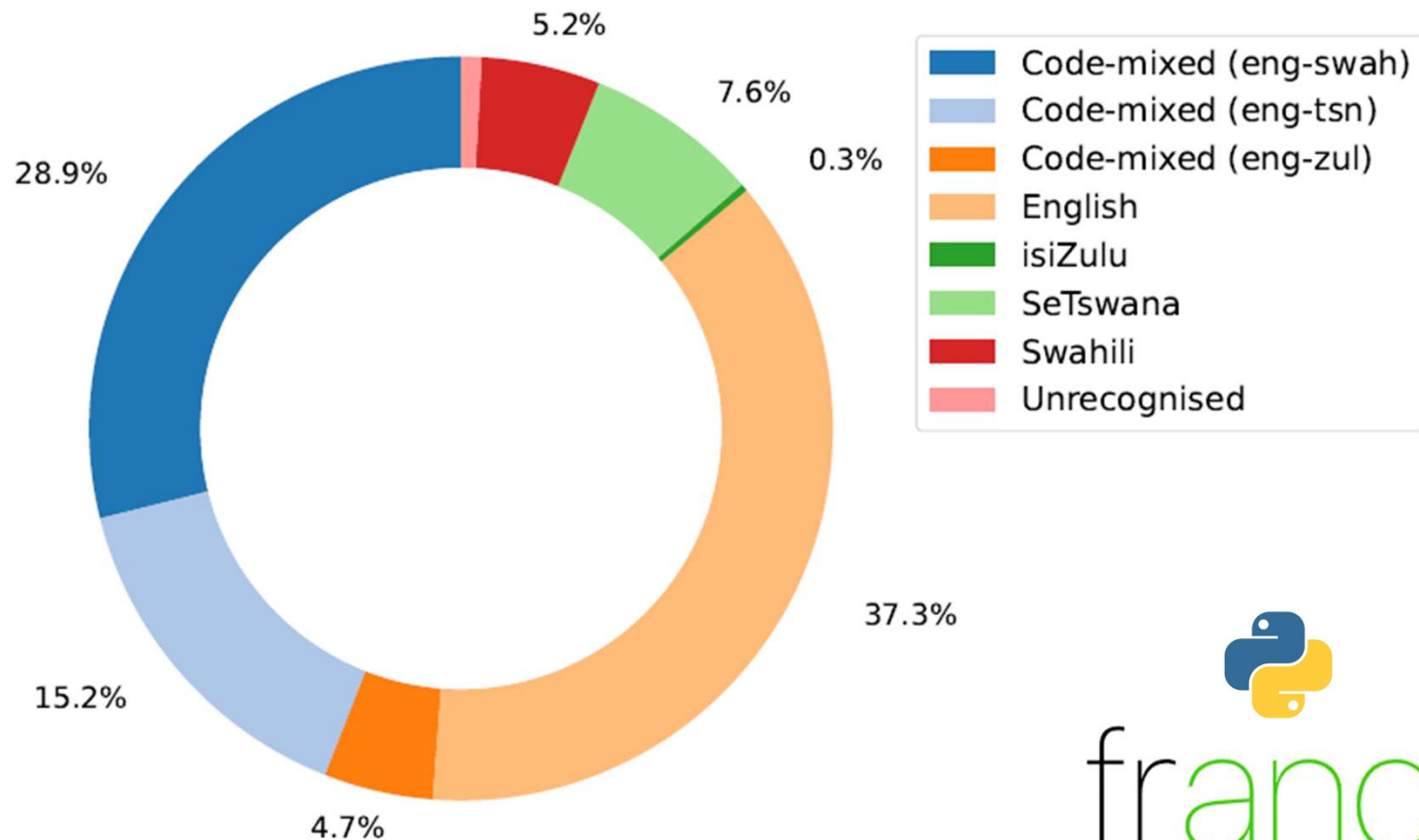
Research (2022 – 2023)



This study explored commuter sentiments on public transport using real-time X (Twitter) data, focusing on rail, mini-bus taxis, and buses. It employed Multilingual Opinion Mining techniques to navigate linguistic diversity and code-switching challenges.

The objective of the research was to provide insight into the **public transport user experience** to identify the main commuter **pain points** and thus **propose possible solutions** beyond infrastructural modifications.

