INTERNATIONAL RAILWAY SAFETY COUNCIL 2022

IRSC Spain

IRSC 2022

INTERNATIONAL RAILWAY
SAFETY COUNCIL

SEVILLA, OCTOBER 16-21, 2022

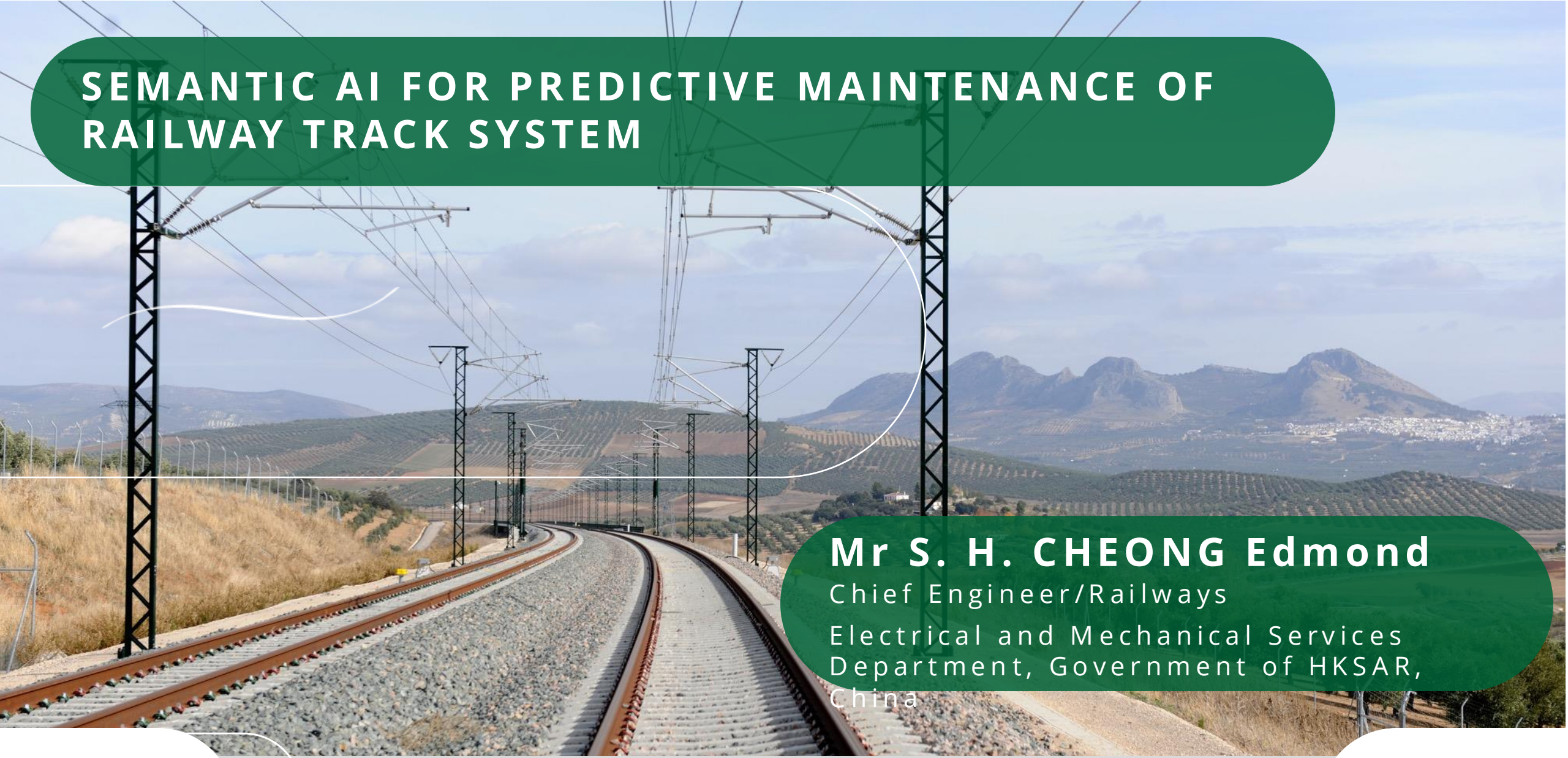
機電工程署
EMSD

EUROPEAN UNION AGENCY FOR RAILWAYS

AESF AGENCIA ESTATAL DE SEGURIDAD FERROVIARIA

adif

SEMANTIC AI FOR PREDICTIVE MAINTENANCE OF RAILWAY TRACK SYSTEM



Mr S. H. CHEONG Edmond

Chief Engineer/Railways

Electrical and Mechanical Services Department, Government of HKSAR, China

IRSC 2022

INTERNATIONAL RAILWAY SAFETY COUNCIL

SEVILLA, OCTOBER 16-21, 2022

AGENDA

Part 1

Background

Part 2

Semantic Artificial Intellgence (AI) Modelling

