Big Data and the Virtuous Circle of Railway Digitization

Howard J. Parkinson*, Gary J Bamford **

**Digital Rail Limited, Lancaster England, ** NTTX Advisory Limited, England*

A B S T R A C T

Presently there are many disruptive technologies and philosophies that will change the way we learn from accidents and carry out safety management on the railway. These ideas and technologies include Big Data, Safety II, the Digital Railway, and Simulation. These have all come to the fore in the last 5 years or so to produce a potential virtuous circle that will provide significant synergy between them. This offers all railway stakeholders emergent opportunities including greater safety, better system reliability and efficiency.

This paper is concerned with railway safety and how it has improved over many years, yet requires new thinking to get further gains. A Big Data Analytics approach is described using the ELBowTie tool and the type of data available. Analytics are discussed and how a digital railway will provide a backbone for this based upon an Internet of Things architecture. Simulation and the use of digital information is discussed and how this can close the virtuous circle.

Keywords: Simulation, Railway Safety, Accidents Learning, Big Data, Digital Rail

1. Introduction

Railways safety has improved over the years often as a result of learning hard lessons from railway catastrophes. The modern approach is to try to envisage the accident before it happens and put in place mitigations. We are entering an age now where we have complex socio-technological systems that rely upon computer control and human-machine interfaces. This is particularly true in the railway industry which is moving towards more sophisticated automatic train protection (ATP). This could mean that the human workers who are still involved will find that a full understanding of the systems and operations is almost an intractable problem (Hollnagel 2015).

The situation is not all bad news though as there are other technologies and philosophies emerging that could, when taken together actually help take the railway to a new low level in accident and incident occurrence. These opportunities are described in this paper as a digital virtuous circle, which is illustrated in Figure 1.

The key elements are: a new way of considering safety termed Safety II (Hollnagel 2015), the advent of Big Data (BD) and the Internet of Things (IoT) (Parkinson and Bamford 2016), the Digital Railway initiative (Network Rail 2015), and the potential innovation of simulation, virtual and augmented reality and how they might fit together.

The next section contains a brief discussion of the philosophy behind railway safety and why we may need a new approach. Section 3 contains a discussion of Safety II, and how the existing techniques are no longer adequate. Section 4 provides an overview of the disruptive technologies that are around at present

adfa, p. 1, 2011.

affecting our approach to safety.

Section 5 introduces an approach to assessing risk, building in the potential of Big Data. Section 6 describes how the digital rail initiative coupled with Systems Engineering (SE) and Computer Aided Design (CAD) data would provide a foundation for the Big Data approach. Section 7 describes how simulation using all the aforementioned techniques might close the digital virtuous circle and Section 8 reviews how this all can support the intent of Safety II.

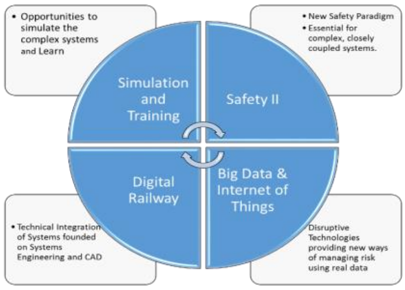

Figure 1. The Digital Virtuous Circle

2. Railway Safety

The unique characteristics of railway safety are that there is a vehicle with a large mass travelling at high velocity with low friction, steel wheel to steel rail contact, making for very long braking distances that are often longer than the driver's field of vision. Also, the train travels on guiding track which allows no opportunity to avoid collision by steering action. High collision speeds and high kinetic energy makes for high severity consequences. To make risks acceptable, therefore, it is necessary to reduce the probability of train accidents to a vanishingly small number.

Railway tragedies over many years have led to numerous safety improvements which include (Erdos 2004):

- Braking Systems: Vacuum brakes, compressor based braking systems, the Westinghouse Brake system,
- Communication Systems: Electric telegraph, telephone and radio systems,
- Vehicle Design: Crashworthiness avoiding telescoping,
- Signalling: Absolute Block Sections, Multiple-Aspect Signals, Track Circuits, Interlocking, Automated level crossings, ATP,
- Platform Interface: Screen doors, Door to traction interlocking

Figure 2 depicts the downward trend in fatal railway train accidents. A plateau seems to have been reached with technical failures and individual worker errors being reduced to very low levels. For example, no track worker fatalities occurred last year in the UK (RSSB 2015). The largest number of deaths result from suicides and trespassers, which swamp the all other causes.