Research (2022 – 2023) Language Identification



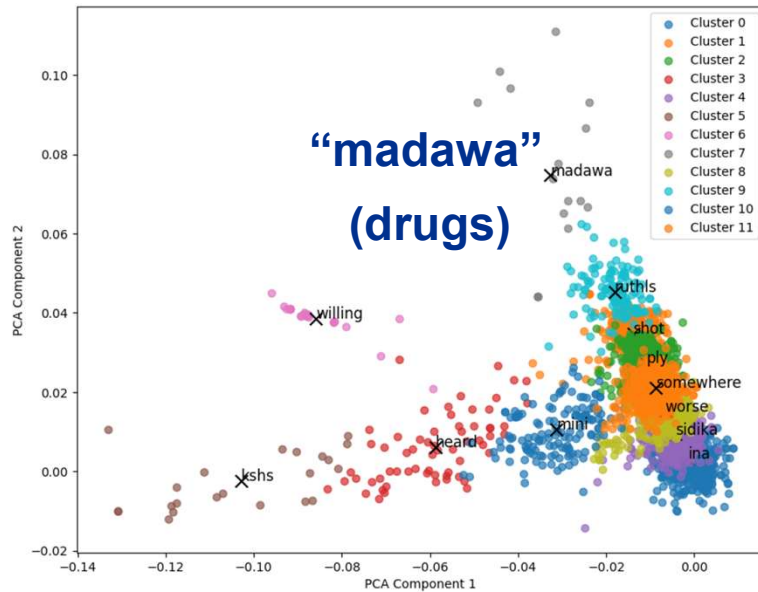
```
from pyfranc import franc
```



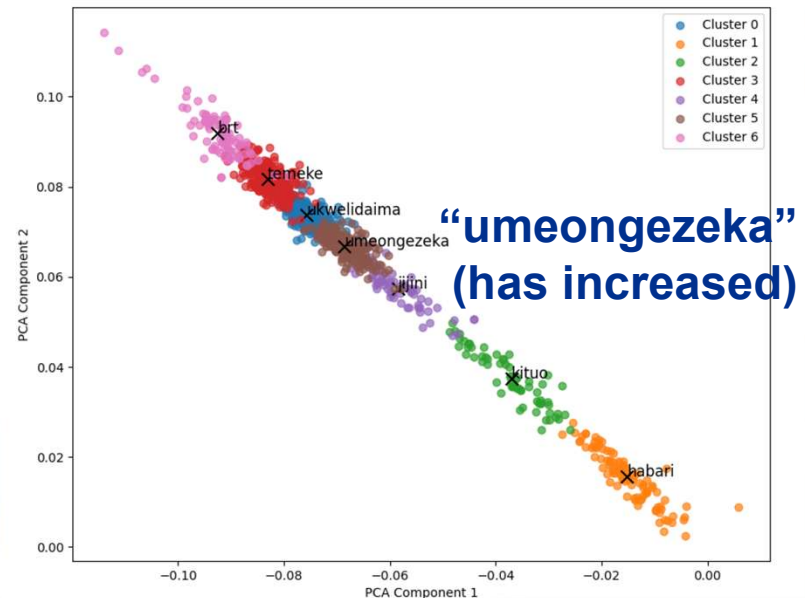
Research (2022 – 2023)