Part 3

Results

Part 4

Conclusion

Part 1
Background

Background

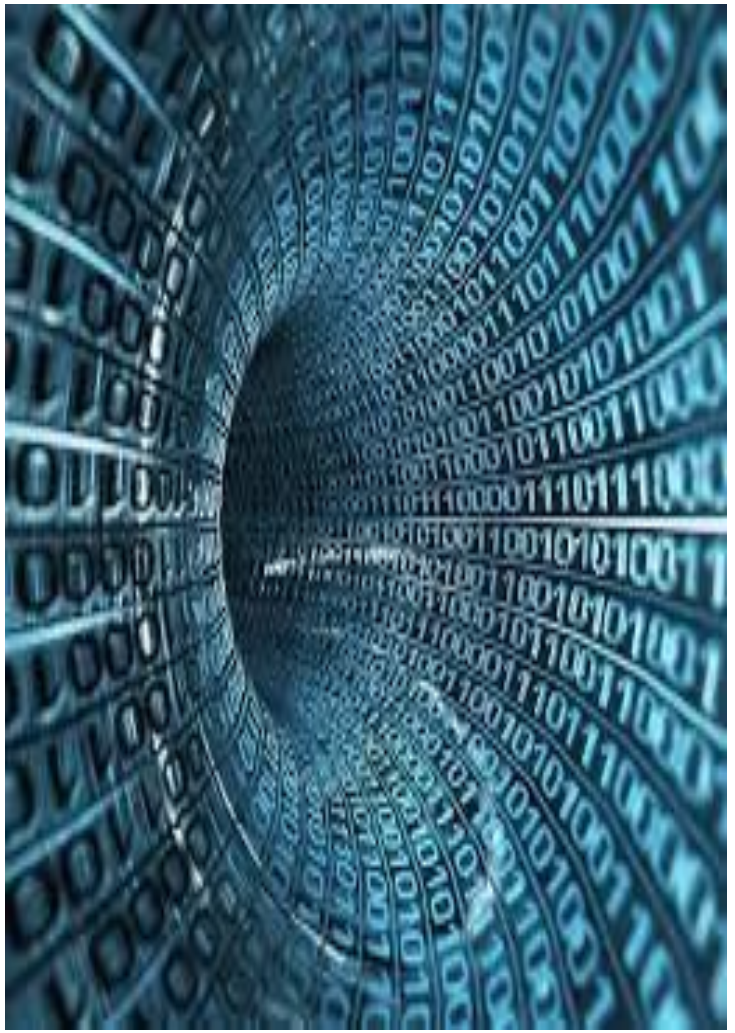
Railway Network of Hong Kong, China

- Over 5 million passenger journeys are made each day
- Very long daily service hours
- Extremely short maintenance window
- Numerous railway assets required maintenance
- Need an effective way to minimize breakdown maintenance and incidents

Background

Data-driven Era

- Wide range of data available
  - Static data
  - Dynamic data
  - Online data
- Extraction of knowledge/ information from data
- Develop Artificial Intelligence (AI) model for decision making / prediction

