AiRAP Automation for Australian Road Safety

Final Report

About this project

This research is funded by iMOVE CRC and supported by the Cooperative Research Centres program, an Australian Government initiative.

The project partners are Transport for NSW (Lead), the University of Technology Sydney and the International Road Assessment Programme.

iMOVE Australia Limited

Lead Project Partner

Transport for NSW

Other Project Partners

University of Technology Sydney

International Road Assessment Programme

Special thanks to the project team

TfNSW
Joseph Le
Akinola Ekundayo
Joyce Tang

iRAP
Rob McInerney
Monica Olyslagers (Project Lead)
Yidi Yan
Luke Rogers
Agustin Centeno

UTS
Arno Schaaf
Simona Mihaita
Adam Berry
Tuo Mao

Anditi
Peter Jamieson
Bohan Chambers

TomTom
Jennifer Loake
Joel Blatchford
Stan Li

This report is produced by the International Road Assessment Programme for, and on behalf of, the project partners. Enquiries can be made to the Project Leader:

Monica Olyslagers
iRAP Safe Cities and Innovation Specialist
monica.olyslagers@irap.org

Photo on cover page by Patty Jansen from Pixabay.

20 May 2022

Foreword

by Transport for NSW (Lead project partner)

Between 2016 - 2020, 1,753 people were killed and over 25,638 people serious injured in road crashes in NSW.

As custodians of the NSW road network, we are guided by two road safety targets.

- The first is the National Road Safety Strategy 2021-2030 which aims to deliver significant reductions in road trauma on our roads and puts us on the path to Vision Zero.
- The second is Towards Zero, the aspirational target of zero fatalities and serious injuries on our roads by 2056.

In the past, crash locations are treated after crashes had occurred. Now, we are moving to a more predictive approach where potential crash locations are risk assessed and mitigated before lives are lost.

AusRAP data is currently used to inform policy, investment and design decisions across the country with more than 300,000km of roads assessed nationwide. AusRAP data has been used by Transport for NSW to inform priority sections of roads across NSW to target for upgrades through the NSW Safer Roads Program ($822 million between 2018/19 and 2022/23). The data has similarly been used by Transport for NSW for funding submissions to the Commonwealth Government.

Transport for NSW has access to this data for the 18,000kms of state managed roads in NSW supporting its overall effectiveness in prioritizing sections of roads for investigation. The existing asset management data held by Transport for NSW also provides additional confidence and supports the calibration and cross-checking of the AusRAP process on state managed roads.

To ensure road safety is addressed proactively across the State, there is a need for more efficient and cost-effective data to drive road safety decision making on State, regional and local government roads. This would improve the management of the State's road assets by ensuring this data is more reliable and up to date, and ensure safety can be adequately monitored and managed using AusRAP. For local government, it would address a critical need for more consistent asset management data and for common and repeatable approaches to improve safety of local government road networks for all road user types.

Transport for NSW initiated this project in partnership with the University of Technology, Sydney, and the International Road Assessment Programme (iRAP) to explore the potential of new accelerated and intelligent data collection methods to guide investment and save lives.

The desired outcome of the iMOVE research project is to assess whether new artificial intelligence and machine-learning based data collection and feature extraction methods, delivered as part of the global *AiRAP* data partnerships, will improve the accuracy, reliability and scalability of capturing AusRAP data for the New South Wales road network and beyond.

NSW GOVERNMENT Transport for NSW

Executive summary

The NSW road network is almost 185,000 kilometres in length and carries more than 60 per cent of the freight moved in the State.[1] Of these, approximately 20,000km is designated as the State road network, 4,200km form part of the national road network, the remaining of which form part of the regional road (~12,800km) or local road (~148,000km) networks.[23]

Over 90 percent of the 7.5 million passenger journeys made each day are by car, and 63 percent of freight movements in Regional NSW by volume are by road.[4] The cost of injury and fatality road crashes in NSW between 2011-15 was estimated to cost $35.7 billion.[5]

Transport for NSW (TfNSW) has an integrated road asset management system, SRRA. SRRA has the capability to serve multiple, complementary uses, including the of safety Star Ratings on an ongoing basis. SRRA uses a consistent data specification for its data inputs which are aligned to those of AusRAP.

Cheaper and more efficient data sources are needed, not just to improve the reliability and performance of SRRA for the State road network, but also extend to that for local government to improve safety awareness and asset management for those larger road networks.

The aim of this research and development project was to explore potential 'off-the-shelf' data sources, such as LiDAR and probe data, which could be integrated into SRRA and used for AusRAP road safety assessments. In doing so, the project outcomes would lead to improved efficiency and accuracy of road infrastructure data and deliver substantial, long-term cost and efficiency savings to TfNSW.

The broader aim of the project was to do this within a scalable and repeatable framework for the benefit of other Australian jurisdictions at the national, State and local levels, as well as internationally. This was to be achieved through a new 'AiRAP' initiative lead by the International Road Assessment Programme (iRAP).[6] The AiRAP framework developed under this project would provide the mechanism for existing and emerging data sources to feed network-level road safety assessments throughout Australia and worldwide.

Project objectives

There were five clear objectives for the project:

1. Develop the AiRAP Framework
2. Prove AiRAP Framework and scalability
3. Prove independent feature extraction
4. Create an AiRAP supplier map and estimate efficiency savings
5. Recommend Phase 2 research priorities

---

[1] Source: TfNSW

[2] Source: Infrastructure NSW

[3] Source: NRMA

[4] NSW Bureau of Transport Statistics 2006, *Journeys to Work in Regional NSW* and Transport for NSW 2012, Draft Transport Master Plan cited in Infrastructure NSW's State Infrastructure Strategy.

[5] Source: NRMA. (2017). The Cost of Crashes: An analysis of lives lost and injuries on NSW roads. Available online.

[6] iRAP is the umbrella organisation for AusRAP and is the custodian of the global data specifications for safety Star Ratings.

Key achievements

Upon completion, the project had successfully:

- ✓ Delivered 20,000km of attribute data (7 attributes) using automated data extraction methods from existing data sources
- ✓ Proved viability of 'off-the-shelf' data for 52 of the iRAP safety attributes with 34 attributes being accredited under AiRAP.
- ✓ Established an accreditation framework for doing the same for other data sources and suppliers
- ✓ Mapped roads carrying 75% of vehicle travel for the State of NSW
- ✓ Established a 'light' Star Rating method for network scanning for local government roads, and
- ✓ Mapped potential data providers and estimated cost savings and efficiency gains of using new data sources and extraction methods.

Advance of the 'state of the art'

This project has pioneered concepts and methods which have long-term implications for TfNSW, other Australian jurisdictions and internationally. The project comes at the beginning of the UN's Second Decade of Action for Road Safety, and coincided with the release of Australia's National Road Safety Strategy and NSW's Road Safety Plan. All three have road risk ratings as a critical component for road safety management and reporting.

Data is needed to drive these road risk ratings on an unprecedented scale. This project is the first of its kind globally, and its activities and outcomes are of interest to stakeholders across Australia and around the world.

The table below highlights progress made across a number of areas since the project commenced.

Situation pre-project (2020)	Situation upon project completion (2022)
No systematic method to convert or verify data from other (non-manual) sources.	The AiRAP accreditation framework provides a standardised process and validation for the conversion and validation of road infrastructure, speed and flow data.
Limited knowledge about possible data providers and sources.	Understanding of current and potential data providers and range of data sources available across Australia and internationally.
Approximately 13 attributes able to be automatically extracted from LiDAR data.	Automatic or accelerated data extraction for all AusRAP infrastructure attributes has been achieved. 34 are now accredited.
Specifications for manual coding processes only.	Data definitions for automatic coding of infrastructure data established and documented.
Precision data, such as LiDAR, is expensive. The cost of bespoke LiDAR surveys approx. $250 per km.	Pre-collected LiDAR data from existing sources proven to meet road data needs at 1/10th of price (approx. $27 per km).
Costs of automatic data extraction unknown.	Costs and potential efficiency savings known.

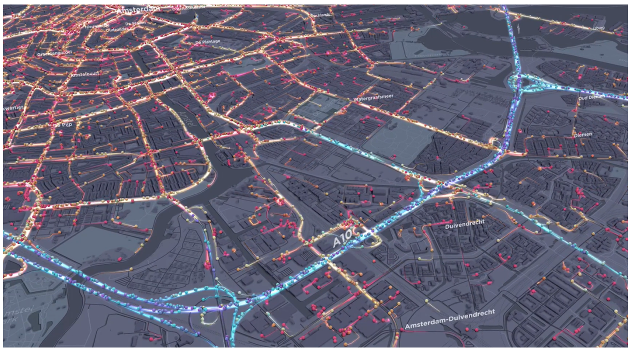
Data sources

The project used pre-captured data supplied by TomTom. Two types of data were considered:

1. Navigation MN-R data which, for the attribute data to be found viable, was used to produce data for the State road network (~20,000km) processed via an FME workbench conversion tool, and
2. Mobile Mapping (MoMa) data, which is both LiDAR and 360-degree imagery, to explore the viability of detecting road attributes using Machine Learning (ML) and other accelerated methods, for potential future use.

TomTom's data coverage extends to 124,000km of roads in NSW in their Net Class 2 1-4 range. This covers the State and regional road networks, as well as a proportion of the local road network.

