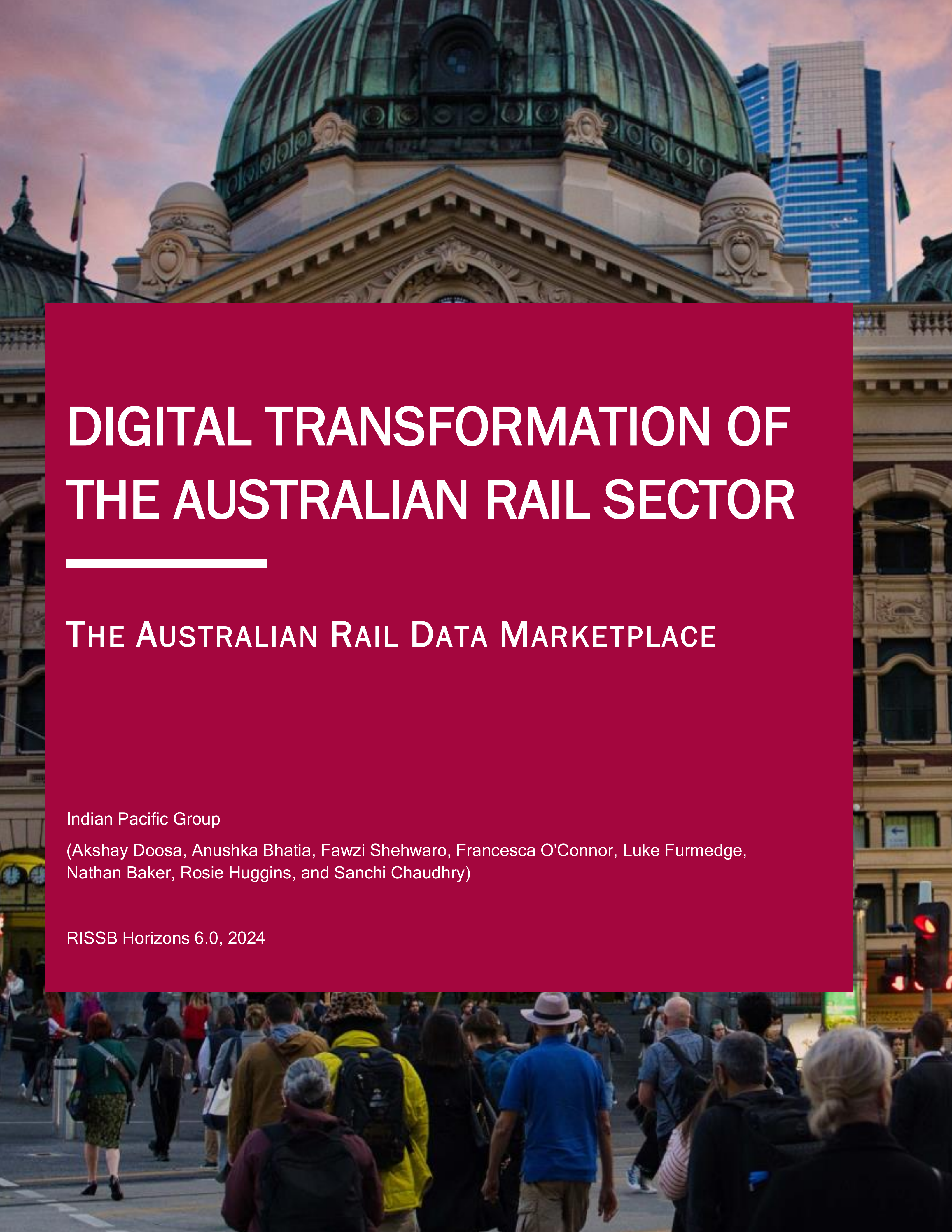


DIGITAL TRANSFORMATION OF THE AUSTRALIAN RAIL SECTOR

The Australian Rail Data Marketplace

Indian Pacific Group

(Akshay Doosa, Anushka Bhatia, Fawzi Shehwaro, Francesca O'Connor, Luke Furmedge, Nathan Baker, Rosie Huggins, and Sanchi Chaudhry)

