Data Sandbox+
Cross-projects meeting

16 January 2020
Welcome and Introduction

Giulia Lorenzini
Senior Partnerships and Grants Manager, RSSB
Objectives of the day

1. Sharing project plans / findings (to date)

2. *For project teams*: start to identify potential synergies and future collaboration opportunities

3. *For industry partners*: gain early visibility of portfolio of work and provide advice / suggestions etc.

4. Network!
Plan for the day

- Update from industry
- Networking session 1
- Data Sandbox+ project presentations
- Networking session 2
- Wrap up & Next steps
RSSB: introduction
RSSB: a brief overview

Through research, standards, and analysis we help our members deliver a better, safer railway
RSSB: R&D Programme

- Driven by industry ideas and needs
- Allows industry to work together to explore new opportunities
- Ensures the insights make a real difference → route to implementation

£10m plus annual grant from DfT

The money comes with requirements to:
- Do truly cross-industry research
- Have mix of high impact/high uncertainty research and shorter term/incremental research
- Leverage co-funding when appropriate
From ideas to implementation

1. Research ideas
   - Driven by industry needs and strategic priorities

2. Defining the research
   - Costs and benefits evaluated and project specified

3. Directly managed
   - Feasibility and demo studies
   - Competitions
   - Strategic partnerships
   - Grant scheme

4. Delivering a project
   - On going monitoring to make sure projects stay on track (time, cost, quality)

5. Implementation of findings
   - Working with partners to ensure research makes the desired impact and where appropriate facilitate implementation

6. Industry
   - Ideas, Priorities
   - Shaping to meet needs

7. Steering Data Access
   - Using the findings
   - Feedback on impact
The Data Sandbox+ Competition
The journey so far...

Summer 2017: Data Sandbox platform was built – powerful example that data from across the industry can be made available...

Oct 2017: £500K ‘call for research’ launched
- 5 projects awarded funding
- Great collaboration from NR + TOCs

April 2018: £1.3m ‘call for research’ launched (co-funded by RSSB and NR)
- Round 1: 4 projects awarded funding
- Round 2: currently shortlisting

Purpose / scope:
- Increased understanding of industry data
- Improved consistency of rail performance from reduced dwell time fluctuation
- Improving disruption management through better recovery from reactionary delays

= Ops savings + improved customer exp.
The ‘expansion’: Data Sandbox +
37% passengers satisfied with how well train company deals with delays (Transport Focus, 2018)

'One team one look' at Victoria station saw a 13% increase in passenger satisfaction

20% Trust, 37% Distrust said they trusted or distrusted train travel

Rail user satisfaction at 10-year low

14.8m 2017/18
15.6m 2018/19
Delay minutes

On time metrics

Performance is still falling back

Rail delays are the 'worst for 12 years'

13% increase in passenger satisfaction
Data Sandbox +

- System Performance
- Timetabling
- Data and Information
The DSP competition so far...

£1.3m competition (co-funded by RSSB/NR)

2 application rounds

Round 1: 4 projects awarded funding

Round 2: currently being shortlisted
Round 1 winners

1. Real-time prediction and mitigation of disruption through personalised passenger communication

2. IntelliDwellTime

3. Data-Driven Robust Timetabling

4. Rail Performance Modelling for Strategic Decision Making
Update from Industry
RSSB: Enabling Better Performance Programme (PERFORM)

Andy Castledine
Performance and Operations Professional Lead
RSSB
The Performance Challenge

Declining trend in PPM

40% passengers satisfied with how well train company deals with delays (Transport Focus, 2019)

Rail passengers face further chaos due to timetables changes and engineering works, report warns

20% Trust 37% Distrust

On time metrics

14.8m 2017/18
16.7m 2018/19
Delay minutes

One team one look at Victoria station saw a 13% increase in passenger satisfaction
The Enabling Better Performance Research challenge (PERFORM)

Fundamental understanding of performance trends and incentives
Gaining better insights on the precursors, influencers and disruptors of performance trends

Managing disruption
Improving the prediction, prevention, response and recovery from disruptions

Operational rules and standards
Revising rules and standards to drive performance

Rail operations and variability
Reducing dwell time, improving consistency of staff behaviours, and effective traffic management on a robust timetable

Value from data
Getting the right data to the right people at the right time

16.7m delay minutes in 2018/19 (12% up since 2017/18)
PERFORM pipeline of key projects underway in delivery, development and ideas