Feature Extraction

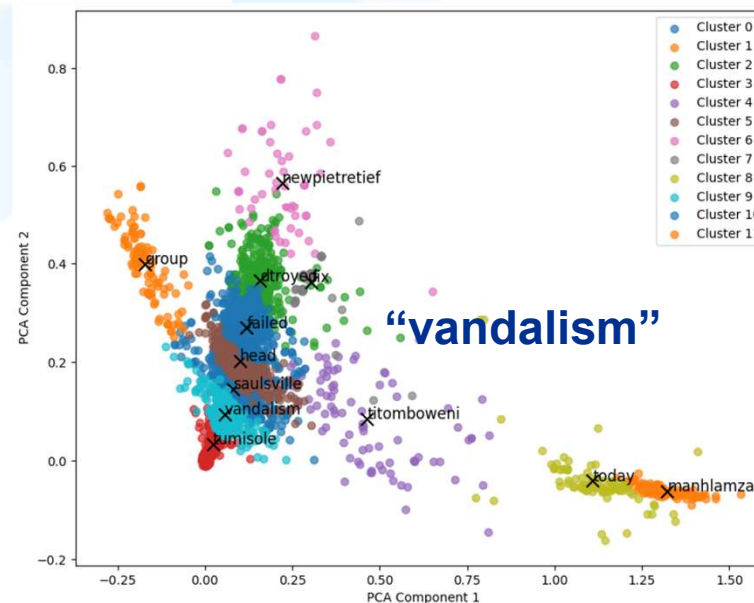
Kenyan Dataset



Tanzanian Dataset



South African Dataset



Word2Vec 



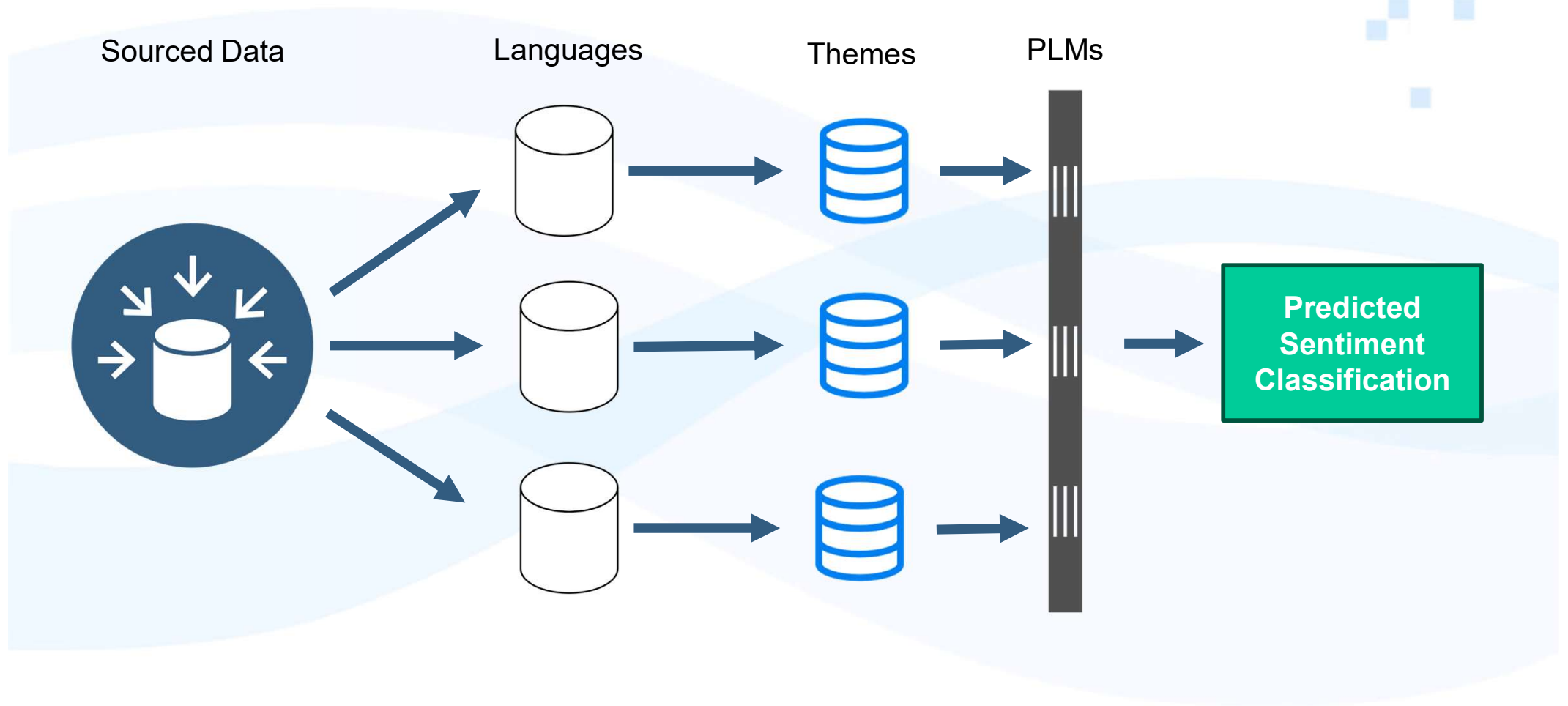
Model Testing and Evaluation

Table 3: Summary of the PLM details

Model	Size	Languages included
AfriBERTa	126M	Afaan Oromoo, Amharic, Gahuza, Hausa, Igbo, Nigerian Pidgin, Somali, Swahili , Tigrinya and Yorùbá
AfroXLMR	550M	Afrikaans, Amharic, Hausa, Igbo, Malagasy, Chichewa, Sesotho, Oromo, Nigerian-Pidgin, Kinyarwanda, Kirundi, Shona, Somali, Swahili , isiXhosa, Yoruba, isiZulu , Arabic, French, and English
AfroLM	264M	Amharic, Afan Oromo, Bambara, Ghomálá, Éwé, Fon, Hausa, Igbo, Kinyarwanda, Lingala, Luganda, Luo, Mooré, Chewa, Naija, Shona, Swahili , SeTswana , Twi, Wolof, Xhosa, Yorùbá, and isiZulu
PuoBERTa	83.5M	SeTswana