RISSB Horizons 6.0, 2024

Executive Summary

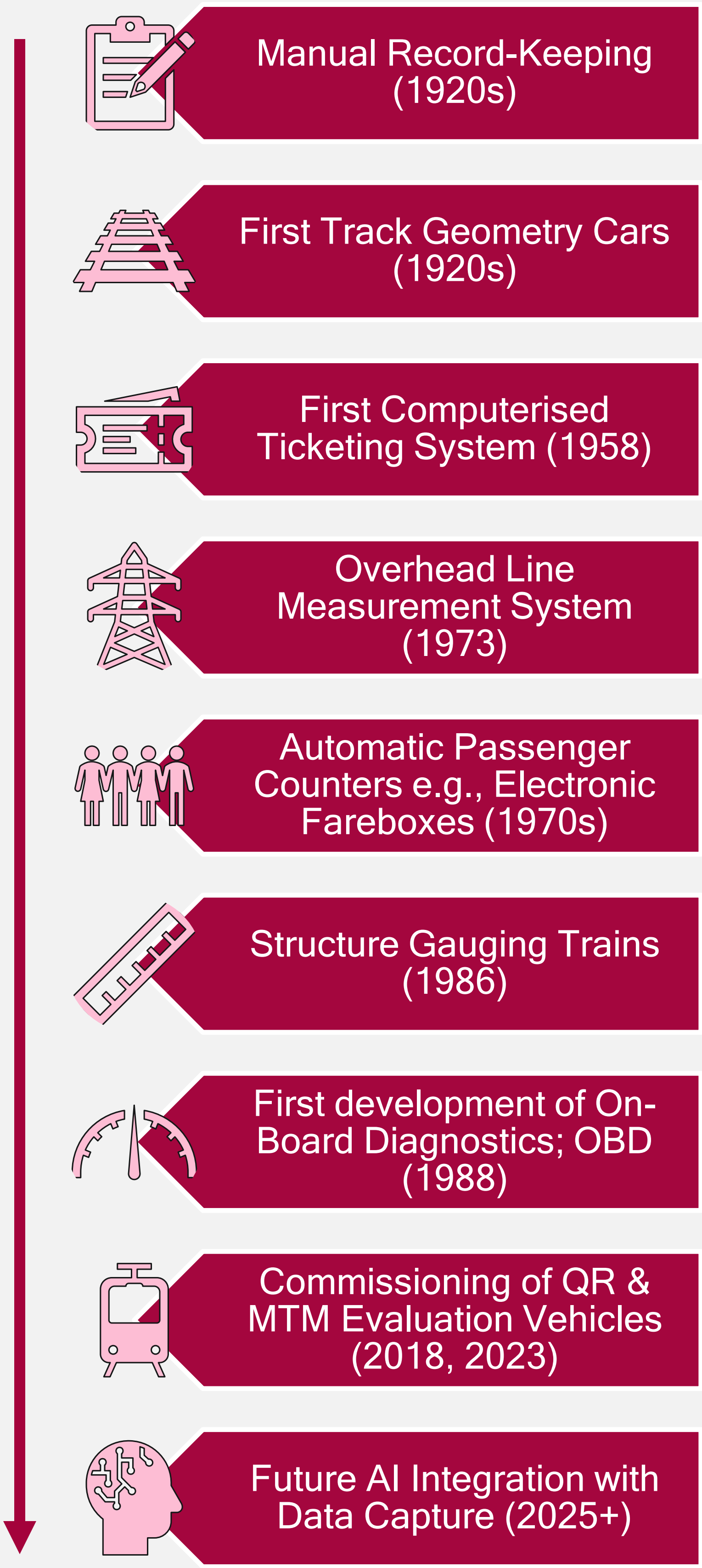
Over the past century, there has been a rapid advancement in rail data collection, from manual record-keeping of the 1920s, all the way to current digital capture. The potential benefits of shared digital data are derived from the availability, proper storage, and industry collaboration, which could transform Australian railways if leveraged correctly.