In delivery
- Options to replace detonators (T1155)
- Alternative rules for detonator use during the assistance to a failed train (IMP-DET)
- Brake tests in snow conditions (T1144)

Project ideas
- Optimising brake isolation rules
- Review of ESW and options on the continued use of temporary block working
- Unlocking greater use of high speed coasting, reviewing and assuring associated rules
- Unlocking greater use of GSMR, reviewing and assuring associated rules
- Options to support the analysis and attribution of reactionary delays (COF-ADA)
- Data sandbox+
- Data requirements for investigating trespass (T1182)

Project ideas
- Role specific data driven dashboards to driver performance improvements in everyday operation
- Options to better understand sub-threshold delays
- Mapping and rationalising interfaces of performance data systems

In development
- Toolkit and guidance for better contingency plans, resource utilisation and control competence (T1154)
- Implementation and impact analysis of the T1154 contingency plan toolkit (IMP-T1154)
- Evidence of the decision making model supporting decision-making during disruption (IMP-T1135)
- Methodologies for detecting and preventing trespass (T1168)

In delivery
- Good practice guide to assessing trespass risk (T1183)
- Optimising technologies for effective staff and customer communication during disruption (knowledge search)

Project ideas
- Guidance and tools to support recovery from non-harm trespass
- Predictive, automated and flexible recovery plans

In development
- Options to support the analysis and attribution of reactionary delays (COF-ADA)
- Data sandbox+
- Data requirements for investigating trespass (T1182)

Project ideas
- Role specific data driven dashboards to driver performance improvements in everyday operation
- Options to better understand sub-threshold delays
- Mapping and rationalising interfaces of performance data systems

Operational rules and standards
- Revising rules and standards to drive performance
- Reducing dwell time, improving consistency of staff behaviours, and effective traffic management on a robust timetable
RSSB R&D research portfolio

- Outputs available now and ready to use
- Outputs requiring further work to implement
- Feasibility outputs illustrating the art of possible
- Funding new research and/or identifying opportunities to progress existing outputs through other funding streams

Industry Risk Maturity model – Performance (RM3P)

- Describes what excellent performance management capability looks like using a five-point maturity scale
- RSSB outputs, products and services available to support industry self-assessments and increase maturity

Performance Innovation funding (PIF)

- Potential for project initiation / implementation of R&D outputs funded through PIF
- T1135 is a good example of successful funding through the PIF to enable implementation of an R&D output
Selected PERFORM projects
T1178 - Good train regulation practice in an ‘on time’ railway

- Mix of performance metrics
- New metrics introduced in CP6
- Impact on train regulation
- Opportunity to:
  - Better understand and evidence good train regulation practice
  - Quantify impact of CP6 metrics on train regulation practice
  - Inform Traffic Management

Background

Current Challenges

1. What does good train regulation practice look like when putting customers first?
2. How does capacity utilisation impact train regulation thresholds?
3. How can we help decision-makers make effective changes and updates to regulation policies?
4. How can we help staff apply regulation policies successfully?
5. How can we gather evidence of good train regulation practice to support continuous learning?
6. How would traffic management systems enable more effective train regulation practice?
T1135 - a tool for improving operational decision making

- Operational decision making has become under valued and is not routinely trained
- Tools and models are widely used in other safety critical industries - these have been used in developing the T1135 model
- Evidence indicates these are only effective when supported by training and a fair culture that empowers decision making
- Training package and model developed published
- A RSSB supported in-service pilot currently underway with EMT & NR LNE
  - Includes measurement of benefit realisation
T1154 Enabling better planning and resource management during disruption

- Service disruption undermines the industry’s commitment to delivering the timetable
- Contingency plans set out the principles by which train services will be amended against the base timetable
- The effectiveness of a plan will be strengthened by good resource management and supporting control competence
- Developing toolkits, underpinned by theory and findings reports, on enabling better contingency plans, resource utilisation, and control competence during disruption
- Implementation trial underway with GTR, Scotrail and WMT for December 2019 TT change

- Contingency plans
- Control competence
- Resource utilisation

- Underdevelopment
- Due October 2019
PERFORM: Get Involved

- Deliverables published on www.sparkrail.org
- Share your performance challenges and contribute to research that could drive a step change across the industry
- Receive more information about the PERFORM programme and projects
- Support the development and delivery of PERFORM projects
- Submit your research ideas on enabling better performance

Contact: andy.castledine@rssb.co.uk OR aaron.barrett@rssb.co.uk
RM3P
Risk Management Maturity Model for Performance
What is RM3P

RM3P is an organisational self-assessment tool

Based on the Risk Management Maturity Model for safety.