TomTom MN-R Probe Data Visualisation and Mobile Mapping (MoMa) vehicles



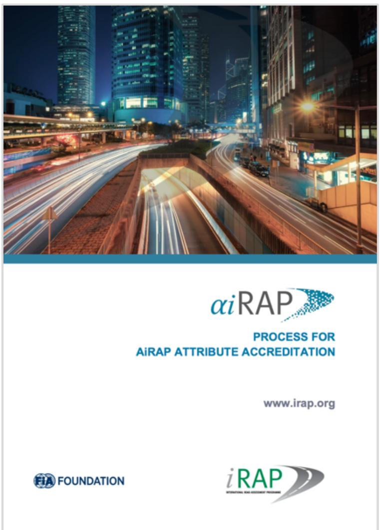
[TomTom MN-R Probe Data Visualisation]

[TomTom Mobile Mapping (MoMa) vehicles]

The AiRAP Framework

A vital outcome of the project was the development and proving of the AiRAP Framework. The framework:

- Provides the framework for converting new and emerging data sources into iRAP-compliant data for Star Rating assessments and road infrastructure safety KPI reporting in Australia and worldwide.
- Is supplier independent and supports the accreditation of multiple data providers and analytical approaches to generate data that meets the iRAP global standard.
- Has a flexible process which accommodates data derived from different types of source data, as well as different collection and processing methods.



The AiRAP accreditation process provides end-users with the confidence that the source data and analytical approach for generating attributes in accordance with the iRAP global standard is viable and repeatable. It also provides an understanding of the reliability of the data for different geographic regions, area types and road types, as well as when and how the data should be used.

The AiRAP accreditation process is now open to any data providers and is managed as part of iRAP's broader [accreditation program].

AiRAP-accredited road attribute data

The two types of source data used in the project (MN-R and MoMa) were demonstrated to be capable of capturing most AusRAP data attributes using either fully automated or accelerated (partially automated) methods.

A total of seven attributes from TomTom's MN-R data and a further 34 attributes extracted by Anditi from TomTom's MoMa data were accredited. This means that the feature extraction techniques used to produce data for these road attributes has been confirmed to meet iRAP's global standard and can be applied anywhere the source data is available.

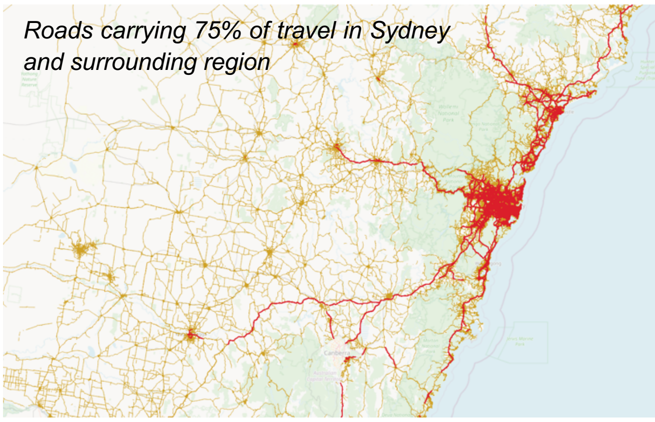
See **Table 8 Attribute coverage resulting from the project** for the full list.

75% of travel mapping

The UN Global Road Safety Target 4 is for more than 75% of travel to be on the equivalent of 3-star or better roads by 2030. Using TomTom Traffic Stats data, roads carrying 75% of vehicular traffic were mapped for NSW.

The analysis estimates that 7.3% of the road network (9,035km) of the total NSW road network carries 75% of vehicular travel.

Recently announced National and NSW State targets are for 80 percent of travel roads to meet road safety targets. The approach used can be adjusted accordingly.

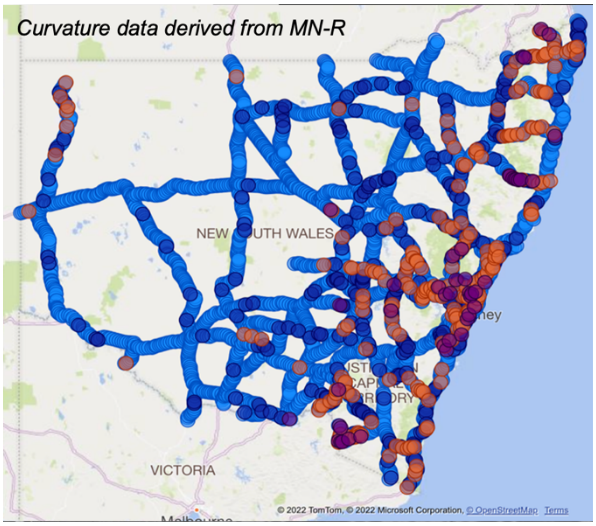


Roads carrying 75% of travel in Sydney and surrounding region

Data produced

The project delivered 20,000 kilometres of road attribute data (7 attributes) for the state road network in New South Wales using TomTom's MN-R data. This data covered:

- Speed limit
- Differential speed
- Number of lanes
- Curvature
- Grade
- Skid resistance (sealed and unsealed only)
- School zone warning (excludes flashing beacons)



Curvature data derived from MN-R

Light Star Rating methodology

In the course of the project, iRAP developed a light Star Rating methodology. This enables statistically valid light Star Ratings to be produced using a reduced set of attributes.

The light data methodology uses a minimum number of road features to predict remaining attribute values to calculate 'light' Star Ratings and FSI estimates. Its application is intended to reduce costs associated for large-scale Star Ratings for 'network-scanning' purposes, or enable adaptation of asset data and AiRAP data for producing Star Rating estimates prior to more detailed assessments of high-priority road lengths.

Recommendations for further research and development

This project was designed as an initial phase of research and development to test the principles and methods for automated extraction of road feature data from existing data sources, and gain a better understanding of the costs and potential of this data. As such, recommendations for second phase R&D project are proposed as follows:

- Continue refinement of AiRAP ready attribute and category definitions and extraction techniques
- Further develop capability for speed and flow attribute data
- Use a data-driven method to model vulnerable road user flows
- Explore ML methods for light data collection for LGs and pilot light Star Ratings
- Expand data supplier base and data marketplace

Contents

Figures

Tables

Acronyms

4WD	Four-wheeled drive vehicle
AADT	Annual Average Daily Traffic
ADAS	Advanced driver assistance systems
AI	Artificial intelligence
AiRAP	iRAP initiative for the 'accelerated and intelligent' collection of RAP attributes
Anditi	Anditi (company)
APAC	Asia Pacific
AusRAP	The Australian Road Assessment Programme
CAD	Computer aided design (software)
CO2	Carbon dioxide
Compass IOT	Compass IOT (company)
CRC	Cooperative Research Centre
CSV	Comma separated values (data file format)
DEM	Digital elevation model
DL	Digital learning
DS/PS	Driver-side/Passenger-side

EPSG	Geodetic Parameter Dataset (spatial referencing)
ETL	Extraction, Transfer and Loading
FMlab	Future Mobility Lab (UTS)
FSI	Fatality and serious injury
GDA94	Geocentric Datum of Australia (1994)
GIS	Geographic information system
GPS	Global positioning system
HD	High definition
HERE	HERE Technologies (company)
iMOVE	Australian national centre for transport and mobility R&D
IMU	Inertial measurement unit
iRAP	The International Road Assessment Programme (global charity and umbrella organisation for AusRAP)
KPI	Key Performance Indicator
LiDAR	Light detection and ranging (remote sensing method)
ML	Machine learning
MN-R	TomTom's MultiNet R
MNR2IRAP	Name of the FME workbench to convert MN-R data into iRAP attributes
MoMa	TomTom's Mobile Mapping
MRWA	Main Roads Western Australia
NC21-4	Net Class 2 roads categories 1-4
NGO	Non-government organisation
NSW	New South Wales
PSG	Geodetic Parameter Dataset (spatial referencing)
QA	Quality Assurance
QGIS	Open-source GIS software
RISM	The European Parliament's *Road Infrastructure Safety Management* directive
SHAP	Shapley Additive Explanations
SR4S	Star Rating for Schools
SRIP	Safer Road Investment Plan
SRRA	Safer Roads Risk Assessment (TfNSW application)
TfNSW	Transport for New South Wales
TomTom	TomTom International BV. (company)
UK	United Kingdom
UN	United Nations
USA	United States of America
UTS	University of Technology Sydney
ViDA	iRAP's online data processing platform used to produce Star Ratings, FSI mapping and SRIPs
VKT	Vehicle kilometres travelled
WA	Western Australia
WGS84	World Geodetic System (1984)
WHS	Workplace health and safety
WSP	WSP Global Inc. (company)

All $ are Australian dollars (AUD).

1 INTRODUCTION

Improving road safety performance is a priority for all levels of government in Australia. As part of the *National Road Safety Strategy 2021-2030*, Australian Transport Ministers set AusRAP Star Rating policy targets and risk assessment requirements for all states and territories. The Star Rating targets also align with the Global Road Safety Performance Targets agreed by the United Nations Member States that form part of the *Global Plan for the Decade of Action for Road Safety 2021-2030*.

As the umbrella charity that supports Road Assessment Programmes worldwide, iRAP has the vision for a world free of high-risk roads. iRAP oversees the global tools and specifications and supports partners in over 100 countries in addressing infrastructure safety and speed management. Within Australia the AusRAP programme is led by Austroads in partnership with all the key road agencies, research agencies, mobility clubs, industry bodies and relevant road safety NGOs.