The Australian Rail Data Marketplace (ARDM) is proposed to be a transformative initiative aimed at unlocking the potential of existing rail data in Australia by improving access, fostering innovation, enhancing operational efficiency to inform better decision-making. The core of the ARDM is its data hub, which aggregates data from multiple sources across the rail industry. This includes both public data and private data that can be accessed under specific agreements.

The ARDM can drive innovation as operators and infrastructure managers throughout Australia benefit from readily available digital solutions from any corner of the country. For instance, the ARDM could help distribute rail training data for machine learning and trained AI models to detect defects in rail. Metadata and cells to build and expand digital twins could be shared on the ARDM, which could serve as building blocks for building models and benefit organisations with limited resources. 3D files for spare parts designed by one organisation could benefit any other organisation with access to a 3D printer.

The Australian Rail Data Marketplace (ARDM) is a pivotal step in the digital transformation of the rail sector. By centralising and democratising access to rail data, the ARDM has the potential to enhance productivity and unlock unprecedented opportunities for collaboration and technological advancement across the Australian railway network.

Advancements in Railway Data Capture



Over the past 100 years, advancements in rail data collection have transformed the industry's ability to monitor and optimize railway operations and maintenance. Initially relying on manual methods of record-keeping, such as schedule books and stationmaster passenger counting, the sector has evolved to adopt infrastructure evaluation vehicles, real-time monitoring systems integrated into fleets and stations, and other state-of-the-art tools for data capture.

These advancements have enabled rail networks to capture comprehensive data on train diagnostics, patronage data, infrastructure condition and asset reliability, vastly improving operational efficiency, safety, and predictive maintenance. Today, technologies like IoT and AI further enhance the accuracy and speed of data collection, providing a near-continuous stream of valuable insights.

Despite these impressive gains, challenges remain in data reliability, storage, and collaboration. Current systems often operate in silos, limiting the full potential of shared data across the rail ecosystem. Addressing these limitations could unlock significant opportunities for enhanced decision-making, cross-industry collaboration, and a more integrated, efficient rail industry.

Historical Data Capture

In the early 20th century, data capture was relatively primitive, with an over-reliance on manual records and book-keeping. Faults were often reactively managed, and incidents were more common due to the lack of overall asset condition monitoring.

It wasn't until the midst of the 'Information Age' in the 1970s that digital technologies were properly integrated into data collection for railway infrastructure maintenance. In Scotland, the first Mobile Electrical Network Testing, Observation and Recording train (MENTOR) came into operation in 1973 to support a programme of electrification. It was used to observe and report on the height, stagger, and contact of the overhead line with the train pantograph.

Pre-19th century
A handheld clicker counter used by a stationmaster to count passengers boarding.

1990s - Early 2000s
A railway worker uses a laser height and stagger gauge to measure the position of the overhead line contact and catenary wire.

Similarly, for operational data management, technologies have advanced from paper-based records for operational, security and emergency incidents to more intricate technologies, such as Video Surveillance (CCTV) and Automatic Passenger Counting (APC) Systems to monitor patronage at stations.

The MTM EV120 is equipped with state-of-the-art sensors, cameras and ground-penetrating radar. 'Evie' examines the condition of tracks, overhead lines, and ballast under the tracks in real time, ensuring any infrastructure faults are promptly identified to maintenance teams.

DATA CAPTURE IN THE PRESENT DAY

Today, Rail Transport Operators (RTOs) utilise sophisticated technology to collect patronage, incident, and infrastructure data.

