



University of  
**Salford**  
MANCHESTER

# Utilizing big data for enhancing passenger safety in railway stations

Alawad, H, kaewunruen, S and An, M

<b>Title</b>	Utilizing big data for enhancing passenger safety in railway stations
<b>Authors</b>	Alawad, H, kaewunruen, S and An, M
<b>Publication title</b>	Proceedings of the 4th World Multidisciplinary Civil Engineering-Architecture-Urban Planning Symposium
<b>Publisher</b>	IOP Publishing Ltd
<b>Type</b>	Conference or Workshop Item
<b>USIR URL</b>	This version is available at: <a href="http://usir.salford.ac.uk/id/eprint/51914/">http://usir.salford.ac.uk/id/eprint/51914/</a>
<b>Published Date</b>	2019

USIR is a digital collection of the research output of the University of Salford. Where copyright permits, full text material held in the repository is made freely available online and can be read, downloaded and copied for non-commercial private study or research purposes. Please check the manuscript for any further copyright restrictions.

For more information, including our policy and submission procedure, please contact the Repository Team at: [library-research@salford.ac.uk](mailto:library-research@salford.ac.uk).

# Utilizing big data for enhancing passenger safety in railway stations

Alawad, Hamad Ali H; Kaewunruen, Sakdirat; An, Min

## *License:*

Creative Commons: Attribution (CC BY)

## *Document Version*

Peer reviewed version

## *Citation for published version (Harvard):*

Alawad, HAH, Kaewunruen, S & An, M 2019, Utilizing big data for enhancing passenger safety in railway stations. in The 4th World Multidisciplinary Civil Engineering-Architecture-Urban Planning Symposium. IOP Publishing Ltd., The 4th World Multidisciplinary Civil Engineering-Architecture-Urban Planning Symposium, Prague, Czech Republic, 17/06/19.

[Link to publication on Research at Birmingham portal](#)

## **Publisher Rights Statement:**

Checked for eligibility: 21/06/2019

## **General rights**

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

## **Take down policy**

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

# Utilizing Big Data for Enhancing Passenger Safety in Railway Stations

Hamad Alawad<sup>1</sup>, Sakdirat Kaewunruen<sup>2</sup>, Min An<sup>3</sup>

<sup>1</sup> Birmingham Centre for Railway Research and Education, The University of Birmingham, Birmingham B15 2TT, UK

<sup>2</sup> Birmingham Centre for Railway Research and Education, The University of Birmingham, Birmingham B15 2TT, UK

<sup>3</sup> School of Built Environment, University of Salford, Manchester, M5 4WT, UK

Hxa724@student.bham.ac.uk

**Abstract.** In light of the increasing demand and capacity in the railway industry, it is imperative to maintain safety in relation to the complexities of the substantial railway stations. Thus, it is important to take note of the time where investments in new technologies directed at the safety of the railway enable safety and protection in this area. Novel technological techniques such as big data analysis (BDA), data mining or machine learning (ML) have been developed and applied in many areas such as sales, banking and healthcare. The development of such methods has important benefits within the context of railway safety, however, these new methods need to be implemented and developed with consideration of whether these operational models can help to solve the various difficulties that currently exist in the risk analysis of railway stations. Moreover, as the adoption of the Internet of thing (IoT) grows, it is expected that analytical needs for handling data will also increase. It has been shown that the progression towards automation and applying such innovative new technologies such as BDA may be a powerful tool for integration in the future of transportation in general and the railway industry in particular, whereby analytical predictions can aid in the development of safer railway stations which have greater potential for ensuring the safety of passengers. In this paper a Bow Tie (BT) framework model has been created to combine BDA into the risk assessment process. The BDA can be beneficial to the risk assessment, support the decision makers in real time, and reduce human errors. This method can be fully integrated into passenger data and the business model for the railway station. Employing the existing safety records utilizing BDA is expected to mitigate risks, predict hazards, raise safety and security efficiency and reduce the cost.