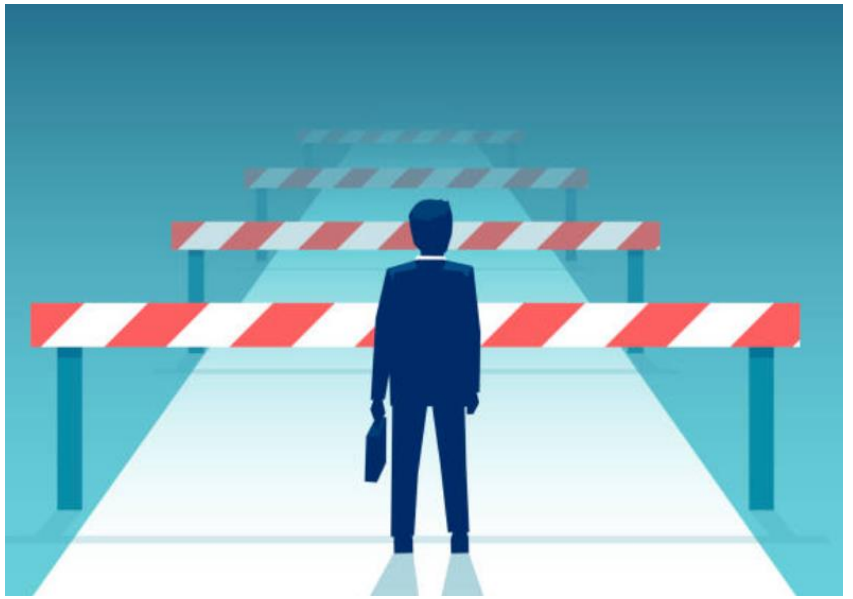


Part 2
Semantic Artificial Intelligence (AI) Modelling

Semantic AI Modelling

Challenges

- Massive / Missing data
- Data in different types and formats
- Missing of correlation among data
- Resources and time consuming for data standardization and cleansing
- Expert knowledge required to transform data into structured knowledge/information