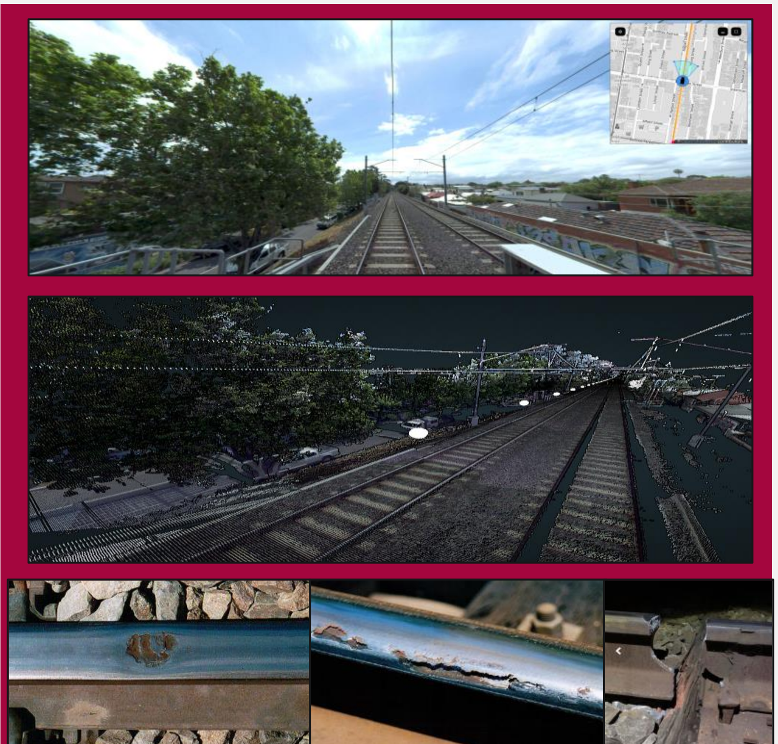
CASE STUDY: Australian Infrastructure Evaluation Vehicles

Metro Trains Melbourne, Sydney Trains, and Queensland Rail each utilise Infrastructure Evaluation Vehicles (IEVs) to continuously monitor and report on the condition of their infrastructure assets. Examples of the technology used on these cutting-edge track machines include the following:

- Track Geometry, Clearance and Rail Profile Measurement
- Rail Joint and Surface Imaging
- Overhead Wiring Measurement and Imagery

Modern evaluation vehicles, such as those utilised by MTM, QR and Sydney Trains, are equipped with advanced sensors, LiDAR, and other imaging technologies, allowing for the automated and continuous collection of vast amounts of data as they travel along railway lines.

These vehicles capture high-resolution imagery of both the tracks and overhead infrastructure in real-time. The quantity of data gathered is difficult to comprehend; a typical annual operation can result in over 60 terabytes. This paradigm shift in technology has significantly enhanced network safety and reduces the time spent performing measurements that were traditionally undertaken manually.



Panoramic data imagery, LiDAR technology, and examples of track defects detected on modern IEVs.

CASE STUDY: Safety Models; ALCAM, ARRM and ONRSR Notifiable Occurrences.

- The Australian Level Crossing Assessment Model (ALCAM) uses a data management system to provide a safety assessment tool for individual level crossings across Australia and New Zealand.
- The RISSB Australian Rail Risk Model (ARRM) collects information from participating rail operators, built around hazardous events that occur on Australian railways. The system produces risk reports around rail safety risk to assist industry in understanding current levels of risk and trend analysis.
- All Rail Transport Operators (RTOs) are required to report notifiable occurrences under Rail Safety National Law (RSNL) to the Office of the National Rail Safety Regulator (ONRSR). Occurrences range in category and severity; ONRSR collect the information and ensure RTOs are effectively reporting and investigating.

Each RTO collects and stores their own safety data (including notifiable occurrences that are required for ONRSR). This data includes safety alerts, near-miss information, safeworking communications / directives, etc.

Additionally, each of the above assessments requires RTOs to distribute their data to meet regulatory reporting requirements and help inform industry-wide models. If this data were loaded into a centralised system that could understand the nuances between RTOs, then it would be able to undertake trend analyses, highlight commonalities / shortfalls in humans' interpretation, and dig deeper into the question of why an alert / near miss occurred in the first place.

FUTURE UTILISATION OF DATA

Despite the incredible opportunity that has been unlocked in these data collection case studies, challenges remain in storage, collaboration, and utility. The sheer volume of data generated is immense, requiring robust and scalable storage solutions.

Moreover, the collaboration between Rail Infrastructure Managers (RIMs), ONRSR, and other industry participants is hindered by fragmented data systems and varying data standards. The ability to effectively process and utilise this data for state and franchise-level decision-making is often constrained by both poor data quality and siloed approaches in determining which infrastructure requires funding.