## 1. Introduction

It is important to take note of the time where investments in new technologies directed at safety of the railway enable safety and protection in this area. The development of such methods has important benefits within the context of railway safety, however, these new methods need to be implemented and developed with consideration of whether these operational models can help to solve the various difficulties that currently exist in the risk analysis of railway stations. One of the main drawbacks of traditional quantitative risk assessment is static management or lack of ability to update the risks over time for developing new circumstances. Over the past decades, risk management methods have

developed models based on static data to discover hazardous situations and raise the safety level in the stations and to support safety, risk management and decision-making (e.g. Bow-Tie model and Reason's Swiss Cheese Model). However, this traditional risk analysis management will be changed by the spread of new technologies, the systems in the stations and new technology such as sensors and customer smart phones which are producing enormous amounts of data which could be linked to improving the safety, security, predicting the risks in real-time and support decision making. Risk methods need to be developed in relation to model uncertainties and to build confidence limits concerning the predictions. It has been shown that the decision maker would have had some discrepancies since many factors can be an influence, such as the subjective evaluation of the experts and the risk assessment process time [1].

It is expected that BDA will improve ways of acquisition, storage and utilization for the smart railway station generation with more efficient and reliable systems; in other words, we have now entered the big data era [2]. Moreover, research has demonstrated that the blending of computing power and large amounts of raw data from a variety of sources could perform an essential role in promoting the safety of the rail industry [3,4]. The sources of BDA in the station can be captured from a variety of places, even historical databases or newly included safety records, reports, emails, text, call audio, images, videos as well as the data from sensors, RFID scans, surveillance, GPS and social media. The properties of Big Data are volume, variety, velocity, variability and complexity and the success of big data analysis depends critically on some technologies, for instance, ML, data mining, time series, cloud computing and visualization [5,6,7].

While the BDA has become an attractive option, it has been suggested that to move forwards with this technique, more clarification of what BDA is and what big data analysis requires are needed. The aims of BDA analysis are largely driven by big data stakeholders' or clients' needs. It has been utilised in many industries such as manufacturing, healthcare and in many businesses such as media, banking, insurance, communication, energy industry and transportation which is the target in this article [6,8,9]. The potential value of BDA is noted and boosted by growing studies and researchers in many fields, particularly within the transportation industry and currently, the notion of a BD revolution is popular in the railway industry. Thus, there is a need for the capability to process growing data where the concept of the smart railway stations offers a successful environment using the big data strategy. It is expected that the transportation industry will benefit from BDA to more effectively manage and contribute to a safer, more efficient operation in the stations which will be reflected in the passenger's travel experience [10,11]. Some station designs are rather complex and are much more than simply a place where trains enter and leave. Indeed, modern railway stations are multi-functional environments. They often feature shops, hotels, and other services, and have become a focal point around which cities have been built. Rail stations have become vital nodes for business and transportation systems.

As a result, railway technology is also becoming more complex since using intelligent assets means unprecedented levels of digital transmutation which then leads to changing the way the industry works, expressly regarding risk management and safety. This change includes city digitalization, passenger needs and to ensure the safety and security of the railway station. Due to advanced powerful computing improvement and the explosion in data availability and development in techniques and algorithms, there should be opportunities to use a BDA and other approaches such as deep learning and ML to proactively identify high-risk scenarios in the railway stations. It has been suggested that utilising the machine intelligence will reduce cost and increase resilience throughout the system's lifecycle including risk analysis and management [10,12,13].