It is designed to help provide insight into organisational MATURITY

It is scalable, and can be applied to a team, business or industry.
At **LOW** levels of maturity, a task is considered “ad-hoc” – a ‘just do it’ approach.

At **HIGH** levels of maturity, there tasks are carried out to a level of excellence. There is a plan and structure to doing things, and everyone consistently works to improve the way they work.
RM3P has 30 segments.

These areas broadly outline the various factors driving performance that need to function well.

Each area has its own level of **MATURITY**
Assessments

Each segment of the wheel has associated EVIDENCE that help to understand what level of maturity has been reached.

Assessments should:

- approach each subject in a non-judgemental, honest way
- consist of a mix of interviews and evidence gathering
- be collaborative and inclusive

OUTPUTS should be action driven, and focus on achievable improvements.
RSSB R&D alignment to the Risk Management Maturity Model for Performance (RM3P)
Mapping RSSB R&D projects to the wheel

- Over 50 R&D projects mapped to RM3P
- Published over the last 5 years or currently underway identifying:
  - Usable knowledge
  - Usable tools
  - Further work needed
- We are working with industry undertaking RM3P assessments to understand:
  - How R&D research can help self-assessments and be implemented to improve maturity
  - Opportunities to implement R&D outputs, support PIF applications, identify gaps for further R&D
- Existing products, services and expertise available to support RM3 (ORR Risk Management Maturity Model)
ITED

Industry Train Event Database
Background

What is ITED:

A whole-industry data warehousing system containing detailed train running information, to address:

- Sub-threshold investigation limitations
- Inconsistent data availability across regions
- Need for more granular data
7 years data available to on-site / NR users.
2 years data cloud based, available to external users

"Data Mart" Output File Visualisation (PowerBI)

Delay and Lateness
Planned Timetable
TSR's
Ref Data Stannox / Tiploc
SRT’s, Mileage, Speed
Actual Consist
Train Stock and Consist
Crew
Planned – Very Short Term
Actual movement GPS ping level
Actual movement Stannox level
Actual movement Berth level

Performance and Delay Metrics
Planned Timetable
TSR’s
Ref Data Stannox / Tiploc
SRT’s, Mileage, Speed
Actual Consist
Train Stock and Consist
Crew
Planned – Very Short Term
Actual movement GPS ping level
Actual movement Stannox level
Actual movement Berth level

"Data Mart" Output File Visualisation (PowerBI)
Data Flow

• ITED contains a number of data sources, summarised below and to the right:

• WEB GEMINI – Planned Consist information
• GENIUS – planned and actual crew information
• PSS data sources
  • Incident
  • Delay
  • Reference
• TABS – Actual consist
• ITPS + VSTP – Long and Short term scheduling information
• GPS – GPS pings from activated trains
• SMART
  • Movement records
  • SClass
• TRUST – Movement records
• PPS – Temporary speed restriction
• BPLAN – Sectional Running times
• Various reference data used to bring all of this information together
Phased Implementation

Phase 1 Implementation

“Data Warehouse”
- Initial training to focus on formation of useful data sets and exporting to third-party tools
- Combined SU-Training and User Acceptance Testing (2 weeks)

Training and Testing

Design Phase 2
End-User Training

Explore Capability

Super Users able to explore new data sets and encouraged to look at new ways of working, provide feedback to National Performance

What worked? Feedback to inform toolset and training development.

Phase 2 Implementation
MS PowerBI

Provide full-access to all users and new MS PowerBI visualisation and sharing tools

Support Processes Published
Use Cases:

- **Headways** and order of trains
- Understanding **conflicting moves**
- **Turnround** and diagram **resilience**
- Understanding **variation** to do with weather / driver behaviour / rolling stock
- **Regulation** analysis
- **Autumn** analysis (breaking and slipping)
- TSR board placement
- Quantifying a **TSR impact**
- Which **line** a train is running on
- Loss of time due to running on slow line
- Better insight into **sub threshold delay**
- Better uses for **attribution** i.e. One Engine Only and coasting for fuel saving.
- **Verifying berth offsets** without having to manually visit a location.
- **Section Running Time** comparisons to Plan
- **Dwell Time** comparisons to Plan
- Performance impact of **adding another train**
Operational Performance Analysis and Improvement

