**REFERENCES**

[1] S. Terabe, T. Kato, H. Yaginuma, N. Kang, and K. Tanaka, ‘‘Risk

assessment model for railway passengers on a crowded platform,’’ *Transp.*

*Res. Rec., J. Transp. Res. Board*, vol. 2673, no. 1, pp. 524–531, Jan. 2019,

doi: 10.1177/0361198118821925.

[2] *Annual Health and Safety Report 19/2020*, RSSB, London, U.K., 2020.

[3] D. M. Blei, ‘‘Probabilistic topic models,’’ *Commun. ACM*, vol. 55, no. 4,

pp. 77–84, Apr. 2012, doi: 10.1145/2133806.2133826.

[4] M. Gethers and D. Poshyvanyk, ‘‘Using relational topic models to

capture coupling among classes in object-oriented software systems,’’

in *Proc. IEEE Int. Conf. Softw. Maintenance*, Sep. 2010, pp. 1–10, doi:

10.1109/ICSM.2010.5609687.

[5] D. M. Blei, A. Y. Ng, and M. I. Jordan, ‘‘Latent Dirichlet allocation,’’

*J. Mach. Learn. Res.*, vol. 3, nos. 4–5, pp. 993–1022, Mar. 2003, doi:

10.1016/B978-0-12-411519-4.00006-9.

[6] H. Alawad, S. Kaewunruen, and M. An, ‘‘A deep learning approach

towards railway safety risk assessment,’’ *IEEE Access*, vol. 8,

pp. 102811–102832, 2020, doi: 10.1109/ACCESS.2020.2997946.

[7] H. Alawad, S. Kaewunruen, and M. An, ‘‘Learning from accidents:

Machine learning for safety at railway stations,’’ *IEEE Access*, vol. 8,

pp. 633–648, 2020, doi: 10.1109/ACCESS.2019.2962072.

[8] A. J.-P. Tixier, M. R. Hallowell, B. Rajagopalan, and D. Bowman,

‘‘Automated content analysis for construction safety: A natural language

processing system to extract precursors and outcomes from unstructured

injury reports,’’ *Autom. Construct.*, vol. 62, pp. 45–56, Feb. 2016, doi:

10.1016/j.autcon.2015.11.001.

[9] J. Sido and M. Konopik, ‘‘Deep learning for text data on mobile

devices,’’ in *Proc. Int. Conf. Appl. Electron.*, Sep. 2019, pp. 1–4, doi:

10.23919/AE.2019.8867025.

[10] A. Serna and S. Gasparovic, ‘‘Transport analysis approach based on big

data and text mining analysis from social media,’’ *Transp. Res. Proc.*,

vol. 33, pp. 291–298, Jan. 2018, doi: 10.1016/j.trpro.2018.10.105.

[11] P. Hughes, D. Shipp, M. Figueres-Esteban, and C. van Gulijk, ‘‘From

free-text to structured safety management: Introduction of a semiautomated

classification method of railway hazard reports to elements

on a bow-tie diagram,’’ *Saf. Sci.*, vol. 110, pp. 11–19, Dec. 2018, doi:

10.1016/j.ssci.2018.03.011.

[12] A. Chanen, ‘‘Deep learning for extracting word-level meaning from safety

report narratives,’’ in *Proc. Integr. Commun. Navigat. Surveill. (ICNS)*,

Apr. 2016, pp. 5D2-1–5D2-15, doi: 10.1109/ICNSURV.2016.7486358.

[13] A. Ferrari, G. Gori, B. Rosadini, I. Trotta, S. Bacherini, A. Fantechi,

and S. Gnesi, ‘‘Detecting requirements defects with NLP patterns: An

industrial experience in the railway domain,’’ *Empirical Softw. Eng.*,

vol. 23, no. 6, pp. 3684–3733, Dec. 2018, doi: 10.1007/s10664-018-

9596-7.