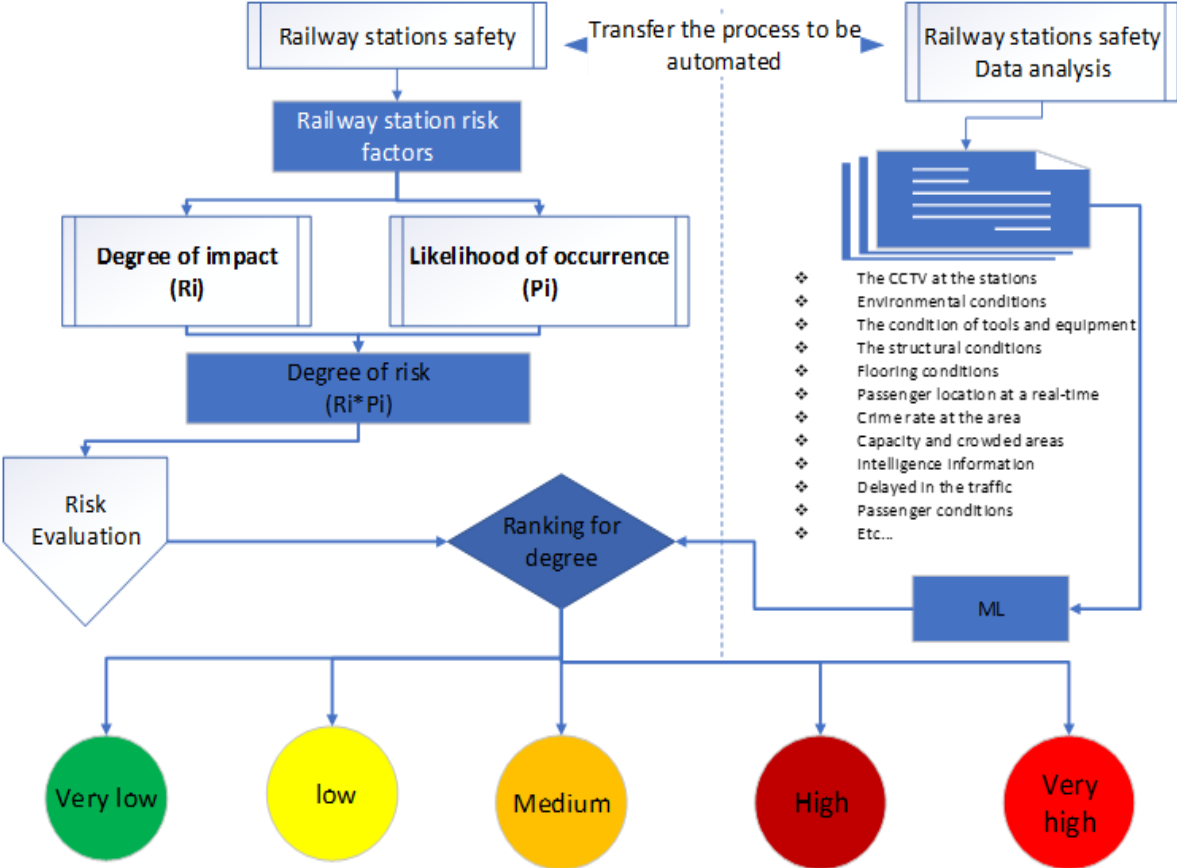
This contribution is made to the Safety Science in general and to risk management specifically in the railway industry. This paper is written from the point of applying new technologies for risk management with respect to operations in the railway stations. The final vision is of a smart railway

station entirely associated with other organizations' data such as smart city, police, and intelligence data etc., where information is shared, and analyses are performed in real time and utilized for passenger safety management, which is a relatively new idea. The advanced evaluation of BDA approaches within data science is known as BDA which has utility in providing this methodology based on inferential statistics while the conventional ways use descriptive statistics. This methodology provides a flexible tool, able to analyse raw data and in an advanced process to learn from "occurrence" detecting patterns, correlations, dependencies, and to predict consequences and responses. Safety in the railway stations can be significantly enhanced from BDA and such methodologies able to gather, store, analyse and visualize the vast amounts of data. In this paper, a framework has been conducted, using the BT model to fit with BDA, and a flow chart to explain the view of ML for risk in the station and transferring the process to be automated.

**2. Machine learning and BDA**

From the literature for dynamic risk management view, it can be seen that artificial neural network (ANN) have been applied, as well as Bayesian for dynamic risk assessment, where Bayes theorem works with the historical data or records and supports the vector machine [14-17].

In this part of work, we present the ML as the technology that will aid in extracting safety information from pure operational past data in the stations, to identify the predictive approach for supervising safety decisions (see Figure 1). This enables the quantification of effects on the passengers and the related systems.