- K. Ogueji, Y. Zhu, and J. Lin. Small data? no problem! exploring the viability of pretrained multilingual language models for low-resourced languages. In Proceedings of the 1st Workshop on Multilingual Representation Learning, pages 116–126, 2021.
- J. O. Alabi, D. I. Adelani, M. Mosbach, and D. Klakow. Adapting pretrained language models to african languages via multilingual adaptive finetuning. In Proceedings of the 29th International Conference on Computational Linguistics, pages 4336–4349, 2022.
- B. F. P. Dossou, A. L. Tonja, O. Yousuf, S. Osei, A. Oppong, I. Shode, O. O. Awoyomi, and C. C. Emezue. Afrolm: A self-active learning-based multilingual pretrained language model for 23 african languages, 2022. URL <https://arxiv.org/abs/2211.03263>.
- V. Marivate, M. Mots’Oehli, V. Wagner, R. Lastrucci, and I. Dzingirai. Puoberta: Training and evaluation of a curated language model for setswana. ArXiv, 2023

Sentiment Analysis Results



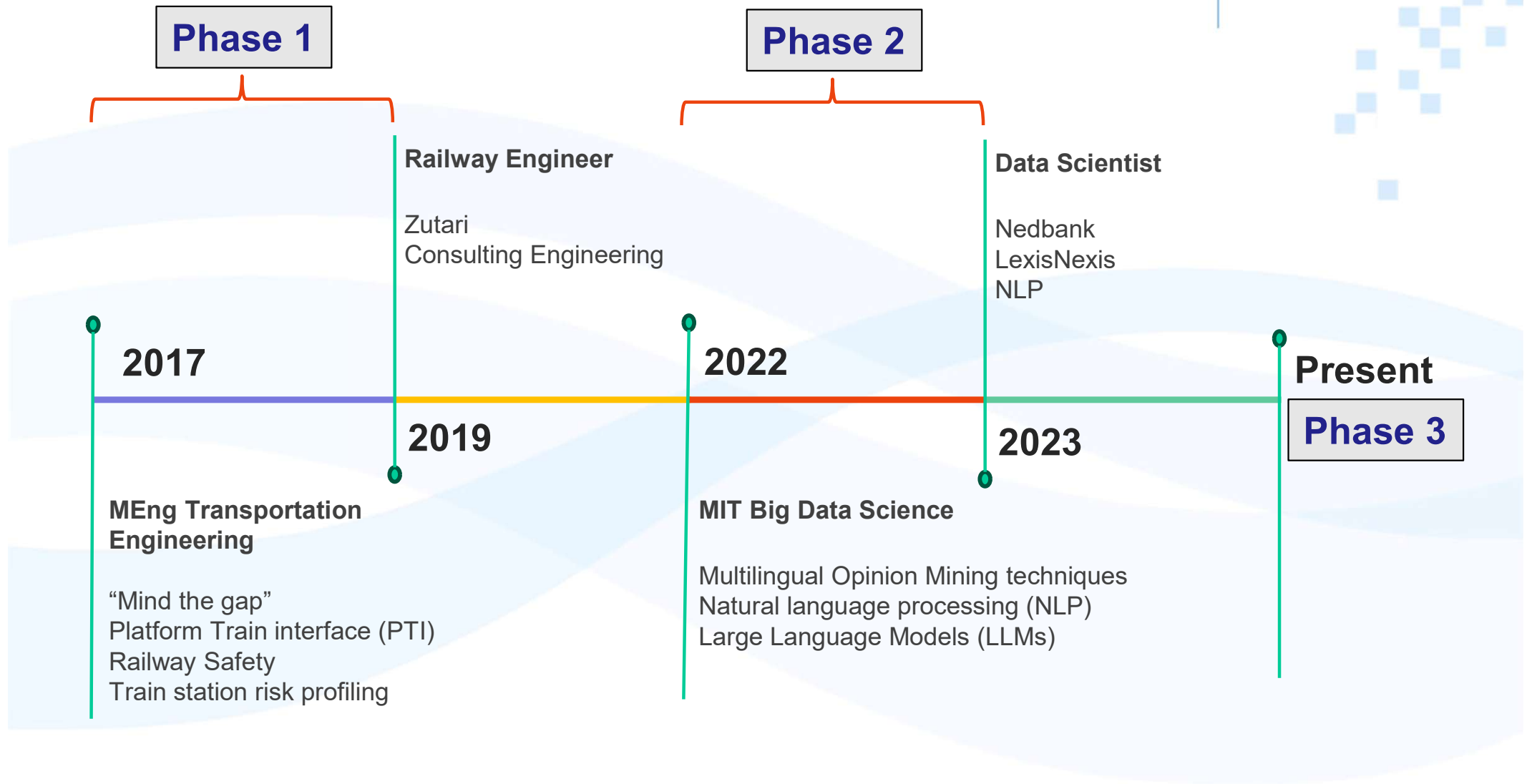


Safety

Price

Government

Research Timeline



Research (2017 – 2019)

J2 : ✕ ✓ fx A male person tried to board moving metro train no. 9013 (38 Set) and fell under the train. He sustained fatal injuries and was taken to Salt River Morgue.