AusRAP data is currently used to inform policy, investment and design decisions across the country. More than 300,000km of AusRAP Star Rating assessments have already been undertaken across the country that includes over 50 attributes coded to the iRAP global standard every 100 metres. These assessments have traditionally been collected using video data and manual coding of attributes.

The accelerated and intelligent collection and coding of this data to a common global standard, being developed as part of the global 'AiRAP' initiative, has the potential to reduce the time and effort required to undertake road safety assessments, reduce the costs and improve the accuracy and frequency of data collection for road safety performance tracking and key performance indicator (KPI) monitoring purposes. Together with road agency and industry expertise, the subsequent AusRAP Star Rating and Safer Road Investment Plan analysis has significant potential to improve road safety and save lives.

AiRAP seeks to use existing stores of readily available commercial and open-source data (including LiDAR, video and satellite data) and automated data analytics techniques including machine learning and artificial intelligence to deliver 'accelerated and intelligent' RAP data for Australia and the world to use.

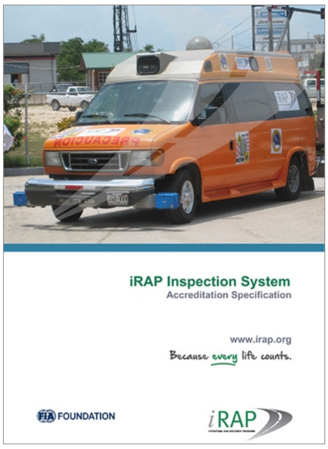
This initiative provides a unique and significant opportunity for the new Australian Office of Road Safety and national, state, and local governments to more cost effectively manage their safety performance and support achievement of the National Strategy and Action Plan Targets. The ability to performance track KPIs over time (including retrospective analysis) in a repeatable and standardized manner to a global standard will be highly valuable to all authorities.

1.1 Project context

Road safety assessment methodologies and models provide an objective and predictive way to assess a road's safety performance based on the physical characteristics of the road environment, as well as road user flows and speeds.

The iRAP Star Rating and associated estimation of fatal and serious injury (FSI) crashes address all road user types and have been adopted by UN Member States at the global level through the specification of 3-star or better performance targets for all road users (1-star is the least safe and 5-star the safest). The associated investment planning and related iRAP tools that seek to measure safety performance and maximise lives saved per dollar spent are well established and provided in a free-to-air environment.

Historically, Star Rating assessments have been undertaken manually using human coding teams and video and street level images in accordance with iRAP specifications. The specifications have been adapted over time to address all of the typical road features encountered across a road network and to inform the more detailed fatality estimations and countermeasure selection and optimisation needs associated with Safer Road Investment Plans.



Specialised data collection vehicles and coding software (collectively known as Inspection Systems) have been progressively improved over time to increase the efficiency and quality of manual data collection methods. iRAP has an Inspection System Accreditation which certifies these systems for use in Star Rating assessments. More than 15 systems are accredited for global use.

Anditi's Roadviewer Inspection System used for this project is the first in the world to be accredited to use mobile LiDAR and imagery combined, opening the pathway to the generation and utilisation of road attributes extracted using accelerated and intelligent techniques as part of the global roll out of the AiRAP initiative.

iRAP has been actively engaging with big-data partners over the last 5-10 years to unlock the potential of alternate data sources and analytical techniques to generate road safety data. With no single system or source dataset able to generate all of the required attributes (in terms of technical and geographical coverage) a more open market approach to generating data is required with the "AiRAP" initiative established to meet that need.

This approach recognises the ever-expanding sources of road data being generated across multiple sectors, such as mapping, vehicle sensors, satellite and telematics, that coupled with the global iRAP specifications and standard, can unlock the potential of more frequent, efficient, cost-effective and accurate road safety data worldwide.

1.1.1 Global needs and demand

In 2021, the United Nations launched its *Global Plan for the Decade of Action for Road Safety 2021-2030*, with the overarching aim to reduce road deaths and serious injuries by 50 percent by 2030. The Plan outlines the *what*, *how* and *who* to achieve the 12 Global Road Safety Performance Targets agreed by Member States. It provides the basis for governments and partners to implement the Safe Systems Approach in the creation and implementation of strategies and programmes for road safety, sustainable mobility and urban design.

Amongst other things, the Plan explicitly recommends that governments, "Specify a technical standard and star rating target for all designs linked to each road user, and the desired safety performance standard at that location." Target 4 is that roads carrying 75% of travel the equivalent of a 3-star or better standard by 2030.

As at the end of 2021, approximately 3 million kilometres of roads have been Risk Mapped or Star Rated by iRAP partners in more than 100 countries worldwide.[7] The Global Plan, as well as regional, national and sub-national policy directives, such as the European Union's RISM Directive and Australia's new National Road Safety Strategy are driving increased demand for the measurement and monitoring of safety performance across road networks. With more than 150 million kilometres of road worldwide, more efficient and scalable methods to collect iRAP data are a priority for many governments and road safety stakeholders.

Within Australia, the Australian government has adopted a target for 80% of travel on 3-Star or better roads by 2030 as part of its *National Road Safety Strategy 2021-2030*. Australia has the world's ninth longest road network, measuring over 823,000 km in length. It comprises 356,000 km of paved roads and over 466,000 km of unpaved roads. The targeting of priority road sections and the completion of regular AusRAP assessments

[7] Non-unique road carriageway kilometres. This means that reassessed roads or multiple carriageway roads may be represented more than once.

and performance tracking of road safety performance is therefore needed to drive cost-effective investment to reduce road trauma and celebrate success.

The AiRAP methods and mechanisms being developed as part of this project, including the AiRAP Framework, data extraction techniques and percent of travel mapping, as well as improved understanding around cost and efficiency, will directly benefit TfNSW, as well as other State, Territory and local governments across Australia. In addition, the outcomes of the project will inform world-leading application of the AiRAP approach that will meet the extended needs of road managers globally.

Figure 1 Data in iRAP's Star Rating processing platform, ViDA (km from 2014 – 2022)

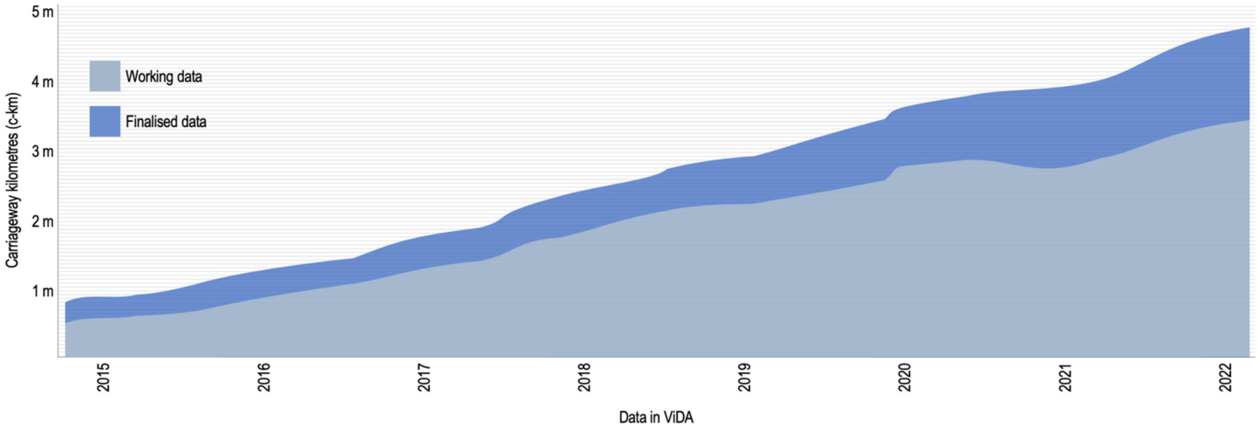
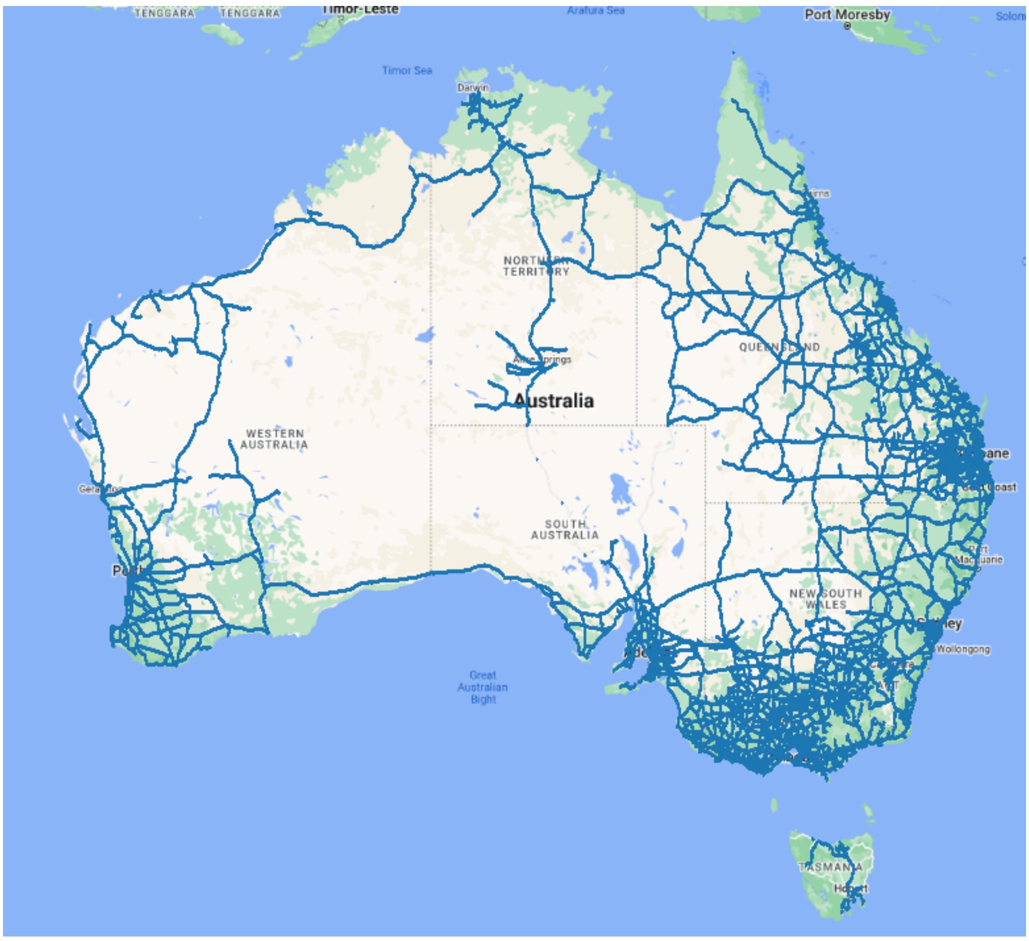