These issues highlight the need for integrated solutions, such as a central data platform, to manage and streamline the use of this valuable information.

The Problem: Data Storage and Collaboration

The Australian railway industry operates on a state-centric approach with public and private companies operating and maintaining their own networks independently as Accredited Rail Transport Operators (ARTOs) and Rail Infrastructure Managers (RIMs).

As a result, rail operators / maintainers, state-level decision-makers, regulators, and other industry participants lack national collaboration and visibility of infrastructure, operational and safety issues. The lack of a centralised platform can cause issues with accurate development of business cases based on up-to-date infrastructure, operational and safety data which would better serve the needs of the community. This challenge is further exacerbated by big infrastructure projects that are delivered by non-operator stakeholders.

Current limitations / constraints across industry include:

LACK OF UTILISATION

Data is not being stored in a centralised system to inform investment.

PIPELINED DECISION-MAKING

Decision-makers are hampered by bias and inaccurate information.

NO SINGLE SOURCE OF TRUTH

Data is spead across a number of platforms with variable quality.

LACK OF COLLABORATION

Data is not aggregated and transparent between parties.

The Solution: An Australian Rail Data Marketplace

The Australian Rail Data Marketplace (ARDM) is proposed to be a transformative initiative aimed at unlocking the potential of rail data in Australia. By improving access to data, the ARDM supports innovation, enhances operational efficiency, and contributes to better decision-making within the rail industry.

A central platform for data sharing in Australia's rail network will improve operational coordination, safety, passenger services, and maintenance. Sharing real-time data between operators and infrastructure managers optimises network management and allows for predictive maintenance and better emergency responses.

Ultimately, the ARDM will help build a more modern, efficient, and passenger-friendly rail network, aligned with broader goals of sustainability and technological advancement. Leveraging the vast data sources of individual RTOs and other independent bodies, the ARDM will apply state-of-the-art digital technologies in the form of AI, digital twin and machine learning to provide ARDM users a tailored and detailed report based on a query input.

The phased implementation of the ARDM is summarised in the figure below. Phase 1 identifies the key stakeholders, carries out consultation and identifies a data governance framework. The data from all parties will be uploaded in the marketplace. Once ARDM is established and data is uploaded, Phase 2 proposes a platform where users could search and access railed related data for the various purposes highlighted further in the report. Phase 3 looks to the future; to optimise data to give recommendations and enable innovative solutions such as 3D printing.

PHASE 1: ESTABLISH ADMINISTRATION

- Establish a data governance framework and engage with stakeholders.
- Implement robust security, compliance and access control measures.
- Foster continuous auditing, training and third party management.

PHASE 2: IMPLEMENT SOLUTION

- Facilitate access to a wide range of rail-related data through a data hub.
- Support a marketplace model with varied data access.
- Caters to various users including rail industry professionals, transport operators, and policy makers.

PHASE 3: DRIVE DIGITAL INNOVATION

- Compare data across Australia and provide detailed recommendations.
- Facilitate sharing of digital assets and AI models.
- Enable innovative solutions such as blockchain for data security and 3D printing for spare parts.

Purpose and Benefits of ARDM

The ARDM proposes to serve multiple purposes to enhance the rail industry in Australia, fostering innovation, operational efficiency, and collaboration across stakeholders.

Enhance Data Accessibility

- Centralises rail-related data for stakeholders (e.g., operators, service providers, government bodies, researchers, and developers).
- Breaks down industry data silos, making data more accessible for innovation and service improvement.
- Sharing data helps identify areas for standardization and harmonization across different networks.

Drive Innovation

- Availability of data enables development of new applications, tools, and services.
- Examples include improved journey planning tools, real-time passenger info systems, and predictive maintenance solutions.

Improve Operational Efficiency

- Facilitates data sharing to optimize train scheduling, capacity management, and disruption response.
- Enhances collaboration across industry organisations.

Reduce Operational Costs

- Reduces the need for additional data collection activities (e.g., site inspections, geotechnical investigations).

Support Environmental / Sustainability Goals

- Enables the development of strategies to reduce the rail carbon footprint and enhance rail operational sustainability.
- Supports innovations for making rail travel more efficient and sustainable.