	A	B	C	D	E	F	G	H	I	J
1	FIN YE	Operator	Date Reported	Time	Date of Occurrence	Day	Province	Place	Category	Description
2	14/15	PRASA RAIL	04 February 2014	7:31	41730,00	Tuesday	Western Cape	Cape Town	H-a	A male person tried to board moving metro train no. 9013 (38 Set) and
3	14/15	PRASA RAIL	16 April 2014	11:50	01 April 2014	Tuesday	KwaZulu-Natal	Charlottedale	H-b	Amanda Zungu alighted from metro train 0245 whilst in motion and sl
4	14/15	PRASA RAIL	16 April 2014	19:50	01 April 2014	Tuesday	KwaZulu-Natal	Pelgrim	H-b	A teenage boy jumped off metro train 0788 whilst in motion and fell o
5	14/15	PRASA RAIL	04 March 2014	21:05	01 April 2014	Tuesday	KwaZulu-Natal	Hammersdale	H-b	Unknown female opened the doors whilst Metro Train 1088 was in mo
6	14/15	PRASA RAIL	25 April 2014	10:03	02 April 2014	Wednesday	Gauteng	Springs	H-b	Ms T Dambuza board metro train 0042 while the train started to move
7	14/15	PRASA RAIL	25 April 2014	15:40	02 April 2014	Wednesday	Gauteng	Johannesburg	H-b	Ms C Ramatsubana was standing between open doors of moving metro
8	14/15	PRASA RAIL	16 April 2014	4:30	03 April 2014	Thursday	KwaZulu-Natal	Kwa Mashu	H-b	Mr. Z. Khumalo attempted to board moving metro train 9158 as a resu
9	14/15	PRASA RAIL	16 April 2014	7:57	04 April 2014	Friday	KwaZulu-Natal	Berea	H-b	Ms Lerato Yengwa was alighted from metro train 0734 whilst the train
10	14/15	PRASA RAIL	25 April 2014	16:30	05 April 2014	Saturday	Gauteng	Dunswart	H-a	Mr S Qeswa was disembarking the metro train 0042, his bag stucked b
11	14/15	PRASA RAIL	25 April 2014	4:45	07 April 2014	Monday	Gauteng	Nancefield	H-b	Ms L Mabube was trying to board metro train 9005 in motion and fell