Figure 2 Unique kilometres of Star Rating related data in Australia (2014-2022)



1.2 Project objectives

The project aimed to prove rapid, scalable, and repeatable methods for road data extraction as part of iRAP's global 'AiRAP' initiative. The initiative aims to ultimately open existing and emerging data sources for network-level road safety assessments throughout Australia and worldwide.

At the outset of the project, the objectives were to:

6. Develop the AiRAP Framework
7. Prove AiRAP Framework and scalability
8. Prove independent feature extraction
9. Create an AiRAP supplier map and estimate efficiency savings
10. Recommend Phase 2 research priorities

As an iMOVE project, the focus has been on a collaborative research and development to improve transport and mobility through the use of new methods of road safety data collection and processing. Priority application of these new methods for TfNSW included (i) finding cost effective and viable methods to support road safety assessments for local governments and (ii) improving vehicle flow (AADT) data for NSW.

1.3 Project partner roles and approaches

The project partners are Transport for NSW (TfNSW – Lead partner), the University of Technology Sydney (UTS) and the International Road Assessment Programme (iRAP).

TomTom (via in-kind contribution) and Anditi (via a supplier agreement with iRAP) provided the data for the project.

The Project Lead was Monica Olyslagers (iRAP), who was responsible for coordinating the partner inputs to deliver the project. The Project Management Committee included Joseph Le (TfNSW), Rob McInerney (iRAP), Monica Olyslagers (iRAP) and Arno Schaaf (UTS).

1.3.1 Transport for New South Wales

Transport for New South Wales (TfNSW) was the Lead Project Partner and the end user of the data to be generated as part of the project. New sources of AusRAP data made possible by this project will form an immediate input into ongoing AusRAP activities and road asset management within the State of New South Wales.

TfNSW, through the AusRAP Technical Working Group, supported the sharing of knowledge and potential application nationwide in support of the Australian National Road Safety Strategy and associated State and Territory goals to reduce road trauma across the country.

1.3.2 International Road Assessment Programme

The International Road Assessment Programme (iRAP) developed the AiRAP Framework that provides the platform and specifications for accelerated and intelligent data collection for road infrastructure safety purposes in Australia and worldwide.

The Framework aligns with the application of the iRAP global standard in a harmonized manner for the benefit of Australian jurisdictions and other transport stakeholders that is aligned with similar global initiatives for the mutual benefit of all iRAP partners worldwide. Using the new AiRAP Framework, iRAP also undertook the accreditation activities for the initial AiRAP attributes generated as part of the project.

Anditi and TomTom

Anditi and TomTom were the data providers for the project.

Anditi licensed TomTom's Multinet-R data to Transport for NSW under their pre-existing value-added reseller agreement. Through that same value-added reseller agreement TomTom provided Anditi with access to their Mobile Mapping (MoMa) source data for a section of road in Western Sydney and a subset of iRAP related attributes from their Multinet-R data for 20,000 kms of road for the project.

In-kind provisions by TomTom included the further development and the execution of TomTom's prototype tool to transform data from TomTom's Multinet-R format to the iRAP specifications. The tools were used to generate AiRAP compliant data at scale across the state road network of NSW as depicted here: https://www.rms.nsw.gov.au/business-industry/partners- suppliers/lgr/documents/map-regions.pdf.

Anditi also led the use of artificial intelligence and other accelerated technologies to utilise the MoMa source data that generated additional AiRAP compliant attributes.

1.3.3 University of Technology Sydney (UTS)

UTS reviewed the AiRAP accredited data and assessed scalability in support of AusRAP activity in Australia. The feature extraction components were also assessed for scalability across NSW.

Based on the available AiRAP datasets, attributes and coverage, UTS have also scoped a detailed research proposal using the data and extended datasets in the TfNSW consolidated data initiative (e.g. crash data, asset data) to support AusRAP data collection and investment decision making that is viable for local government road networks. The proposal will help shape further research priorities as part of any extension of the AiRAP iMOVE partnership.

2 THE AiRAP FRAMEWORK

The AiRAP Framework, developed under this project, is the mechanism by which new and emerging data sources can be converted into iRAP-compliant data for application in Star Rating assessments and road infrastructure safety KPI reporting in Australia and worldwide.

The accreditation process provides end-users with the confidence that the source data and analytical approach for generating attributes in accordance with the iRAP global standard is viable and repeatable. It also provides an understanding of the reliability of the data for different geographic regions, area types and road types, as well as when and how the data should be used.

The process is designed to be flexible and accommodates data derived from different types of source data, as well as different collection and processing methods that will support an open and competitive data marketplace.

2.1 Processes and procedures for AiRAP accreditation

A vital outcome of the project was the development and proving of the AiRAP Framework. The framework:

- Provides the framework for converting new and emerging data sources (e.g. big data, satellite, telematics, LIDAR, feature recognition cameras etc.) into iRAP-compliant data for Star Rating assessments and road infrastructure safety KPI reporting in Australia and worldwide.
- Is supplier independent and supports the accreditation of multiple data providers and analytical approaches to generate data that meets the iRAP global standard.
- Has a flexible process which accommodates data derived from different types of source data, as well as different collection and processing methods.

2.1.1 AiRAP accreditation process

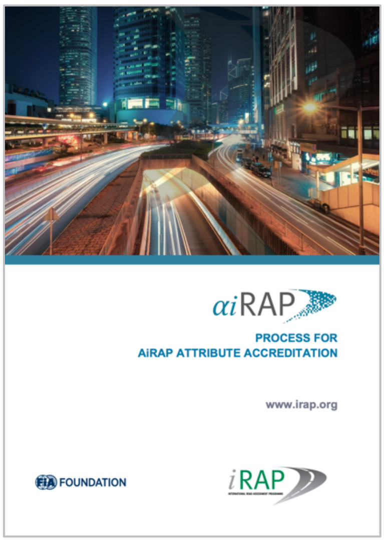
The aim of the accreditation is to verify that the algorithms being used meet iRAP coding specifications consistently, accurately and in a repeatable way. The accreditation focuses on the verification of the conversion process rather than the output data that should always be subject to relevant quality assurance.

The AiRAP accreditation process provides end-users with the confidence that the source data and analytical approach for generating attributes in accordance with the iRAP global standard is viable and repeatable. It also provides an understanding of the reliability of the data for different geographic regions, area types and road types, as well as when and how the data should be used.

The AiRAP accreditation process is now open to any data providers and is managed as part of iRAP's broader accreditation program.

The accreditation process is described in the *Process for AiRAP Attribute Accreditation* document. Specifically, the document:

- Defines what AiRAP is, how the AiRAP attribute accreditation process will work and what is required of suppliers
- Guides prospective suppliers through the process required to attain AiRAP attribute accreditation, and
- Sets out the relationship with iRAP for the AiRAP accreditation process and support, the ownership and sharing of intellectual property, and the role iRAP will play in promoting AiRAP attributes and connecting suppliers to consumers.