**Figure 1.** The ML framework for risk management in the railway stations.

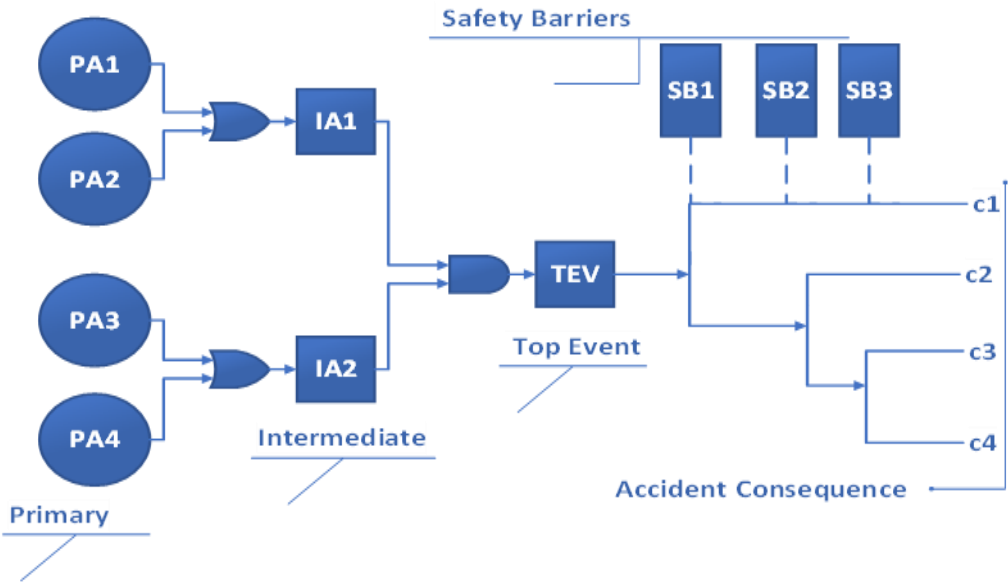
There is a necessity for further smart models like forecasting and nowcasting methods for the railway assets to improve capacity, punctuality and other performance indicators. Thus, BDA can be used along with the ML algorithms to analyse big data for further improvement in the operations and services in the stations.

BDA is a modern theory that can advance the overall accuracy of ML in the risk management context by improving the predictive potential of risk models. Moreover, ML techniques are useful for data analysis and pattern discovery which can play an important role in managing risks. The implementation of such technology related to the railway station is still not so widespread, and this is owing to many factors such as unstructured data, cost and the high skilled manpower required. It is essential that risk is addressed from an operational perspective, from a passenger’s view, and particularly in terms of accident prevention and high levels of security and safety. The BDA can aid operational and regulators to predict real-time symptoms of risks and acting in time to prevent or mitigate them [18-20].

**3. Bow Tie and risk assessment with BDA**

In this section, a BT model is selected which is very popular. A BT framework has proven to be valuable in analysing past accidents and signifying improvements to avoid further reoccurrence of undesired events which have also demonstrated unwanted event prediction. Moreover, it has been shown that for BT, techniques can be used in the process of risk analysis, both qualitative and quantitative. The BT consequence side can make an interface with the decision models, and ultimately decisions taken will be reflected back toward the causes [21-23].

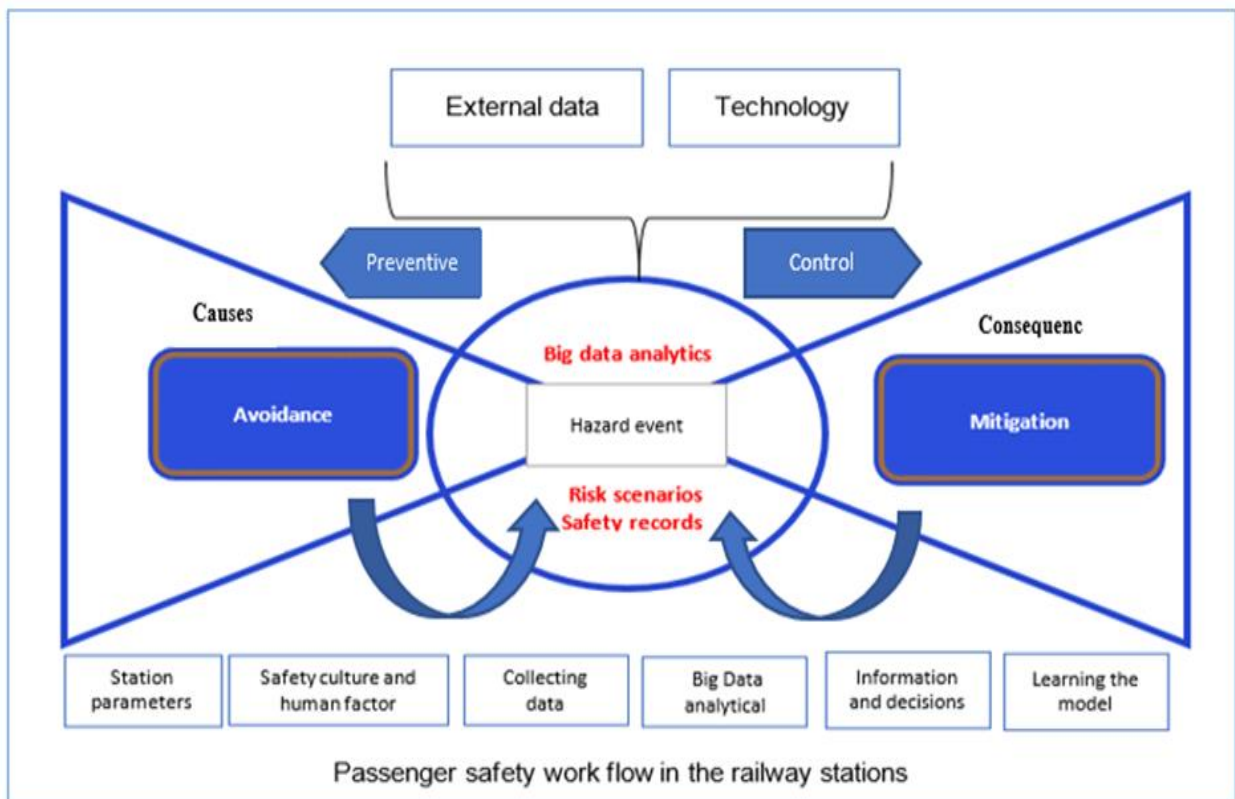
BT framework has been used for occupational risk management, and it is selected in this work as it is described as a tool that is both proactive and reactive and systematically works through the risk management. BT can be employed to explain how the related safety management system detail provisions are met with respect to the control and management of danger and risk circumstances [24-28]. For more effective analysis and prediction of an accident, it is essential to understand the accident scenario altogether including causes and consequences. Figure 2 illustrates an example of BT. However, a typical BT has a static structure limiting its ability to capture risks in real-time and monitoring, and this consequently delays the probability of updating, which is a fundamental determinant in dynamic risk management.