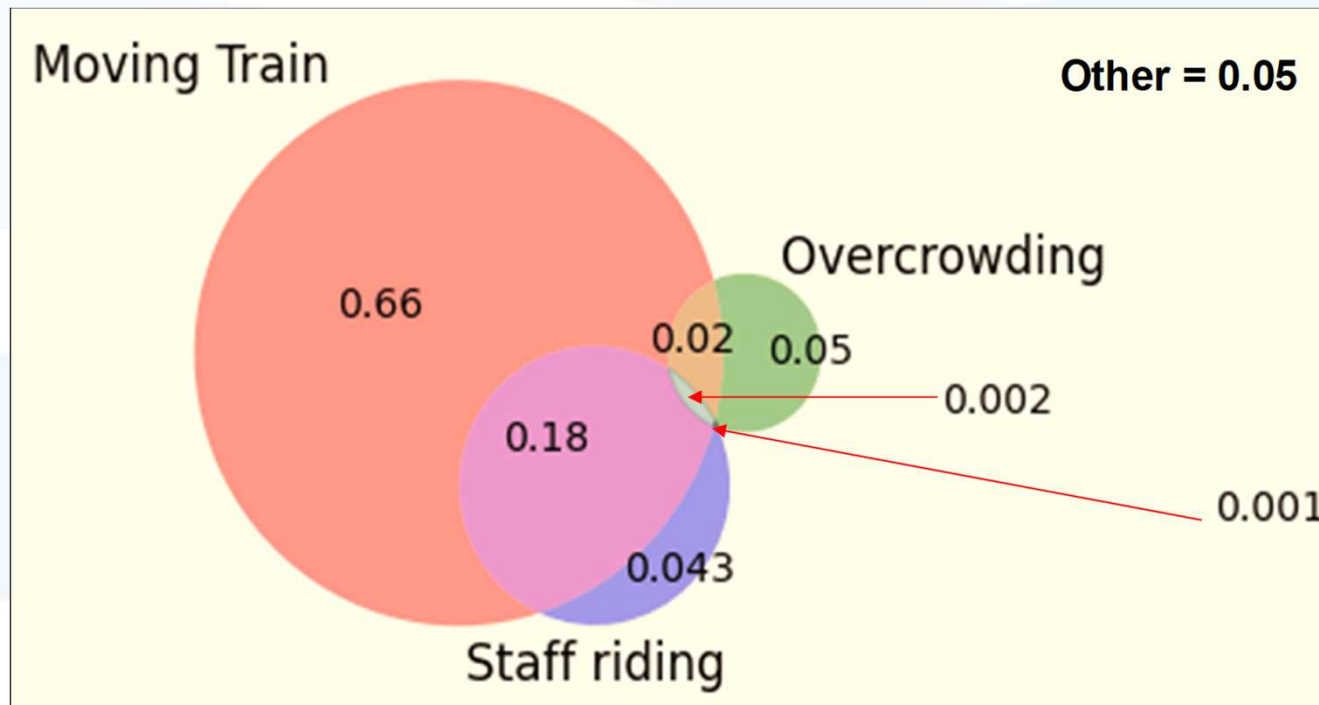
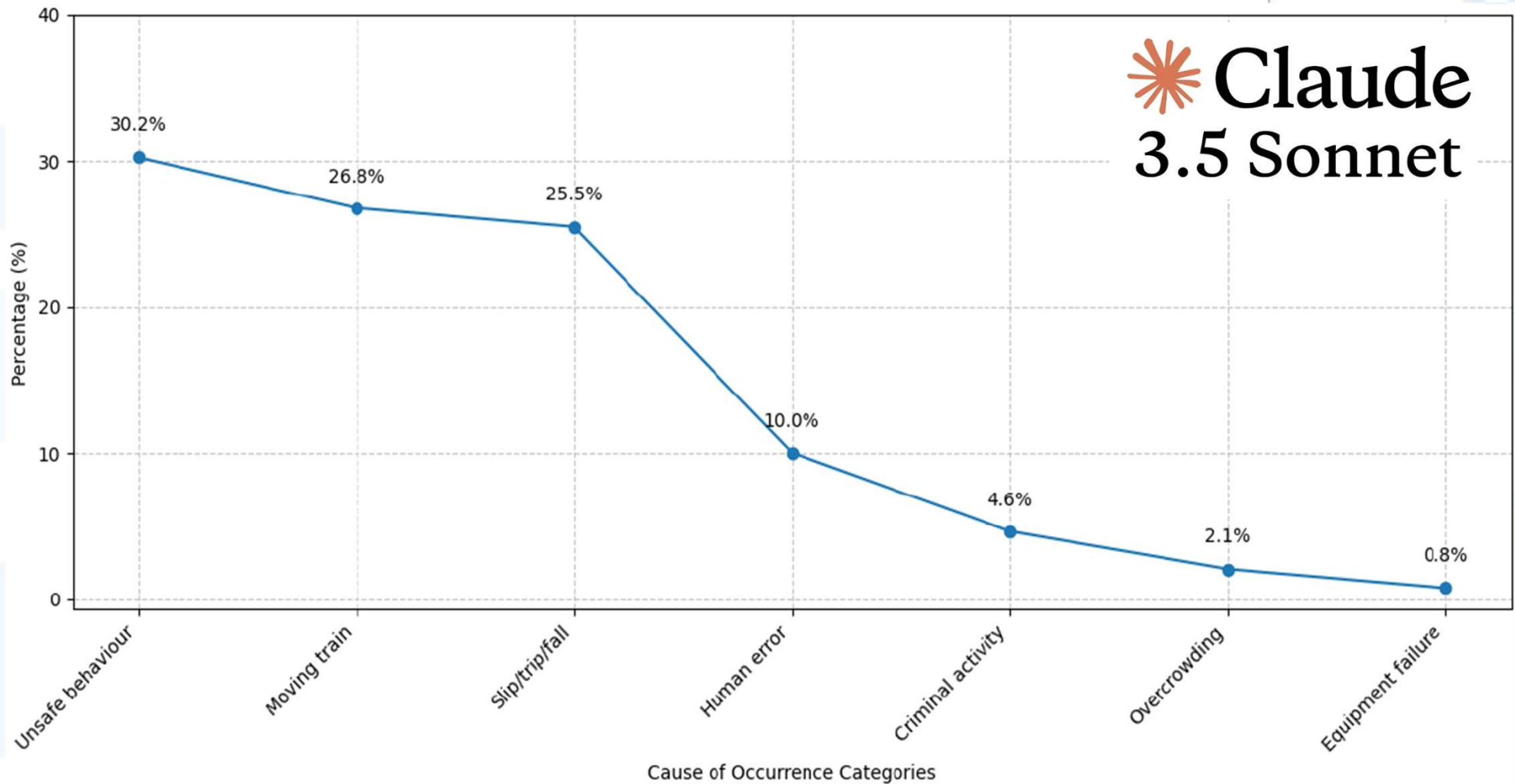


Fig 12: PTL occurrences per 1000 passengers versus gap sizes measured larger than 200mm

