GB railway

• More than 15,000 km of track
• More than 1.5 billion passenger journeys per year
• Passenger numbers have doubled in the past 20 years
• Currently holding a good record for safety
Safety Risk Model

- Fault-tree-based model underpins railways’ SMSs.
- Allows queries only within the structure of the model.
- We often don’t have data for failures that have not occurred.
- Keeping the model up-to-date is laborious: it can take one to two years to reissue the model.
Real-time data is what we expect the future will be
It’s easy to imagine an alternative...

With modern technology it is easy to envisage:

- automatic collection of real-time data;
- data from many sources: traditional sensors, video, mobile phones, text-based reports, ...

<table>
<thead>
<tr>
<th>In-house data:</th>
<th>External data:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• train movements</td>
<td>• weather</td>
</tr>
<tr>
<td>• passenger movements</td>
<td>• sports events</td>
</tr>
<tr>
<td>• status of signalling</td>
<td>• road traffic conditions</td>
</tr>
<tr>
<td>• maintenance works</td>
<td>...</td>
</tr>
</tbody>
</table>

...real-time picture of risk on the railway
How to do this in practice?

Combining many sources of data has problems, for example how to:

– combine conflicting data?
– use data that goes out-of-date?
– capture usable data from audio, video, text?
– include new data sources as they become available?
– deal with interruptions in data feeds?
– abstract raw data into railway risk assessment knowledge?
– avoid spurious correlations?
– obtain knowledge from a blend of numeric and non-numeric data?
– ...?
The vision

Data processing

Ontology

Visualisation

Enterprise architecture
Can we build it?

Data processing

Ontology

Visualisation

Enterprise architecture
Are we building it?


Stow, Julian, Zhao, Yunshi and Harrison, Chris (2016) Estimating the frequency of trains approaching red signals: A case study for improving the understanding of SPAD risk. IET Intelligent Transport Systems. ISSN 1751-956X (In Press)

Figueres-Esteban, Miguel, Hughes, Peter and Van Gulijk, Coen (2016) Visual analytics for text-based railway incident reports. Safety Science, 89. pp. 72-76. ISSN 0925-7535


Does the industry want it?

- Potters Bar, North London, 10 May 2002
  - Seven people died following a derailment of a passenger train
  - The derailment was caused by degraded points
  - In the preceding days, drivers had reported abnormal track conditions

- With hindsight we can see that information necessary to address the problem had been available before the accident

- When the world is full of data, how can we get the safety information we really need?