**Figure 2.** Hierarchy of an example BT.

The BT evaluation methodology has been created to integrate BDA into the risk management process. The BT analysis has been enhanced through associating related Industry Data Taxonomy sources to each of the factors of the BT. These data sources can include many aspects, for instance, monitoring, social media, and safety management and so on. BDA analysis can be beneficial to the risk assessment, support the decision makers in real time, and reduce human errors.

This method can be fully integrated into passenger data and the business model for the railway station. The present BT approach is updated to be a dynamic structure, and this includes the BDA which will reflect on the occurrence and probability of accident consequences and an acceptable safety level. The process will occur in real-time by reliable models and internal or external information in the station and will be periodically updated using indicators of the risks (see Figure 3).



**Figure 3.** The BT and BDA analysis.

The main aim is to use a proposed BT framework and a developed risk management model to evaluate and prioritise risk factors in the railway station based on real-time information from many sources, which consequently makes the process dynamic and has the ability to make predictions. The proposed framework is for BDA, and the workflow describes the steps of the process referring to learning or training the model and the risk scenarios [21-29]. BT framework integrates the fault and event trees which is valuable for analysing hazardous events, as their causes and outcomes are continuously linked together. BT has a suitable level of simplification of the causal factors in order to be able to abstract large quantities of data into a relatively small number of common scenarios, which can cover the majority of accidents in the railway stations. Furthermore, these scenarios can illustrate

the link between an accident and all its potential causes. Additionally, this suggested BT provides an integrated process, feeding the BDA from both sides of the accident (causes and consequences) and varied information available (internal and external) and from history and real-time safety information that support the decision making and risk management [1, 22].

#### **4. Conclusions**

Overall, the advanced risk management model can achieve the objectives of the BDA framework in the railway stations. This model can assist risk managers and decision makers and that is expected to reflect on the service, safety, security and efficiency in the station. It can also aid in deciding whether to take preventive actions for causes or corrective actions during the risk mitigation phase and in real time with more robustness, fewer uncertainties and human errors. This will lead to proceeding towards a dynamic and predictive risk management process utilising new technologies artificial intelligence (AI) in real time. Employing the existing safety records utilizing BDA is expected to mitigate risks, predict hazards, raise safety and security efficiency and reduce the cost. Here, the safety records are vital information and need to be a part of an automated process. This type of technology is interesting because it makes use of historical data of safety records such as the history of investigating accidents to predict future events. Railway stations can maintain safety levels proactively and obtain unexpected notices by modelling the experience that has to be avoided and classifying and controlling precursors to those events. Moreover, the new data analytics method will encourage the railway industry to recognise the significance and advantage of the safety data. Also, the BDA will improve the perception of the risk features which are expected to reflect on the service and passenger satisfaction and then reduce the overall cost. The data is valuable since it holds a definitive record of all the activity and behaviour of the passengers, transactions, applications, servers, mobiles or other devices in the railway stations. BDA technology has been used to monitor and improve safety and can be useful in promoting international safety and security in the railway industry. Finally, it has been shown that the progression towards automation and applying innovative new technologies such as BDA, which is a powerful tool for integration in the future of transportation in general and the railway industry in particular, means analytical predictions can aid in the development of safer railway stations which have greater potential for ensuring the safety of passengers. Then, applying such technology is expected to provide direction in bridging the existing gaps in research and development in risk management.