There is, of course, still the danger of large unexpected railway accidents such as the recent ones in Bavaria (BBC 2016) and Santiago de Compostela in Spain (RSSB 2015) together resulting in over one hundred deaths. These are attributable to organisational type accident factors and are always complex in nature and usually the result of human error (Hollnagel 2015). The railway appears to be catching people out with more and more complex systems and rules.

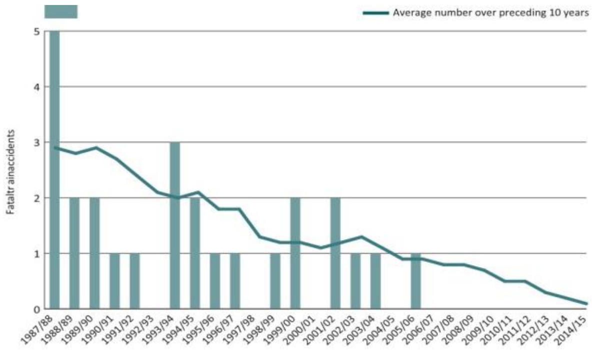

Figure 2. Train accidents with passenger or workforce fatalities (RSSB 2015)

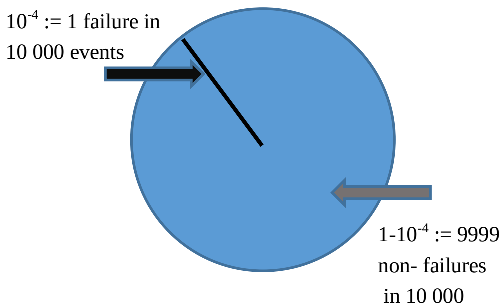

Figure 3. The imbalance between things that go right and things that go wrong (from Hollnagel 2015).

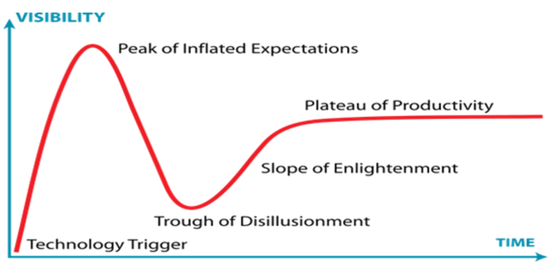

Figure 4. The Uptake in New Technology (Gartner 2014)

So, in the railway as in other industries, there are tightly coupled technologies with non-linear functioning, producing 'outlier' events that usually result in significant impacts which human supervisors find difficult to control in the normal and degraded operation. The next section contains a discussion of these implications in a little more detail and what might be needed to control them.

3. Safety II

Hollnagel (2015) has suggested that instead of focusing upon failure we focus upon how the system actually works in the real world and not as it is imagined to work by a development team. Having an understanding of how the system interacts with humans, it can then be established how best to safely manage them. Hollnagel (2015) has described this as Safety II which would supersede the traditional Safety I approach, which would still be suitable for non-complex systems.

Figure 3 taken from Hollnagel (2015) illustrates the imbalance that occurs when the focus is on what goes wrong and not on how things go right.

Safety I has its focus on analysing failure and has evolved from the analysis of more simple electro/mechanical system though 3 ages as describe by (Hale and Hovden 1998), i.e. the age of technology, the age of human factors and the age of safety management. Many of the techniques still used are founded in simpler times such as, failure mode and effect analysis (FMEA) and fault tree analysis (FTA) amongst others. A thorough understanding by the practitioner of past incidents is essential in all these processes. Without this knowledge it will be impossible to understand errors, failures, and how hazards propagate into accidents.

As technical failures have been reduced, and as the complexity and difficulty in understanding complex systems has increased, more and more accidents are being attributed to human error. It is becoming increasingly difficult to train workers to respond to the various hazardous states that these complex systems can find themselves in. So, new approaches are needed to understand, analyse and optimise day-to-day safety management. Modern systems are very reliable and focussing upon failure is the wrong way to analyse these systems in order to achieve the next step change in safety.