The accreditation process can be summarised as:

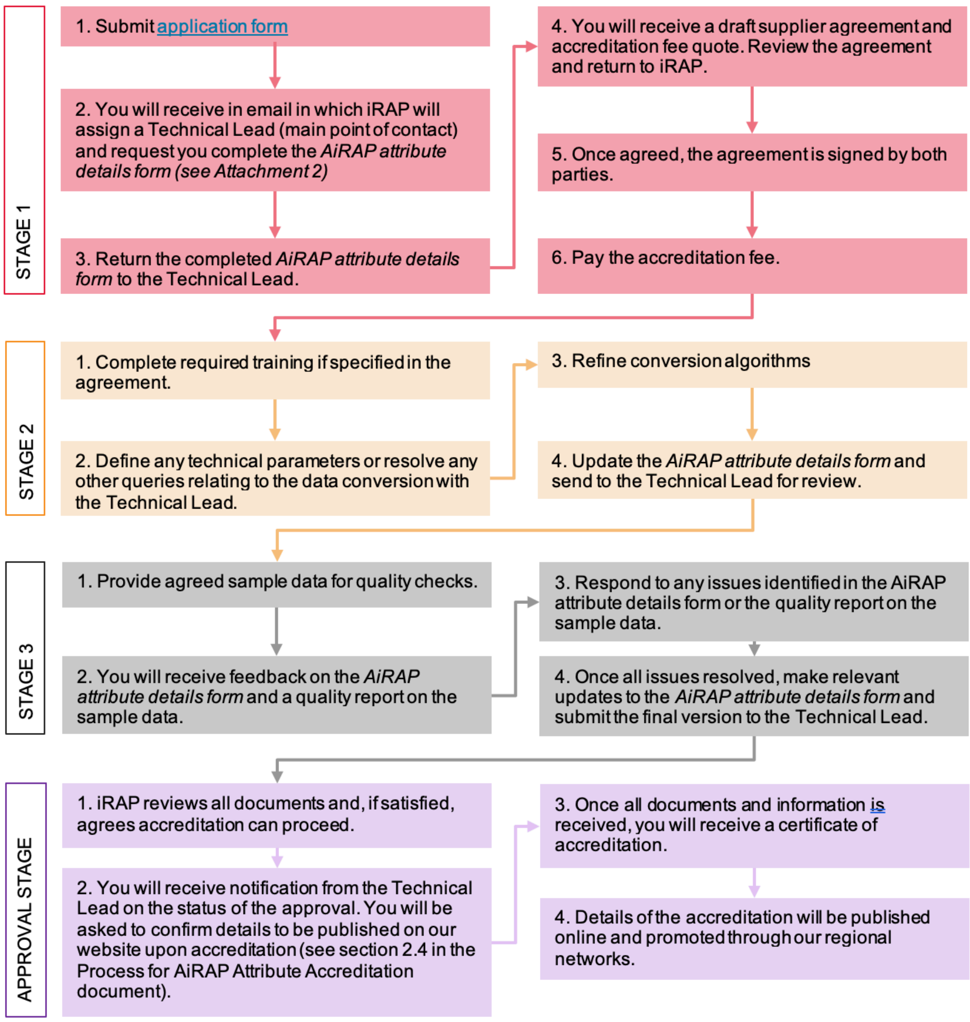
i. Initial application where the applicant submits basic information which is reviewed and refined. A fee is paid by the supplier to ensure the process is self-sustaining and can scale with demand and iRAP assigns a technical lead to oversee the process.

ii. Details on the source data and conversion process are submitted by the applicant and are then reviewed by iRAP. Any issues are identified and corrected by the applicant.

iii. Once the conversion process is verified, the applicant provides sample data to prove it works. This data is then reviewed by iRAP. Any issues identified are corrected and the validation process repeated.

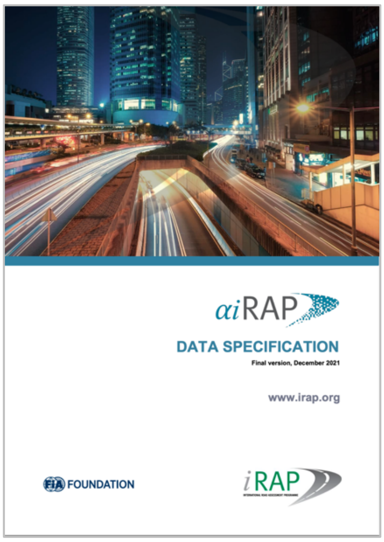
iv. Once the conversion process is validated, conditions for accreditation[8] are determined and the accreditation is granted. Details of the accredited attribute, the supplier and other relevant details are then published on the iRAP website.

Figure 3 AiRAP accreditation process map



[8] Once attribute accreditation is granted, conditions may be placed on its application. For example, its use may be limited by road class/type or by geographic location. Conditions will be determined based on the extent to which the data is available, can be relied on for quality, can be converted accurately and can produce consistent results.

2.2 Digital data specifications



To support the AiRAP Framework, iRAP developed an *AiRAP Data Specification*. Clear data definitions to underpin the ML algorithms and accelerated techniques used for the automated extraction of road feature data are critical for the reliability and repeatability of such processing.

The purpose of this document is to provide potential data suppliers with specific guidance and requirements to support the conversion of source data into AiRAP accredited data.

This document provides:

- Background and contextual information about the Star Rating models and how data is used.
- A full list of attributes, including priority attributes (those that play a significant role in the model and for which specific road user).
- Data segmentation and formatting.

The document will also be progressively updated to ensure it reflects a consistent set of agreed parameters required for converting data from a variety of data sources.

How many attributes are used for Star Ratings?

There are a total of 78 attributes recorded as part of a standard Star Rating. These include:

- Road details and contextual information (e.g. road name, location, etc.)
- Volume and flow data (who is using the road), including observed flows and road user flow data
- Vehicle speed information, including speed limits and operating speeds
- The physical features of the road, including mid-block attributes, roadside features, intersection features and vulnerable road user (VRU) facilities, and
- Other information.

However, not all attributes have a direct influence on Star Ratings. Some provide referential information for the data collected. Others provide details which assist with analysis, quality checks, fatality and serious injury (FSI) estimates, or maybe useful for modelling data where it is not available (such as pedestrian flows).

The table below provides a breakdown of these attributes by function.

Road details and context (reference information required for all attribute data)

1. Coder name
2. Coding date
3. Road survey date
4. Image reference
5. Road name
6. Road section
7. Distance
8. Length
9. Latitude
10. Longitude
11. Landmark (optional)
12. Comments (optional)
13. Area type

Direct impact on Star Rating

1. Vehicle flow (AADT)
2. Operating speed – 85th percentile
3. Speed limit
4. Median type
5. Delineation
6. Number of lanes
7. Lane width
8. Curvature
9. Speed differential
10. Skid resistance
11. Road condition
12. Roadside severity – driver-side distance
13. Roadside severity – passenger-side distance
14. Roadside severity – driver-side object
15. Roadside severity – passenger-side object
16. Paved shoulder – driver-side
17. Paved shoulder – passenger-side
18. Service Road
19. Lane width
20. Quality of curve
21. Vehicle parking
22. Grade
23. Sight distance
24. Centre line rumble strips
25. Shoulder rumble strips
26. Intersection type
27. Intersection quality
28. Intersection channelization
29. Intersecting road volume
30. Property access points
31. Sidewalk – driver-side
32. Sidewalk – passenger-side
33. Pedestrian crossing facilities – inspected road
34. Pedestrian crossing facilities quality
35. Pedestrian crossing facilities – side road
36. Pedestrian fencing
37. Speed management
38. Street lighting
39. Motorcycle facilities
40. Bicycle facilities

No impact on Star Rating

1. Motorcycle observed flow
2. Bicycle observed flow
3. Ped observed flow across
4. Ped observed flow along – driver-side
5. Ped observed flow along – passenger-side
6. Motorcycle speed limit
7. Truck speed limit
8. Upgrade cost
9. Land use – driver-side
10. Land use – passenger side
11. Motorcycle % (influences motorcyclist FSI Estimation)
12. Operating speed – mean (used for FSI Estimations for all road users)
13. Roads that cars can read
14. Vehicle Occupant Star Rating policy target
15. Motorcyclist Star Rating policy target
16. Pedestrian Star Rating policy target
17. Bicyclist Star Rating policy target
18. Annual fatality growth multiplier

Limited impact on Star Rating

1. Carriageway label	Affects smoothing of results
2. Pedestrian peak hour flow along – passenger side 3. Pedestrian peak hour crossing flow 4. Pedestrian peak hour flow along – driver side 5. Bicycle peak hour flow	Switches Star Ratings 'on' for these users if present. Affects FSI Estimates for these road users.
6. Roadworks	Switches Star Ratings 'off' where major road works are present.
7. School zone warning 8. School zone crossing supervisor	Slightly reduces pedestrian risk factors if school and signs/beacons present.

2.3 AiRAP Accreditation Levels

Based on the quality assurance checks and evaluation undertaken as part of the accreditation process, each attribute is provided an accreditation level. The levels range from 1-3 as shown below. The level refers to the degree to which the attribute categories can be captured and produced in accordance with the iRAP specification. It does not refer to the quality or reliability of the data, and assumes that recommended post-processing quality assurance checks are completed as described in the accreditation reports.

Table 1 AiRAP accreditation levels

Level 3	There are no limitations on the use of the data for iRAP applications. It is suitable for network screening, light data approaches, KPI reporting, Star Ratings, Fatality Estimations and Investment Plans.
Level 2	Due to limitations, the data should be used with caution (or in conjunction with additional data) for iRAP applications. It is suitable for network screening, light data approaches and KPI reporting, and may be suitable for Star Ratings and Fatality Estimations.
Level 1	Due to limitations, the data is not suitable for all iRAP applications. It is suitable for network screening, light data approaches and KPI reporting.

In most cases, the accreditation level will correspond with whether an attribute is Full (F), Partial (P) or Simple (S), where:

- Full (F) indicates that all attribute categories are available
- Partial (P) indicates that only some attribute categories are available, and
- Simple (S) indicates that a simplified version of the attribute categories are available.

A decision to use partial (P) attributes for standard iRAP applications needs to be made on the basis of the likely effect it will have on the assessment outcomes. For example, if a particular type of barrier cannot be detected, but that barrier type is not used in a particular country, then the data can be used with confidence that it is going to have minimal (if any) impact on the results.