#### **Acknowledgment(s)**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 691135. The second authors wishes to thank to the European Commission for the financial sponsorship of the H2020-RISE Project No. 691135 "RISEN: Rail Infrastructure Systems Engineering Network", which enables a global research network that tackles the grand challenge of railway infrastructure resilience and advanced sensing in extreme environments ([www.risen2rail.eu](http://www.risen2rail.eu)) [30]. The second author is sincerely grateful to the Australian Academy of Science and the Japan Society for the Promotion of Sciences for his Invitation Research Fellowship (Long-term), Grant No. JSPS-L15701 at the Railway Technical Research Institute and the University of Tokyo, Japan.

#### **References**

- [1] F. Khan, S. Rathnayaka, and S. Ahmed, "Methods and models in process safety and risk management: Past, present and future," *Process Safety and Environmental Protection*, vol. 98, pp. 116–147, Nov. 2015.
- [2] H. J. Watson, "Tutorial: Big Data Analytics: Concepts, Technologies, and Applications," *Communications of the Association for Information Systems*, vol. 34, 2014.
- [3] Figueres-Esteban, Miguel, Peter Hughes, and Coen Van Gulijk. "The role of data visualization



in railway big data risk analysis." In *Proceedings of the 25th European Safety and Reliability Conference, ESREL 2015*, pp. 2877-2882. CRC Press/Balkema, 2015.