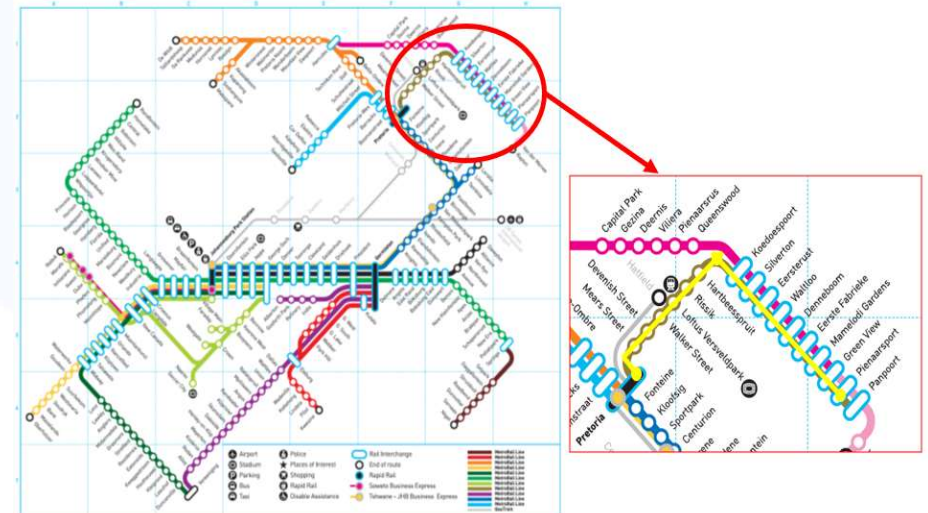
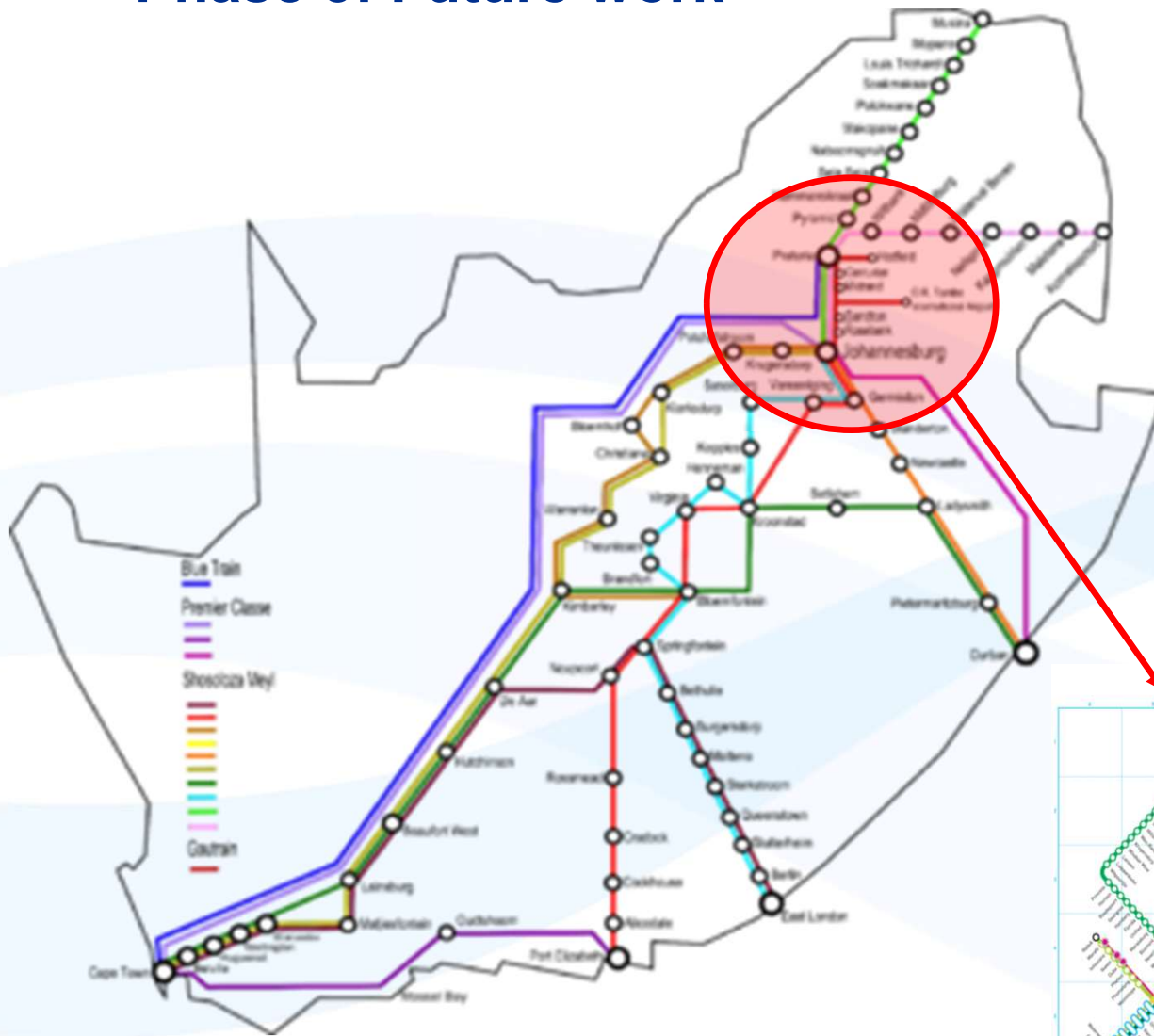