Simple (S) attribute categories are different to standard ones, meaning they cannot be used directly in standard Star Rating assessments. For example, instead of determining what intersection type there is, the simple attribute just indicates if an intersection is present or not. However, simple attributes can be used for generating light star ratings or for KPI reporting.

The role of the accreditation levels is to ensure that the end data user (e.g. road authority) is provided information about the viability and limitations of the data and is able to be make informed decisions about the data's use based on their needs and requirements. How 'fit-for-purpose' the data is for any specific application is ultimately the responsibility of the agency procuring the data.

3 DATA AND ANALYSIS

To date, collecting road attribute data for safety assessment and asset management has predominantly relied on the collection of video survey data which can be used to manually code the information. A combination of LiDAR and 360-degree imagery offers a much more enriched and accurate data source compared to video. This expands the potential application as it provides the ability to take precise measurements and be used to inform road design. However, bespoke LiDAR capture has in the past been prohibitively expensive for large scale use across networks.

The project sought to explore the viability of using existing sources of pre-collected data as a more efficient and cost-effective alternative to bespoke data collection activities.

To do this, this project considered two types of pre-collected data sources which could potentially meet the needs of TfNSW. Both were provided via in-kind contribution from TomTom Global.

1. Navigation MultiNet-R (MN-R) data which, for the attribute data to be found viable, was used to produce data for the State road network (~20,000km) processed via an FME workbench conversion tool, and
2. Mobile Mapping (MoMa) data, which is both LiDAR and 360-degree imagery, to explore the viability of detecting road attributes using AI and accelerated and intelligent technologies, for potential future use.

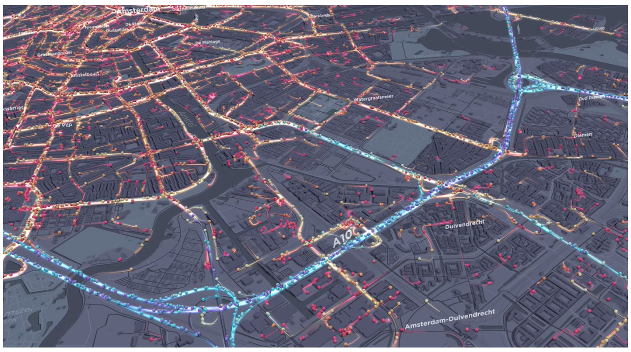
TomTom's MN-R data coverage extends to 124,000km of roads in NSW in their Net Class 2 1-4 range. This covers the State and regional road networks, as well as a proportion of the local road network.

3.1 About TomTom source data

The MN-R data is collected via anonymous GPS measurements from its community of 600+ million drivers using their mapping products. This input is combined with knowledge harvested from local experts alongside a multi-source approach integrating data from survey vehicles, GPS traces, governmental sources, and vehicle sensor data to build the TomTom maps. MN-R data is stored on TomTom's Map Content Portal with archived versions updated weekly.

The MoMa data comes from imagery and point cloud sensor data from their global fleet of mobile mapping vehicles equipped with sensors and cameras. These historically capture the top 4 road classes with a refresh rate of <24 months in Australia and many other countries. Archive MoMa data is stored in TomTom's cloud storage facility. Data is typically pre-collected and extracted from TomTom's archives as needed. However, custom capture can be requested from clients if TomTom's recent archive does not cover their scope of work.

Figure 4 TomTom MN-R Probe Data Visualisation and Mobile Mapping (MoMa) vehicles



TomTom MN-R Probe Data Visualisation

TomTom Mobile Mapping (MoMa) vehicles

Table 2 Summary of TomTom's MN-R and MoMa data coverage and availability

	MN-R source data	TomTom MoMa (mobile LiDAR and 360-degree imagery)
Availability	164 countries and 35 territories, including Australia	60 countries, including Australia
Coverage	Approx. 68 million kms of navigable road All geographic regions and area types (i.e. rural/urban etc.) Focussed on higher class roads (NC2 1-4)	Approx. 10 million kms of road All geographic regions and area types (i.e. rural/urban etc.) Focussed on higher class roads (NC2 1-4)
Refresh rate	Captured continuously with weekly map updates	Updated at a rate of approximately 3-4 million kms per year

3.2 MN-R data conversion and results

This section summarises the methodology used for the MN-R source data conversion, production and accreditation. See Appendix A for more detail.

3.2.1 Identification of attributes

An initial analysis of the attributes available in the MN-R data was completed to determine which attributes were appropriate to be accredited by iRAP. This analysis considered the feature level and road class. Of the attributes analysed, Speed limit, Differential speed, Number of lanes, Curvature, Grade, Skid resistance and School zone warning proceeded to accreditation. Operating speed, Land use, and Intersection channelisation were eliminated due to lack of data.

3.2.2 Data produced for State road network

Using the FME workbench, the attribute data was then produced for the seven attributes deemed viable, for approximately 20,000 kilometres of the state road network in New South Wales.

3.2.3 Accredited attributes

A total of seven attributes were accredited from TomTom's MN-R data as shown in the table below. All were found to be able to produce the full range of attribute categories except for Skid resistance and School zone warning. Skid resistance was Level 1 accredited as it could only be used to distinguish if a road was sealed or unsealed, rather than the full five categories which takes into account the quality of the surface. School zone warning does not record flashing beacons separate to regular signage.

See AiRAP Accreditation Levels for more information.

Table 3 Accredited attributes from TomTom MN-R data (Accredited supplier: TomTom)

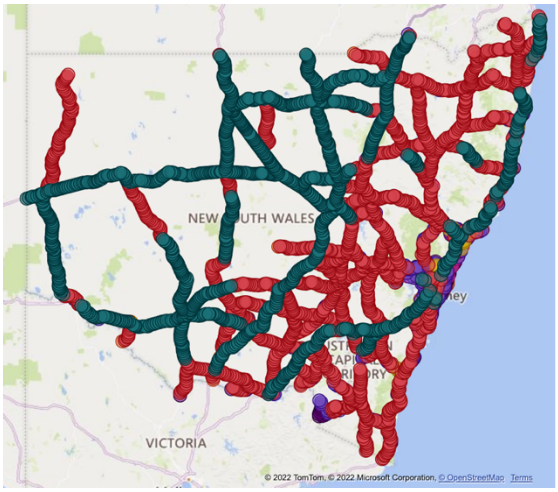
Attribute	**Full (F) Simple (S) Partial (P)**	**Area** (Urban, Rural, All)	**Road types** (Specific [Sp.]; All)	**Quality checked location/s**	**QA min. % accuracy** (<85%*; ≧85%; ≧90%; ≧95%)	**Accreditation Level**
Speed limit	F	All	Sp.	NSW, Australia		Level 3
Differential speed	F	All	Sp.	NSW, Australia	≧95%	Level 3

Attribute	Full (F) Simple (S) Partial (P)	Area (Urban, Rural, All)	Road types (Specific [Sp.]; All)	Quality checked location/s	QA min. % accuracy (<85%*; ≧85%; ≧90%; ≧95%)	Accreditation Level
Number of lanes	F	All	Sp.	NSW, Australia	≧95%	Level 3
Curvature	F	All	Sp.	NSW, Australia	≧95%	Level 3
Grade	F	All	Sp.	NSW, Australia	≧95%	Level 3
Skid resistance	S	All	Sp.	NSW, Australia	≧95%	Level 1
School zone warning	P	All	Sp.	NSW, Australia	≧95%	Level 2