As stated earlier new techniques, data and initiatives are available that will enable a new focus upon how systems actually work. In addition, technologies to simulate environments exist to help increase the ability of the human worker to deal with system perturbation and increase the resilience of the safety system. The next section contains an overview of some of these new technologies.

4. Disruptive Technologies

It is a widely held belief that we are entering a revolution in technology and computer intelligence, driven primarily by increases in computing power and the development of intelligent algorithms. This is being described as the 3rd Industrial Revolution (Ford 2015). There are several so-called disruptive technologies present which include a Big Data (BD), the Internet of Things (IoT), Virtual Reality (VR), Augmented Reality (AR), and Machine Learning (ML) to name a few.

This interaction of disruptive technologies can be viewed as a key part of the virtuous circle. The pace at which ML is happening is now hard to deny (Parkinson, Bamford and Kandola 2016). Many of these technologies have promised great things in the past and not delivered. For example, neural networks were very much the flavour in the 1990s, however, due to the lack of computational power and the lack of understanding of the how to combine them with other approaches, they did not deliver the required accuracy for safety related predictions (Iwnicki and Parkinson 1999). Gartner (2014) has identified a curve that describes this trend in the uptake of new technology as shown in Figure 4.

Many of these disruptive technologies are now beginning to climb the slope of enlightenment and are approaching the plateau of productivity.

5. Big Data Approach

We are now in the Big Data Analytics (BDA) age where mountains of data should enable us to understand how complex socio technical systems actually work. This is a central idea of Safety II.

In addition, visualisation will be the key to providing the 'users' with a view of system dynamics. Data/information overload is a problem that will need to be overcome. Possibly by only presenting information critical to the business or the individual. This new area of critical visualisation is being called the Internet of Me - what matters for you view of the world (IET Turing lecture 2016).

Spending in railway systems in the next twenty years is set to explode, and it is critical that best practices in safety, data analysis and systems engineering are employ. (Parkinson and Bamford) have defined a BDA approach to learning from accidents which will enable the identification of heightened risk. The approach uses an accident causation model called an ELBowTie, illustrated in Figure 5. This model is based upon the bowtie methodology that is already well utilised in the safety engineering world (Yaneira 2013).

(Parkinson and Bamford) first establish a railway Enterprise Data Taxonomy (EDT) which lists the available sources of data that can be linked to railway operational risks. These include, for example: condition based monitoring information from sensors, either analogue or digital, that would provide digital information, including vibration (accelerometers), machine vision, heat, displacement, strain, humidity, particle ingress, etc. which would be classified as 'Real Time, remote monitoring', which is already an accepted means of classifying this type of data. Other data types are less well defined, for example, data from industry reports, staff morale, organisational culture, but can be equally as important in safety evaluations.

The EDT is aimed at providing a mechanism for systematically classifying data related to safety incidents to facilitate assessment of data sources in a traceable and consistent manner, giving the new approach its name: "Enterprise data taxonomy Linked Bow-Tie" (ELBowTie),

The research by (Parkinson and Bamford) looked a series of railway accidents and established their potential hazardous causes and conditions. These causes and conditions were then linked to the EDT. The EDT is then linked to the bowtie elements to predict heightened railway risk. It will be necessary to develop a bowtie for every railway hazard to develop the full risk picture.

At the heart of the tool are a series of Machine Learning algorithms that will be trained to analyse a stream of data and recognise when the system is outside of its normal bounds. These outputs are then amalgamated to take into account the complex interaction of the typical accident scenario (Parkinson, Bamford and Kandola 2016). The data is linked to the barriers on either side of the bowtie to enable real time interventions to take place either to prevent hazards occurring or to prevent hazards propagating into accidents. When a red flag situation transpires, the particular initiators are known and thus enabling automated warnings to be activated. Figure 6 depicts this relationship. The paper by Parkinson, Bamford and Kandola (2016) also has an extensive discussion of the various machine learning approaches that might be taken and how these could be amalgamated.