Empower Passengers

- Provides real-time data for better journey planning, disruption avoidance, and improved travel control.
- Increases transparency and informed decision-making for passengers.

Facilitate Research and Policy Making

- Offers comprehensive data for tracking performance, analyzing trends, and making informed policy decisions.
- Sharing data can help ensure compliance with regulatory requirements and industry standards.

Phase 1: Establish Administration

To effectively implement the ARDM, establishing a Board of Directors is essential. This Board will include industry experts, financial professionals, legal advisors, experienced executives, stakeholders from RTOs, and representatives from regulatory bodies like ONRSR. The Board will set strategic direction, ensure regulatory compliance, oversee financial health, guide operational effectiveness, and establish strong data governance and ownership. This comprehensive approach will ensure data security, compliance, and operational efficiency while safeguarding intellectual property (IP) in line with Australian legislation and standards to optimize the process.

The following ten principles outlines such an approach:

1. Data Governance Framework

- **Governance Structures:** Develop a data governance framework featuring a Data Governance Committee responsible for overseeing data management practices and legislative compliance. At Sydney Trains, this function is managed by the Cybersecurity team. Ensure IP management policies are integrated to protect patents, trademarks, and copyrights as per the Australian Patents Act 1990 and the Trademarks Act 1995.
- **Data Ownership and Stewardship:** Assign data owners and stewards to manage various data assets. Data owners are accountable for the accuracy, security, and compliance of data, while data stewards handle day-to-day data management and IP governance. These roles are supervised by the Data Governance Manager, who supports the Chief Information Officer (CIO) in implementing and overseeing the overall data strategy of the ARDM.

2. Data Classification and Management

- **Data Classification:** Implement a data classification scheme to categorise data based on sensitivity and value. This will ensure that IP contributed to the ARDM is subject to appropriate security measures and compliance controls.
- **Data Management Policies:** Develop and enforce policies for data including IP collection, usage, retention, access, and disposal. Ensure these policies comply with the Australian Privacy Principles (APPs) as outlined in the Privacy Act 1988 and relevant IP laws.

3. Access Control

- **Role-Based Access Control (RBAC):** Implement RBAC to restrict access based on user roles and responsibilities. Ensure only authorised personnel can access sensitive IP stored on the ARDM.
- **Multi-Factor Authentication (MFA):** Require MFA for accessing the ARDM to add an additional layer of security.

4. Data Security and Compliance

- **Encryption and Protection:** Encrypt ARDM data both at rest and in transit to protect against unauthorised access and breaches. Adherence to the Australian Cyber Security Centre's Essential Eight and ISO/IEC 27001 standards will reinforce data security.
- **Compliance with Legislation:** Ensure ARDM's data governance practices comply with Australian laws, including the Privacy Act 1988, the Security of Critical Infrastructure Act 2018, and IP legislation such as the Patents Act 1990 and the Trademarks Act 1995 to protect IP rights.

5. Network and Physical Security

- **Network Segmentation and Security:** Implement network segmentation to restrict access to sensitive data, perhaps releasing sensitive data as a redacted version followed by a formal data request to enhance protection against breaches. Utilise firewalls and intrusion detection/prevention systems to secure network boundaries.
- **Physical Security:** Enforce stringent physical access controls and monitoring to safeguard ARDM's data storage facilities and equipment.

6. Data Backup and Recovery

- **Regular Backups:** Ensure that the ARDM's critical data is backed up regularly and securely. Data owners should verify the completeness and accessibility of backups.
- **Disaster Recovery Planning:** Develop and regularly test ARDM's disaster recovery plans to ensure data can be restored and operations can continue following a major incident.

7. Compliance and Auditing

- **Regular Audits:** Conduct periodic data audits on the ARDM to verify compliance with governance policies, legal requirements, and IP management. Data owners should oversee these audits and address any issues identified.
- **Documentation and Reporting:** Maintain comprehensive records of ARDM's data governance practices, security measures, and compliance activities. Report data breaches / IP breaches or compliance failures as mandated by the Notifiable Data Breaches (NDB) scheme, including notifications to affected individuals and the Office of the Australian Information Commissioner (OAIC) when breaches are likely to cause serious harm.