Semantic AI Modelling

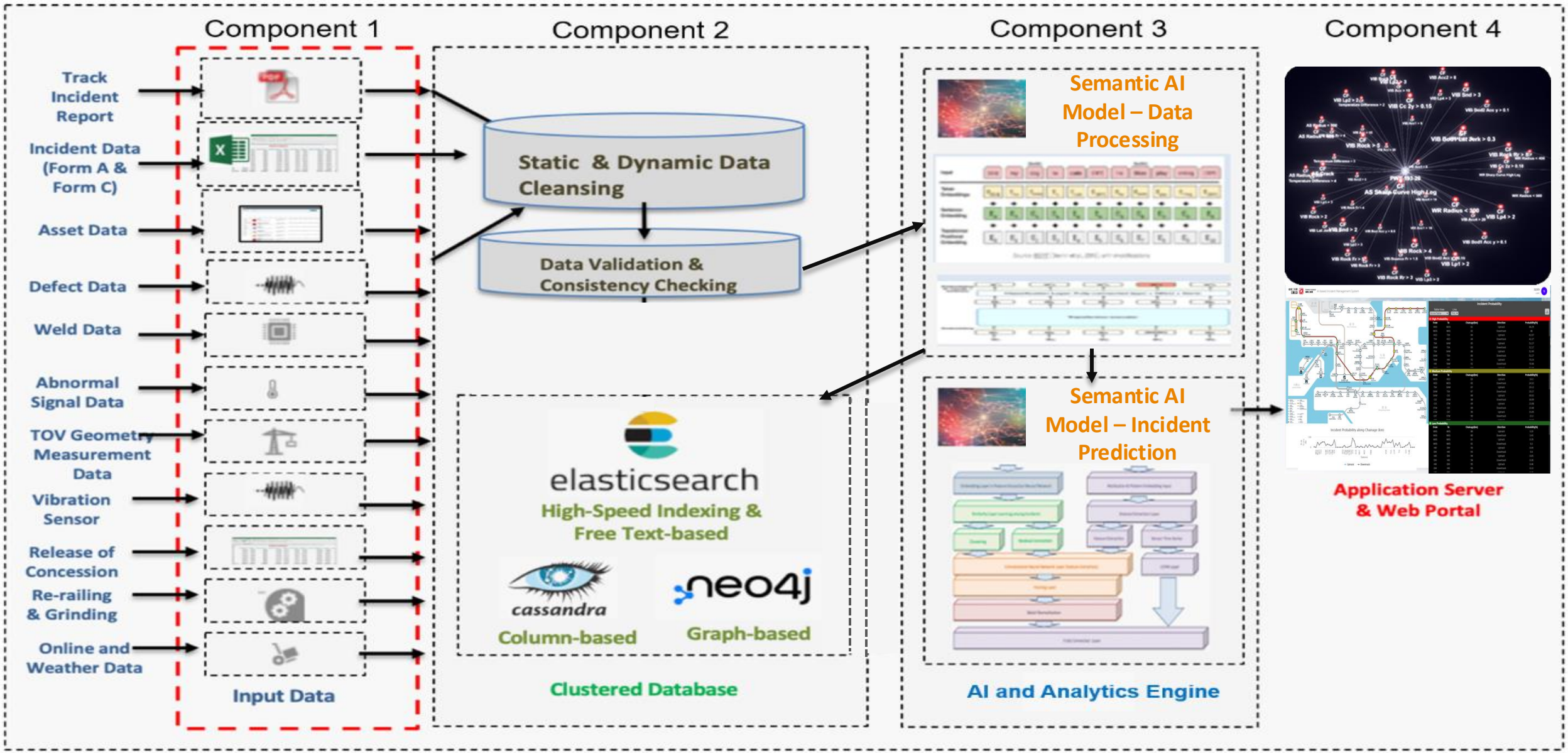
Novel Approach on AI Modelling

- Standardization – Railway Schema
- Adopt Semantic AI Technology
  - Transform raw data into structured data
  - Find out correlation, contribution factors and ranking of incident
- Develop AI Predictive Maintenance Model – based on incident prediction probability
- Pilot trial application on Permanent Way (Pway) System in Hong Kong railway network