Network Rail and Amey Consulting
Amey History

Network Rail South East - 2016

Analyst led performance improvements - 2018-19
Section Runtime Validation
TSR Impact Assessment
Operational performance Investigations
Wessex, Anglia, LNE, ARL, Southeastern

Tool development and Rollout - 2017-19
Cosmo Analytics
Quartz – Real-time Station Dashboard
South East, Anglia - GTR, Greater Anglia, ARL, MTR

Current Engagement – 2019/2020
Shape of the Team

Two dedicated individuals in each Region: a Quartz rollout champion, and a data scientist

Embedded in Region performance teams, reporting to heads of performance

Flexible – resources will move between Routes or Regions as and when necessary

Team of six data scientists embedded in NPAT (four Amey, two NR)

Parallel software support team responsible for Quartz product management and ongoing development

Whole team comes together weekly in London to share analysis, knowledge and ideas
Future focus

- Station Readiness
- Signal Operations
- Maintenance Execution

Knowledge Transfer and Upskilling
Reactionary Delay
TRUST and TD

**TRUST**
- arrivals and departures at stations and junctions
- actual vs planned times
- same root source as TD, but times truncated to the minute
- unique train id

**TD**
- trains passing signals
- accurate to the second
- non-unique train id

**SMART berth offsets**
- maps signals to stations
- includes time offsets for each signal (how long it takes to get from signal to platform)
TRUST and TD

TRUST
- arrivals and departures at stations and junctions
- actual vs planned times
- same root source as TD, but times truncated to the minute
- unique train id

TD
- trains passing signals
- accurate to the second
- non-unique train id

SMART berth offsets
- maps signals to stations
- includes time offsets for each signal (how long it takes to get from signal to platform)

Amey database
- arrivals and departures at stations and junctions
- actual vs planned times
- times accurate to the second (assuming offsets are correct)
- unique train id
- since September 2017

real-time join

real-time join

real-time feed

real-time join

static file

real-time feed

real-time join

Amey analysts/developers
NR analysts
Database access

- For train-movement data since September 2019, ITED is the recommended source.
- ITED data is more granular (including every signal, not just stations/junctions).
- Data before September 2019 is not in ITED, so our database is an alternative for station-level timings
- We have provided access to a read-only replica database to NR analysts, via a flexible hosted analytics environment (Jupyter Hub):

  ```python
  import rolldb as rdb
  e = rdb.get_ra_engine()
  df = rdb.read_sql(e, 'SELECT * FROM public."mergedMovements" WHERE train_service_code = '25905000' and inferred_origin_date = '2019-01-03'')
  df.to_csv('/w/data/merged_movements_service_code-25905000.csv')
  ```

- Discussions are underway to provide similar access to RSSB Sandbox project teams
Overview of DISIC and Rail Data Council

Tim Wood, RDG and Liz Davies, RSSB
The purpose of the Data and Information System Interface Committee (DISIC) is to provide a focus for the governance, management, and exploitation of data, information and knowledge in a safe, cost-effective and efficient manner.

DISIC members shall collectively provide expertise and strategic knowledge of:

(i) Data, information and knowledge architecture and management
(ii) Systems architecture and interfaces
(iii) Data and information governance
(iv) Cyber security
(v) Data exploitation, reporting and analysis
Rail Data Council – governance of RSD and JRDP

- Hugh Clancy + Enrique Fernandez
- James Bain
- Other tbc?
- Aidan Hancock
- Emma Campbell

**Rail Data Council**

- **Working groups**
  - Data use & access
  - Data standards & quality
  - Better value from data
  - Innovation, culture & skills

- **Led by...**
  - Simon Moorhead: RDG
  - Karl Butler-Garnham: Network Rail (in RSSB’s DISIC)
  - Ian McLaren: TOC
  - John Easton: UKRRIN

**Focus**

- **Marketplace**
  - Value generation from an open and collaborative industry providing simple access to data

- **Governance**
  - Industrialise and standardise, access, quality, standards & operating model allowing simple use of rail data

- **Value**
  - Business case justification based on GB Rail Data driving Customer Satisfaction then P&L enhancements

- **Innovation & Culture**
  - Review required to drive future culture state to become more inclusive and easy to work with/in
Vision: To drive maximum value in customer service, industry efficiency and growth through better and easier access to data.