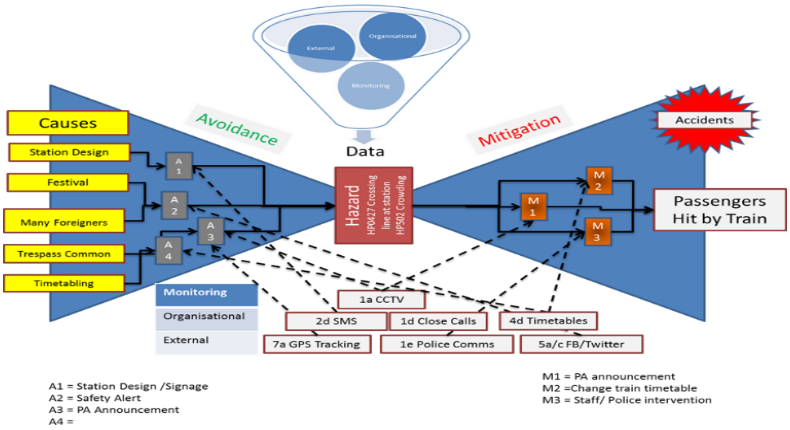

Figure 5. ELBowTie Big Sat Railway Risk Assessment Tool (Parkinson and Bamford 2016)

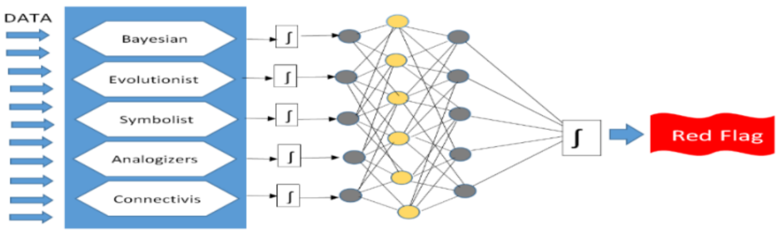

Figure 6. Proposed Structure for Combining Data (Parkinson, Bamford and Kandola 2016)

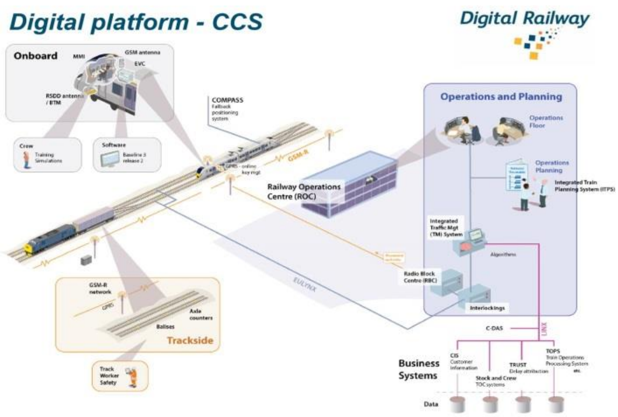

Figure 7. Digital Rail Architecture (Network Rail 2015)

6. The Digital Railway

The Digital Railway is an initiate in the UK mainly driven by Network Rail (Network Rail Ltd 2015) and is depicted in Figure 7. Network Rail are the organisation responsible for maintaining the mainline railway infrastructure in Great Britain. Several other organisations called TOCs (Train Operating Companies) are responsible for running trains and there are a myriad of private maintenance contractors and engineering support companies.

The core element of the Digital Railway will be the ERTMS (European Rail Traffic Management System) which, simplistically, comprises an on-board computer that provides automatic train protection by calculating a safe train movement authority based upon track conditions and the status of the railway. The interlocking sends signals to the train, relating whether a route is set and whether other trains are present. Around this core system all the communication, passenger information, status, etc. will be provided.

The Digital Railway will provide a platform for an IoT approach with enhanced communication systems and inbuilt sensors in assets.

7. Simulation and Training

Finally, we have moved into areas of simulation and training. With the aid of the digital models created during system development, and using AR and VR, we can create scenarios for workers to investigate how systems really function. This allows managers and designers to communicate with the various stakeholders about how the system works in all its complexity and not simply how they "think" it will work. This will enable effective training to be designed.

To build upon the ability to simulate systems, the recent training development of gamification is useful. Research suggest this is the optimum approach (Kapp 2012) to training were trainees are stimulated to learn through competition, collaboration and reward. E-learning has struggled to deliver on its promise due to lack of good training design and creativity. However, the latest simulation technology and gamification opportunities for blended delivery will improve the situation. Blended delivery is the combination of face to face delivery with e-learning and online training interventions.