8. Employee Training and Awareness

- **Security Awareness Training:** Provide continuous training on data protection, cybersecurity practices, and compliance requirements. Ensure ARDM staff understand their roles in safeguarding data including IP.
- **Role-Specific Training:** Customize ARDM's training programs to address the specific responsibilities of data owners, stewards, and other relevant personnel.

9. Vendor (Data Provider) Management and Third-Party Risk

- **Vendor Assessments:** Since data including IP will be collected from third parties, it is vital to evaluate third-party vendors for compliance with data security and privacy standards. Ensure commercial agreements include data protection clauses to assist in correctly downloading potentially sensitive data from the marketplace.
- **Third-Party Data Handling:** Monitor and manage third-party access to data, ensuring vendors adhere to proper data handling, IP security practices and commercial agreements for monetisation of data.

10. Incident Response and Management

- **Incident Response Plan:** Develop and implement a comprehensive incident response plan for ARDM to manage data breaches or security incidents, including those involving IP. Data owners should lead response efforts and coordinate with relevant stakeholders of the marketplace.
- **Incident Documentation and Reporting:** Thoroughly document all incidents and breaches to the marketplace and report them as required by Australian legislation, including the Privacy Act's NDB scheme and IP related regulations.

Phase 2: Australian Rail Data Marketplace Implementation

For implementation of the ARDM, the principle features are as follows:

1. Data Hub

The core of the ARDM is its data hub, which aggregates data from multiple sources across the rail industry. This includes both public data and private data that can be accessed under specific agreements. Examples of data that may be shared include:

Station Accessibility	Quantity of accessible carparks	Platform heights and offsets	Boarding ramp dimensions
Environmental	Carbon footprint per journey	Noise pollution levels	Heritage protected areas
Ticketing	Fare calculations	Price per km travelled	Concession / senior discounts
Patronage	Lost customer minutes	Daily / Total station entries	Customer feedback / complaints
Incidents	Trespasser management	Safety investigations	Network downtime due to planned disruptions
Power Consumption	Peak vs. offpeak wattage	Energy use per fleet	Untinterruptible power supplies
Infrastructure	Track gradient / curvature	Structure gauge clearances	Overhead wiring structure types
Train performance	Line speeds	Train on-board diagnostics	Punctuality and delivery targets

2. Marketplace Model

The ARDM operates on a marketplace model, where data providers can make their data available, and users can discover, access, and use that data. Some data might be available for free, while other data sets could be offered on a commercial basis, depending on the agreements between providers and users.

Access to the marketplace requires authorisation or subscription, and the level of access may depend on the user's role and the nature of their data needs. Accessing the ARDM requires contacting marketplace administrators.

3. Collaboration and Community Features

The platform includes features that facilitate collaboration between different stakeholders, such as forums, user groups, and partnership opportunities. This helps create a community around rail data, fostering sharing of ideas and best practices.

4. Users

The ARDM is a specialized platform that provides access to a range of rail-related data and analytics. Access to this marketplace typically involves several categories of users.

- Rail Industry Professionals - Includes rail operators, infrastructure managers, and other stakeholders involved in rail operations.
- Data Providers - Organizations supplying data related to rail operations, such as sensor data and performance metrics.
- Researchers and Analysts - Individuals or institutions engaged in rail industry research or analysis for academic studies or industry reports.
- Technology Providers - Companies developing solutions for the rail industry, such as software developers and system integrators.
- Regulatory and Policy Makers - Government agencies and policymakers needing data for regulations, policies, or industry performance monitoring.

Phase 3: Digital Innovation

Below are some use cases for ARDM to enable innovative use of digital technology in railway infrastructure, operations and maintenance.

INFRASTRUCTURE

The Cross River Rail project for instance has built a detailed and up-to-date 3D model of the project and Brisbane.