[14] G. Fantoni, E. Coli, F. Chiarello, R. Apreda, F. Dell’Orletta, and

G. Pratelli, ‘‘Text mining tool for translating terms of contract into

technical specifications: Development and application in the railway

sector,’’ *Comput. Ind.*, vol. 124, Jan. 2021, Art. no. 103357, doi:

10.1016/j.compind.2020.103357.

[15] G. Yu, W. Zheng, L. Wang, and Z. Zhang, ‘‘Identification of significant

factors contributing to multi-attribute railway accidents dataset (MARA-D)

using SOM data mining,’’ in *Proc. 21st Int. Conf. Intell. Transp. Syst.*

*(ITSC)*, Nov. 2018, pp. 170–175, doi: 10.1109/ITSC.2018.8569336.

[16] Y. Wang, W. Zheng, H. Dong, and P. Gao, ‘‘Factors correlation mining

on railway accidents using association rule learning algorithm,’’ in *Proc.*

*IEEE 23rd Int. Conf. Intell. Transp. Syst.*, vol. 22, Sep. 2020, pp. 1–6, doi:

10.1109/ITSC45102.2020.9294317.

[17] S. Sarkar, S. Vinay, R. Raj, J. Maiti, and P. Mitra, ‘‘Application of

optimized machine learning techniques for prediction of occupational

accidents,’’ *Comput. Oper. Res.*, vol. 106, pp. 210–224, Jun. 2019, doi:

10.1016/j.cor.2018.02.021.

[18] H. Jelodar,Y.Wang, C.Yuan, X. Feng, X. Jiang,Y. Li, and L. Zhao, ‘‘Latent

Dirichlet allocation (LDA) and topic modeling: Models, applications,

a survey,’’ *Multimedia Tools Appl.*, vol. 78, no. 11, pp. 15169–15211,

Jun. 2019, doi: 10.1007/s11042-018-6894-4.

[19] S. W. Thomas, ‘‘Mining software repositories using topic models,’’ in

*Proc. 33rd Int. Conf. Softw. Eng. (ICSE)*, May 2011, pp. 1138–1139, doi:

10.1145/1985793.1986020.

[20] H. U. Asuncion, A. U. Asuncion, and R. N. Taylor, ‘‘Software traceability

with topic modeling,’’ in *Proc. ACM/IEEE 32nd Int. Conf. Softw. Eng.*,

vol. 1, May 2010, pp. 95–104, doi: 10.1145/1806799.1806817.

[21] Z. Jiang, X. Zhou, X. Zhang, and S. Chen, ‘‘Using link topic model to

analyze traditional Chinese medicine clinical symptom-herb regularities,’’

in *Proc. IEEE 14th Int. Conf. e-Health Netw., Appl. Services (Healthcom)*,

Oct. 2012, pp. 15–18, doi: 10.1109/HealthCom.2012.6380057.

[22] M. J. Paul and M. Dredze, ‘‘You are what you tweet: Analyzing Twitter for

public health,’’ in *Proc. Int. AAAI Conf. Weblogs Social Media (ICWSM)*,

2011, pp. 265–272.

[23] W. Zhao, W. Zou, and J. J. Chen, ‘‘Topic modeling for cluster analysis

of large biological and medical datasets,’’ *BMC Bioinf.*, vol. 15, no. 11,

pp. 1–11, Oct. 2014, doi: 10.1186/1471-2105-15-S11-S11.

[24] H.-M. Lu, C.-P. Wei, and F.-Y. Hsiao, ‘‘Modeling healthcare data using

multiple-channel latent Dirichlet allocation,’’ *J. Biomed. Informat.*, vol. 60,

pp. 210–223, Apr. 2016, doi: 10.1016/j.jbi.2016.02.003.

[25] S. Bauer, A. Noulas, D. Ó. Séaghdha, S. Clark, and C. Mascolo, ‘‘Talking

places: Modelling and analysing linguistic content in foursquare,’’ in

*Proc. ASE/IEEE Int. Conf. Privacy, Secur., Risk Trust, ASE/IEEE Int.*

*Conf. Social Comput. (SocialCom/PASSAT)*, Sep. 2012, pp. 348–357, doi:

10.1109/SocialCom-PASSAT.2012.107.