Training and simulation enables the understanding of safety concepts and the working principles to be properly laid out. The training coarse attendees discover the skills and knowledge by playing the game and come away more motivate and better able to apply the principles.

8. Closing the Virtuous Circle

A virtuous circle has been described that could take railways to a lower level of safety risk. Disruptive technologies could be threats to the way the railway is operated and its continual viability but if the chances for efficiencies and improvements in safety are taken then the railway's prospects are improved. Machine learning and supporting technologies (Ford) are coming together now that will make the way we undertake much white collar work radically different in 5 years' time and much potentially redundant in 10 years' time. This is likely to apply to railway risk and safety management activities too.

How does this affect the future of safety management? It will mean the replacement of risk assessment with real data. The sacred cow of risk assessment and quantification no longer works for socio technological systems as the world of railways is becoming too complex. Quantification is ineffective as it seeks to model the world using a gross simplification by deriving numbers that imply a large level of accuracy. Approaches like the (RSSB) safety risk model, which uses historical data to determine quantified risk figures for hazards on the railway, would be made redundant as the approach described in the ELBowTie would use real data in real time.

A data driven approach will replace "conventional safety engineering". If one looks through the railway safety and reliability standards, CENELEC EN50126 and the recommended activity at each life-cycle stage it can be seen that nearly all the items are data driven. By using Big Data and Machine Learning the majority of the work undertaken now by safety engineers could be eliminated.

9. Conclusion

The railway industry has advanced a long way with safety methods over the past fifty years and accidents are at a record low, but now there must be a major change if further improvements are to be realised. The approach of blaming 80% of our accidents on human error (Hollnagel 2015) is clearly not right. Safety II, Big Data, the Digital Railway, and Simulation have all come along in the last 5 years or so to produce a virtuous circle that will provide massive synergies. The complexity of modern socio technological systems requires that a new approach is required and the elements discussed in this paper provide a way forward. This will offer railway stakeholders striking opportunities including greater safety, happier workers, better reliability, greater efficiency and less waste.

REFERENCES

BBC Report Bad Aibling Rail Disaster 2016
http://www.bbc.co.uk/news/world-europe-35530538 (accessed 23/3/2016)

CENELEC EN50126

Erdos, G.D., (2004) The Evolution of Rail Safety IRSE Technical Meeting: Adelaide 29 Oct 2004.

Ford M., The Rise of the Robots, Oneworld, London 2015

Gratner_2016 (accessed 23/3/2016) http://www.gartner.com/technology/research/methodologies/hype-cycle.jsp

Hale A, Hoyden J. (1998). Management and culture: The third age of safety.

Hollnagel E., Safety I and Safety II, Ashgate 2015.

IET and BCS Turing Lecture 2016: The Internet of me: It's all about my screens http://www. http://conferences.theiet.org/turing/ (accessed 29/3/16)

Iwnicki S. & Parkinson H. J., Fault Free Infrastructure: Effective Solutions to Improve Efficiency' Derby, England, 23-24 November 1999. Hardcover Published Nov-1999 ISBN 1860582338.. "Assessing railway vehicle derailment potential using Neural Networks". Int. IMechE Conference '

Kapp K. M. The Gamification of Learning and Instruction, Pfieffer, Wiley 2012

Network Rail 2015, Digital Railway Program http://www.networkrail.co.uk/Digital-Railway-Introduction.pdf (Accessed 26/2/2016)

Parkinson H. J., and Bamford, G., The Potential for Using Big Data Analytics to Predict Safety Risks by Analyzing Rail Accidents, Accepted for 3rd International Conference on Railway Technology: Research, Development and Maintenance, Cagliari, Sardinia, Italy 5-8 April 2016

Parkinson H J, Bamford G, and Kandola B, The Development of an Enhanced Bowtie Railway Safety Assessment Tool using a Big Data Analytics Approach, IET Conference ICRE Brussels May 2016.

RSSB Annual Safety Performance Report 2014/15 http://www.rssb.co.uk/Library/risk-analysis-and-safety-reporting/2015-07-aspr-key-findings-2014-15.pdf Accessed 26/2/2016

Yaneira E. S., et al., 2013, Bowtie Diagrams in Downstream Hazard Identification and Risk Assessment, 8th Global Congress on Process Safety Houston, TX, April 1–4, 2012. American Institute of Chemical Engineers.