3.2.4 MN-R derived attribute maps

Figure 5 Speed limits

All speed limit categories shown



Speed limits < 100km/h shown

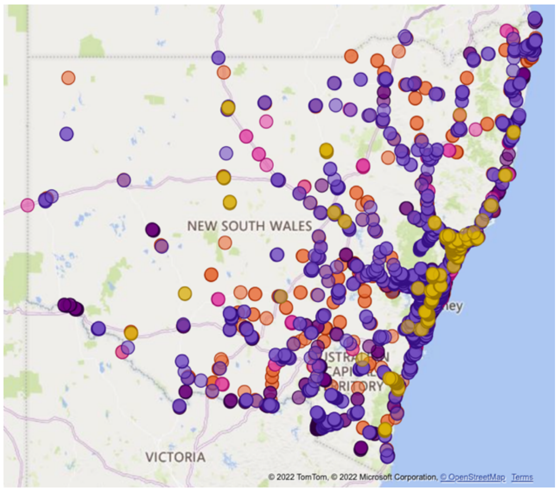


Figure 6 Differential speeds present

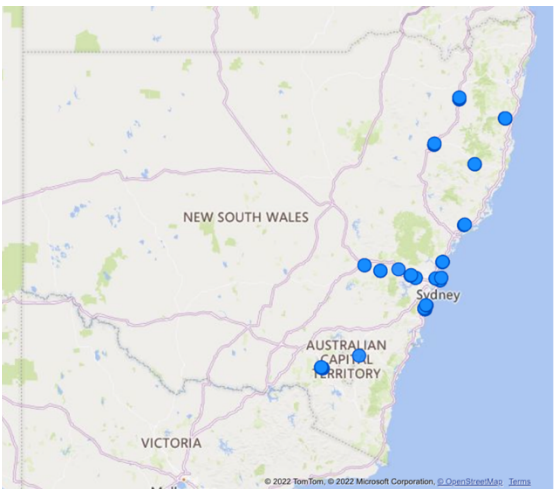
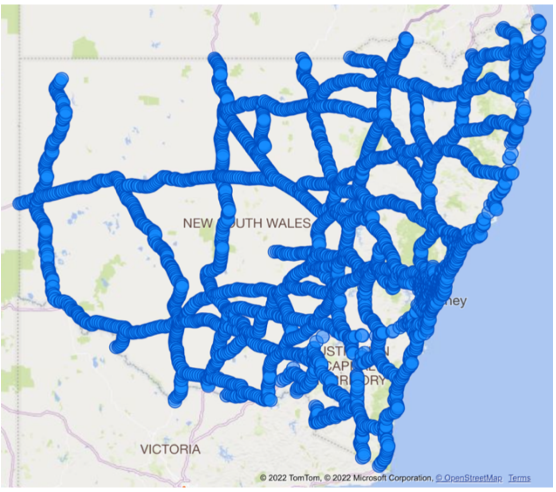
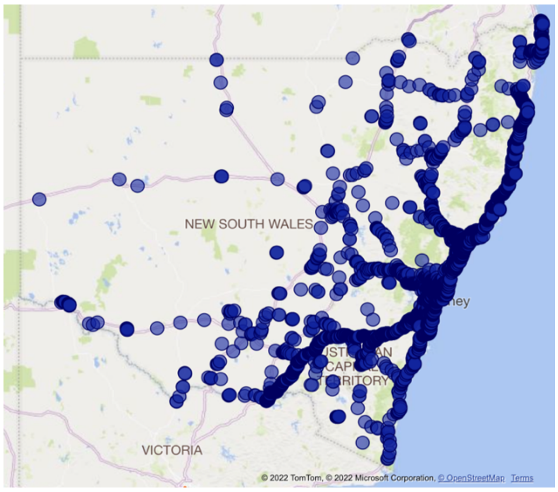


Figure 7 Number of lanes

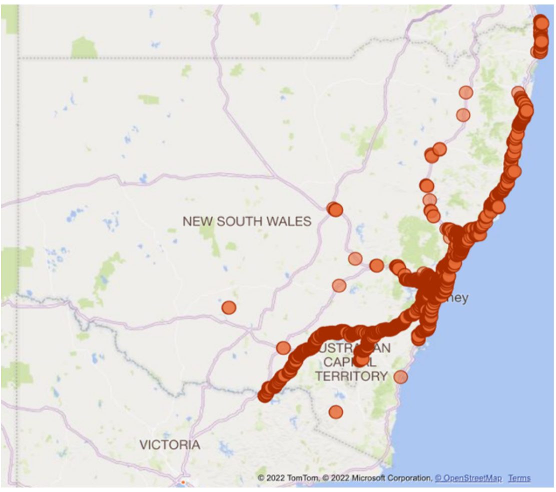
One lane each direction



Two lanes each direction



Three lanes each direction



Four or more lanes each direction

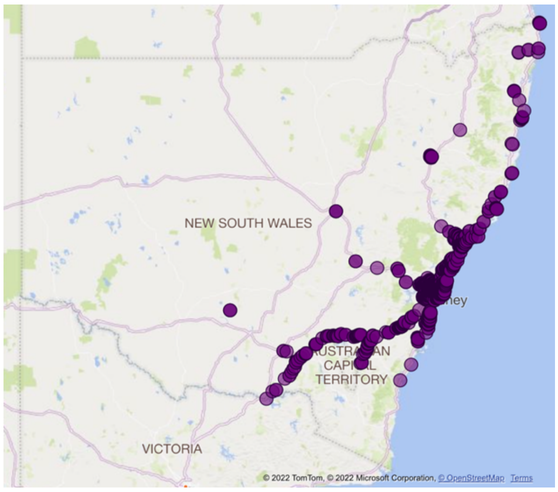
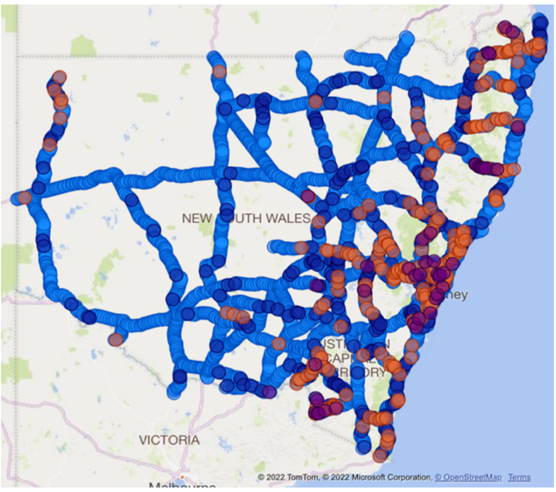
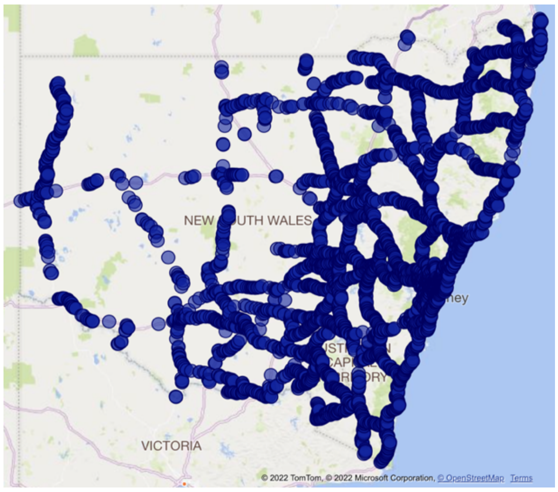


Figure 8 Curvature

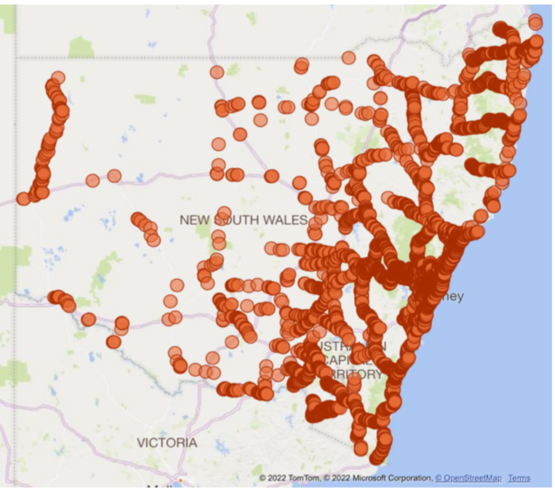
All categories



Moderate curve



Sharp curve



Very sharp curve

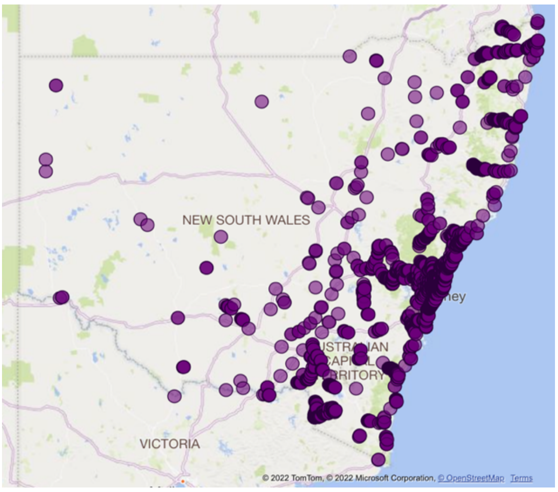
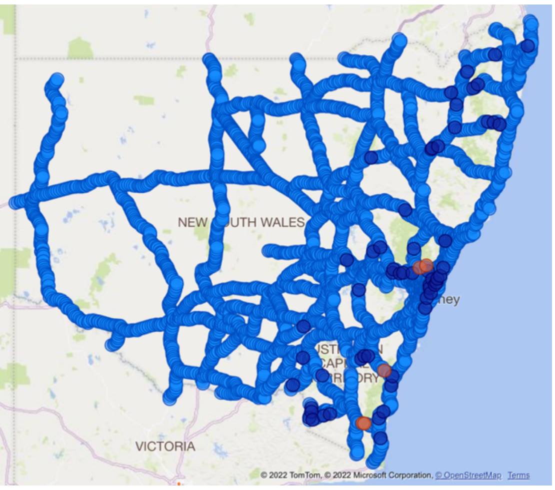


Figure 9 Grade

All categories



Moderately steep 7.5% to <10% (dark blue) and very steep >10% (orange) only

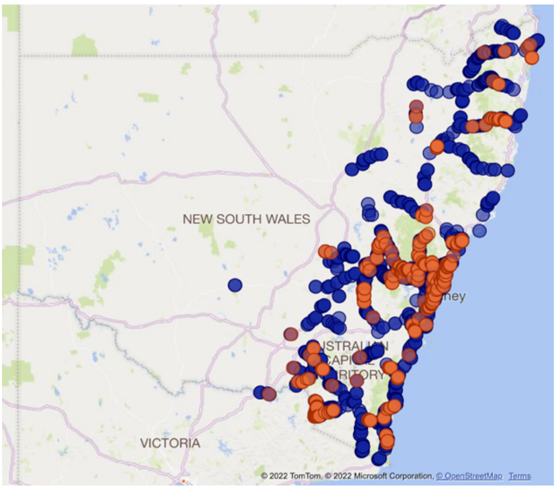
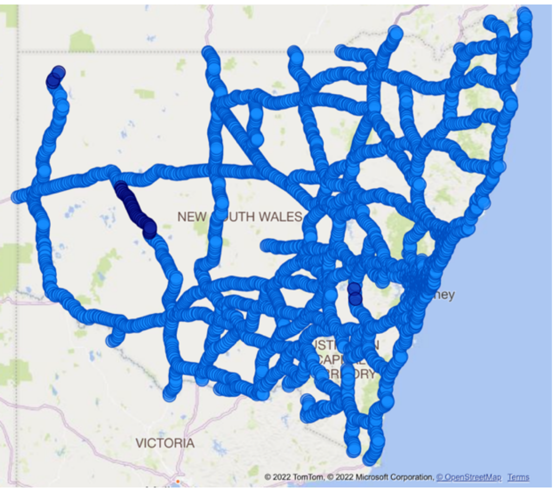