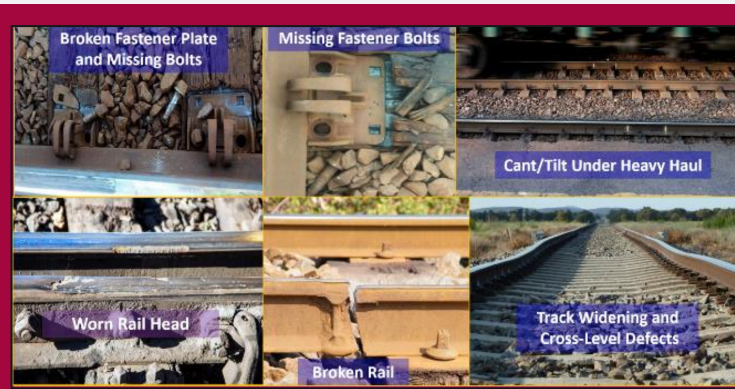
Digital assets used on the project could be shared on ARDM to serve as building blocks for project or city models, benefiting organisatons with limited resources.

Woolloongabba station model by the Cross River Rail Delivery Authority

OPERATIONS AND MAINTENENCE

Machine learning algorithms can be trained on images of railway defects to perform real-time detection and automatic reporting of defects.

These images of defects from all over the country can be hosted on the ARDM and used to train an AI model to detect faults on any rail network. Trained AI models do not need to be state or region specific. Similarly, trained AI model weights can be hosted on the ARDM to benefit operators and infrastructure managers with limited resources.



Track Integrity Monitoring System (TIMS) by Avante

Additionally for operational security, blockchain distributed ledger technology can be coupled with the ARDM for a decentralised and immutable ledger of data transactions, making operational data obtained from the marketplace verifiable and traceable.

ROLLINGSTOCK

Spare part costs make up the biggest cost percentage of rollingstock maintenance at 13.8%. On top of affordability, availability of rollingstock spares is one issue that can be resolved by innovative use of 3D printing rollingstock spares. 3D files for spare parts could be hosted on the ARDM to benefit any organisation with access to a 3D printer.

3D printed rollingstock parts by Roboze

Conclusion

In conclusion, the Australian Rail Data Marketplace (ARDM) signifies a major advancement in the digital transformation of the rail sector. By centralising and sharing access to rail data, ARDM enhances productivity and unlocks new opportunities for collaboration and technological progress.

This initiative is set to revolutionise the Australian railway network by driving innovation and operational efficiency through improved data access and industry-wide cooperation. Transitioning from paper-based record-keeping to an extensive data ecosystem, ARDM integrates and makes vast amounts of data available with simple storage and cataloguing, offering a digital library to solve problems more efficiently. Its potential to distribute data, support real-time responses, and identify trends exemplifies its transformative impact.

By building on existing platforms to create a common data system and leveraging available solutions to overcome challenges such as security, storage and awareness, the ARDM can ultimately driving productivity across the industry. This makes ARDM a cornerstone for future developments in the Australian rail industry.

Further Reading

1. Shi, X., Zhou, Z., & Wang, H. (2022). Cost of rolling stock maintenance in urban railway operation: Literature review and direction. ResearchGate. https://www.researchgate.net/publication/359741814_Cost_of_Rolling_Stock_Maintenance_in_Urban_Railway_Operation_Literature_Review_and_Direction
2. Esri Australia. (2023, July 12). Australia's fastest-growing city: A digital twin guides rail expansion. https://esriaustralia.com.au/blog/australias-fastest-growing-city-digital-twin-guides-rail-expansion
3. Roboze. (n.d.). 3D printing in the rail industry: Reduce the need for spares with Roboze. http://roboze.com/en/resources/3d-printing-in-the-rail-industry-reduce-the-need-for-spares-with-roboze.html?utm_source=pocket_shared
4. Avante Technology. (n.d.). Avante track integrity monitoring system (TIMS). https://www.avantetech.com/avante-track-integrity-monitoring-system-tims
5. Najjar, M., Borana, A., & Alomainy, A. (2023). Detection of railway defects using machine learning and edge computing. arXiv. https://arxiv.org/abs/2408.15245
6. Rubinsztejn, Y. (2011). Automatic detection of objects of interest from rail track images (Doctoral dissertation).