Silver medal award wining project
Invention of Geneva 2022

Semantic AI Modelling

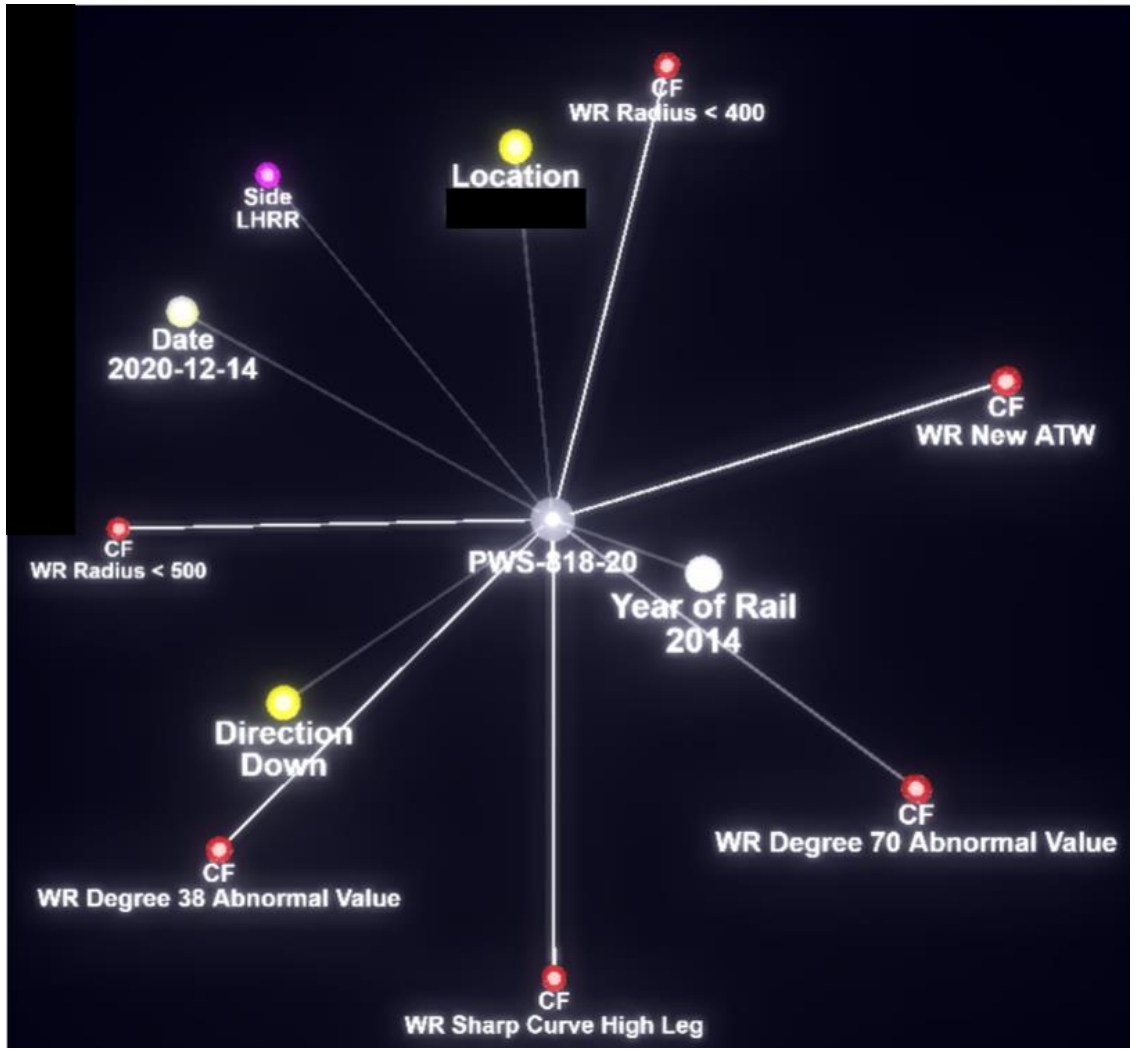
System Hierarchy



Semantic AI Modelling

Output

- **Knowledge Graph**
  - Visually Display
  - Relationship between Contributing Factors and incident
  - Strength of Edges



Semantic AI Modelling

Output

- Probability of incident along Pway chainage
- Ranked
  - High Risk
  - Medium Risk
  - Low Risk

Table View: Score Table

Incident Probability

High Probability

From	To	Chainage(km)	Direction	Probability(%)
Location A1	Location B1	1	Uptrack	64.54
Location A2	Location B2	2	Uptrack	81.00
Location A3	Location B3	3	Uptrack	72.99
Location A4	Location B4	4	Uptrack	85.45
Location A5	Location B5	5	Uptrack	81.46
Location A6	Location B6	6	Uptrack	50.56
Location A7	Location B7	7	Uptrack	73.71
Location A8	Location B8	8	Uptrack	86.67
Location A9	Location B9	9	Uptrack	95.03
Location A10	Location B10	10	Uptrack	84.90

Medium Probability

From	To	Chainage(km)	Direction	Probability(%)
Location A11	Location B11	11	Uptrack	19.35
Location A12	Location B12	12	Uptrack	42.13
Location A13	Location B13	13	Uptrack	39.49
Location A14	Location B14	14	Uptrack	36.14
Location A15	Location B15	15	Uptrack	39.50
Location A16	Location B16	16	Uptrack	36.01
Location A17	Location B17	17	Uptrack	24.37
Location A18	Location B18	18	Uptrack	21.19
Location A19	Location B19	19	Uptrack	30.24
Location A20	Location B20	20	Uptrack	25.82