Phase 3: Large Language Model (LLM) insights



Phase 3: Future work

- Incident Detection
- Information Gap
- User-centred design



Conclusions

1. Transport safety issues are complex and multifaceted
2. User experience is central to transport effectiveness
3. Multilingual NLP and social media mining are powerful tools
4. AI and LLMs significantly improve analysis
5. There is potential for real-time, intelligent monitoring systems
6. User-centred research drives smarter transport planning



Thank you!

Rozina Myoya

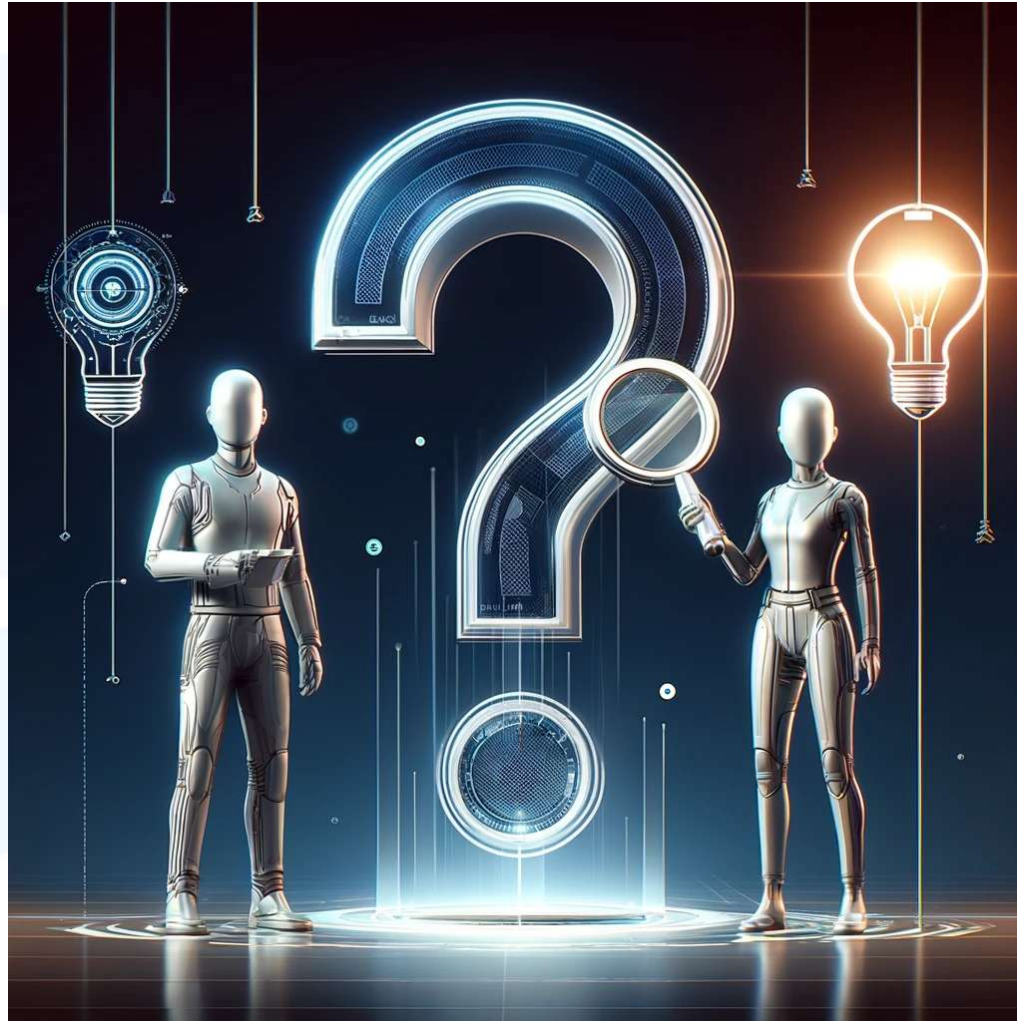
Email: Rozina.Myoya@tuks.co.za

Publications: <https://linktr.ee/codewithroz>



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dinlalefi

Any Questions?



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Denkleiers • Leading Minds • Dikgopolo tša Dihalefi