- [4] Figueres Esteban, Miguel, Peter Hughes, and Coen van Gulijk. "Big Data for Risk Analysis: the future of safe railways." In XII Congreso de ingeniería del transporte. 7, 8 y 9 de Junio, Valencia (España), pp. 347-353. Editorial Universitat Politècnica de València, 2016.
- [5] S. Bryson, D. Kenwright, M. Cox, D. Ellsworth, and R. Haines, "Visually exploring gigabyte data sets in real time," *Communications of the ACM*, vol. 42, no. 8, pp. 82–90, Aug. 1999.
- [6] J. Bughin, J. Livingston, and S. Marwaha, "Exploiter le potentiel du « big data »,," *L'Expansion Management Review*, vol. 146, no. 3, p. 52, 2012.
- [7] K. Kaloyanova, T. Hristov, I. Naydenova, and Z. Kovacheva, "Information Management Technologies for Big Data," *Proceedings of the Second International Conference on Advanced Wireless Information, Data, and Communication Technologies - AWICT 2017*, 2017.
- [8] "GRANULARITIES AND INCONSISTENCIES IN BIG DATA ANALYSIS | International Journal of Software Engineering and Knowledge Engineering," *International Journal of Software Engineering and Knowledge Engineering*, 2013. [Online]. Available: <http://www.worldscientific.com/doi/abs/10.1142/S0218194013500241> .
- [9] S. Miele and R. Shockley, "IBM Global Business Services Business Analytics and Optimization Executive Report Analytics: The real-world use of big data How innovative enterprises in the midmarket extract value from uncertain data," Apr. 2013.
- [10] H. Alawad and S. Kaewunruen, "Wireless Sensor Networks: Toward Smarter Railway Stations," *Infrastructures*, vol. 3, no. 3, p. 24, Jul. 2018.
- [11] A. I. Torre-Bastida, J. Del Ser, I. Laña, M. Iardia, M. N. Bilbao, and S. Campos-Cordobés, "Big Data for transportation and mobility: recent advances, trends and challenges," *IET Intelligent Transport Systems*, vol. 12, no. 8, pp. 742–755, Oct. 2018.
- [12] L. Bertolini, "Nodes and places: complexities of railway station redevelopment," *European Planning Studies*, vol. 4, no. 3, pp. 331–345, Jun. 1996.
- [13] S. Kohli, A. V. S. Kumar, J. M. Easton, and C. Roberts, Eds., "Innovative Applications of Big Data in the Railway Industry," *Advances in Civil and Industrial Engineering*, 2018.
- [14] M. Marseguerra, M. Ricotti, and E. Zio, "Approaching system evolution in dynamic PSA by neural networks," *Reliability Engineering & System Safety*, vol. 49, no. 1, pp. 91–99, Jan. 1995.
- [15] M. Kalantarnia, F. Khan, and K. Hawboldt, "Dynamic risk assessment using failure assessment and Bayesian theory," *Journal of Loss Prevention in the Process Industries*, vol. 22, no. 5, pp. 600–606, Sep. 2009.
- [16] G. Dong, L. Xue, Y. Yang, and J. Yang, "Evaluation of hazard range for the natural gas jet released from a high-pressure pipeline: A computational parametric study," *Journal of Loss Prevention in the Process Industries*, vol. 23, no. 4, pp. 522–530, Jul. 2010.
- [17] L. Ni, J. Jiang, and Y. Pan, "Leak location of pipelines based on transient model and PSO-SVM," *Journal of Loss Prevention in the Process Industries*, vol. 26, no. 6, pp. 1085–1093, Nov. 2013.
- [18] F. Corea, "Big Data & Risk Management in Financial Markets (Part I)," *Medium*, 07-Nov-2016. [Online]. Available: [https://medium.com/@Francesco\\_AI/big-data-and-risk-management-in-financial-markets-part-i-eed4e245f3db](https://medium.com/@Francesco_AI/big-data-and-risk-management-in-financial-markets-part-i-eed4e245f3db). [Accessed: 23-Apr-2019].
- [19] I. Bose and R. K. Mahapatra, "Business data mining — a machine learning perspective," *Information & Management*, vol. 39, no. 3, pp. 211–225, Dec. 2001.
- [20] A. Thaduri, D. Galar, and U. Kumar, "Railway Assets: A Potential Domain for Big Data Analytics," *Procedia Computer Science*, vol. 53, pp. 457–467, 2015.
- [21] K. Mokhtari, J. Ren, C. Roberts, and J. Wang, "Application of a generic bow-tie based risk analysis framework on risk management of sea ports and offshore terminals," *Journal of Hazardous Materials*, vol. 192, no. 2, pp. 465–475, Aug. 2011.

- [22] J. Groeneweg, *Controlling the controllable: the management of safety*. Leiden: Dswopress, Leiden University, 1998.
- [23] L. J. Bellamy et al., "Storybuilder—A tool for the analysis of accident reports," *Reliability Engineering & System Safety*, vol. 92, no. 6, pp. 735–744, Jun. 2007.
- [24] J. E. Cockshott, "Probability Bow-Ties," *Process Safety and Environmental Protection*, vol. 83, no. 4, pp. 307–316, Jul. 2005.
- [25] K. Mokhtari, J. Ren, C. Roberts, and J. Wang, "Application of a generic bow-tie based risk analysis framework on risk management of sea ports and offshore terminals," *Journal of Hazardous Materials*, vol. 192, no. 2, pp. 465–475, Aug. 2011.
- [26] "Dnv [2001]- Frequency Analysis of Accidental Oil Releases From Fpso Operations in the Gulf of Mexico - [PDF Document]," *vdocuments.site*, 2014. [Online]. Available: <https://vdocuments.site/dnv-2001-frequency-analysis-of-accidental-oil-releases-from-fpso-operations.html>. [Accessed: 23-Apr-2019].
- [27] C. Jacinto and C. Silva, "A semi-quantitative assessment of occupational risks using bow-tie representation," *Safety Science*, vol. 48, no. 8, pp. 973–979, Oct. 2010.
- [28] Trbojevic, Vladimir M and B. J. Carr, "Risk based methodology for safety improvements in ports," *Journal of Hazardous Materials*, vol. 71, no. 1–3, pp. 467–480, 2015.
- [29] N. Khakzad, F. Khan, and P. Amyotte, "Dynamic risk analysis using bow-tie approach," *Reliability Engineering & System Safety*, vol. 104, pp. 36–44, Aug. 2012.
- [30] S. Kaewunruen, J. M. Sussman, and A. Matsumoto, "Grand Challenges in Transportation and Transit Systems," *Frontiers in Built Environment*, vol. 2, 2016.