Low Probability

From	To	Chainage(km)	Direction	Probability(%)
Location A21	Location B11	21	Uptrack	7.51
Location A22	Location B12	22	Uptrack	7.09
Location A23	Location B13	23	Uptrack	10.30
Location A14	Location B14	24	Uptrack	2.03
Location A15	Location B15	25	Uptrack	10.81
Location A16	Location B16	26	Uptrack	4.97
Location A17	Location B17	27	Uptrack	5.43
Location A18	Location B18	28	Uptrack	9.21
Location A19	Location B19	29	Uptrack	7.18
Location A20	Location B20	30	Uptrack	10.20

Part 3
Results

Results

Incident Probability Prediction

- ✓ Accuracy Test (Round 1)
  - ➢ Training data: Jan 2016 – Aug 2021 (68 Months, Line 1 data)
  - ➢ Testing data: Sep 2021 – Apr 2022 (8 Months, Line 1 data)
  - ➢ Accuracy is **67%** with Probability Threshold at 0.7

Probability Threshold	Correct Predictions	**Accuracy (True Positive & True Negative)**	False Positives	False Negatives
0.1	59	**51%**	35%	8%
0.2	68	**59%**	26%	9%
0.3	74	**64%**	19%	11%
0.4	77	**66%**	15%	13%
0.5	76	**66%**	14%	15%
0.6	78	**67%**	10%	16%
0.7	78	**67%**	8%	19%
0.8	73	**63%**	6%	25%
0.9	74	**64%**	2%	28%

Results

Incident Probability Prediction

✓ Accuracy Test (Round 2)

- Training data: Jan 2016 – Dec 2021 (68 + 4 Months, Line 1 data)
- Testing data: Jan 2022 – Jun 2022 (6 Months, Line 1 data)
- Accuracy is **72%** with Probability Threshold at 0.7

Probability Threshold	Correct Predictions	**Accuracy (True Positive & True Negative)**	False Positives	False Negatives
0.1	69	**59%**	36%	4%
0.2	73	**63%**	27%	10%
0.3	77	**66%**	20%	14%
0.4	83	**72%**	15%	14%
0.5	83	**72%**	9%	19%
0.6	83	**72%**	6%	22%
0.7	83	**72%**	4%	24%
0.8	85	**73%**	3%	24%
0.9	87	**75%**	0%	25%

Accuracy increased by 5%
False Positives improved by 4%

Results

Incident Probability Prediction

- ✓ Accuracy Test (Round 3)
  - ➢ Training data: Jan 2016 – Dec 2021 (68 + 4 Months, Line 1 data)
  - ➢ Testing data: Jan 2022 – Apr 2022 (4 Months, Line 2 data)
  - ➢ Accuracy is **56%** with Probability Threshold at 0.7

Probability Threshold	Correct Predictions	**Accuracy (True Positive & True Negative)**	False Positives	False Negatives
0.1	39	**63%**	32%	5%
0.2	39	**63%**	32%	5%
0.3	51	**82%**	13%	5%
0.4	51	**82%**	13%	5%
0.5	45	**73%**	8%	19%
0.6	35	**72%**	3%	40%
0.7	34	**56%**	0%	45%
0.8	33	**55%**	0%	47%
0.9	32	**52%**	0%	48%

Part 4
Conclusion

Conclusion

Semantic AI Model

- Schema standardization allows integration of wide range data into a clustered database more easily
- Semantic technology (Q&A module) enables recovery and visualization of structured knowledge/ information from data more efficiently
- Identified contributing factors and not easily noticeable indicators (Pway incidents in this case)
- Incident prediction accuracy improves with more data are used to train up the AI model
- Ranking of incident probability empower maintainer with predictive early warnings, historical case matching, and actionable intelligence (predictive maintenance)
- Application in other railway systems is feasible

Acknowledgement

Organizations participated in this pilot trial

- Mass Transit Railway Corporation Limited
  Infrastructure Maintenance

- City University of Hong Kong
  School of Data Science

Thank You !

INTERNATIONAL RAILWAY SAFETY COUNCIL 2022
IRSC Spain

機電工程署
EMSD

www.irsc2022.com

EUROPEAN UNION AGENCY FOR RAILWAYS

AESF AGENCIA ESTATAL DE SEGURIDAD FERROVIARIA

adif