Figure 10 Skid resistance

Sealed (light blue) and unsealed (dark blue)



Unsealed only

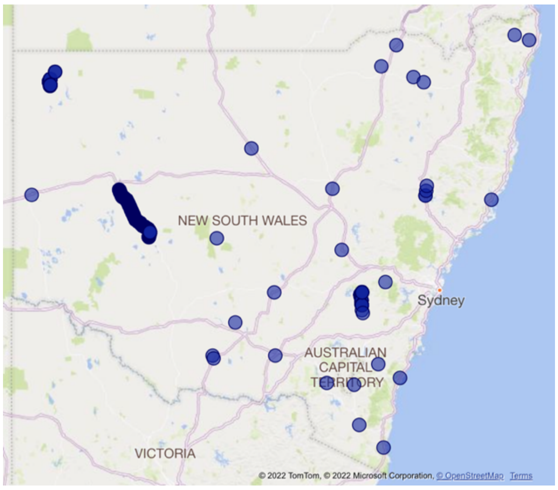
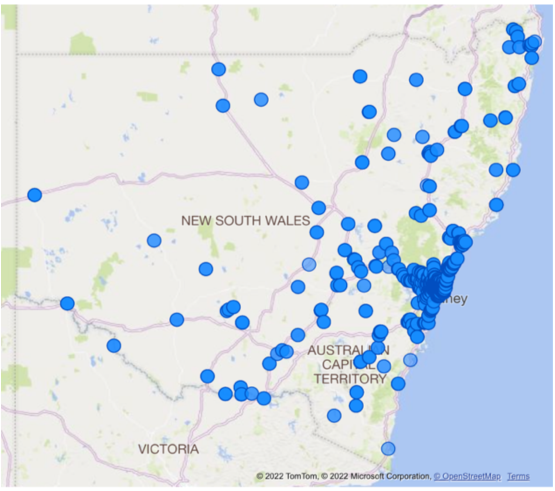


Figure 11 School zone warning

School zone warning present



3.3 MoMa data feature extraction

This section summarises the methodology used for feature extraction from MoMa data and accreditation. See Appendix A for more detail.

3.3.1 Identification of attributes

For the past four years, Anditi has been researching and developing automated and accelerated techniques to utilise mobile LiDAR and 360-degree imagery for Star Rating of roads.

Mobile LiDAR and 360-degree imagery is only suitable for the extraction of road infrastructure attributes. It is not suitable for others, such as traffic speed and road user flows.

3.3.2 Feature extraction methods

AI and other technologies used by Anditi in deriving attributes for Star Rating of roads include, LiDAR point cloud processing, image processing/computer vision algorithms, and fusion of information extracted from imagery and point cloud.

In undertaking this analysis, Anditi applies various techniques of Signal Processing, Optimization and Artificial Intelligence including:

i. Data driven modelling and learning

ii. Convolutional and Deep Neural Networks

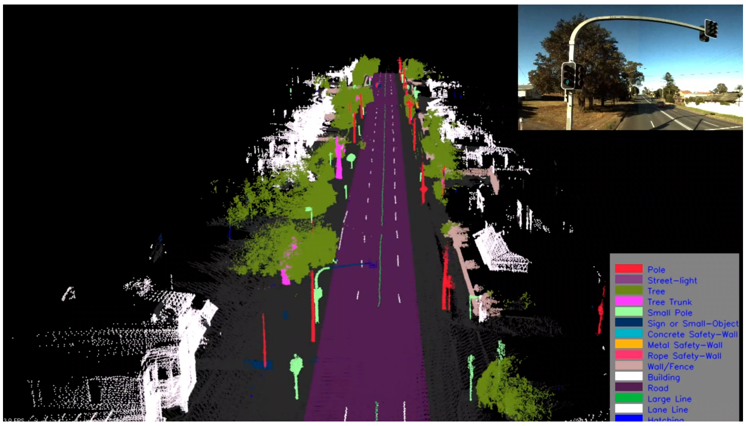
iii. Convex optimization

iv. Supervised and unsupervised clustering

The source data, after quality checking, is processed using Anditi's Roadviewer technology which converts it into iRAP attributes for subsequent viewing, checking and editing where required. Roadviewer is then used to export the iRAP coded data into a VIDA compatible format.

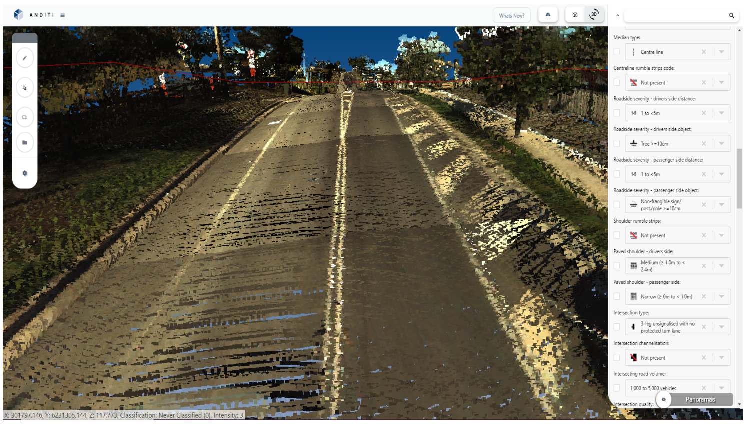
Figure 12 – Anditi Roadviewer Technology

3D Feature Extraction – Canterbury Road

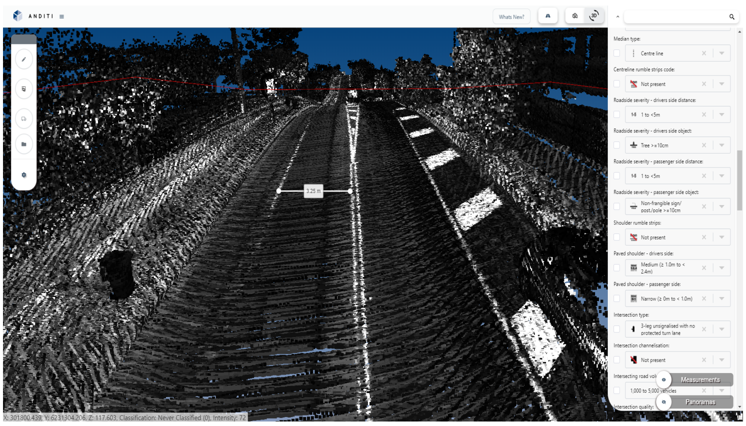


Roadviewer Inspection System - 360-degree imagery

Roadviewer Inspection System - Colourised Point Cloud



Roadviewer Inspection System - LiDAR Intensity of Return



3.3.3 Accredited attributes

A total of 52 attributes were able to be extracted by Anditi from TomTom's MoMa data with 34 of these being accredited under AiRAP.

Anditi's data conversion process includes both automated and 'accelerated' (partly automated) functions. The extensive quality assurance processes developed by Anditi are fully documented and are provided as part of the accreditation and coding system.

All 34 accredited attributes provide the full range of sub-attributes for urban and rural areas, and hence are accredited to 'Level 3'. See AiRAP Accreditation Levels for more information.

As the accreditation was based on TomTom MoMa data, it is accredited to N2C1-4 road types as defined by TomTom. However, as Anditi's conversion algorithms are also adaptable to alternative data providers, this may be revised in future if alternative data sources can provide the data for other roads at the same standard.

Table 4 Accredited attributes from TomTom MoMa data (Accredited supplier: Anditi)

Attribute	**Full (F) Simple (S) Partial (P)**	**Area** (Urban, Rural, All)	**Road types** (Specific [Sp.]; All)	**Quality checked location/s**	**QA min. % accuracy** (<85%*; ≧85%; ≧90%; ≧95%)	**Accreditation Level**
Median type	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Roadside severity – distance (driver side/passenger side)	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Roadside severity – object (driver side/passenger side)	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Paved shoulder (driver side/passenger side)	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Intersection type	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Number of lanes	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Lane width	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Speed limit	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Speed management	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Curvature	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Skid resistance	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Road condition	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3

Attribute	Full (F) Simple (S) Partial (P)	Area (Urban, Rural, All)	Road types (Specific [Sp.]; All)	Quality checked location/s	QA min. % accuracy (<85%*; ≧85%; ≧90%; ≧95%)	Accreditation Level
Vehicle parking	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Grade	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Sight distance	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Delineation	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Street lighting	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Centreline rumble strips	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Shoulder rumble strips	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Intersection channelization	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Property access points	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Land use (driver side/passenger side)	F	All	Sp.	NSW & WA, Australia	≧85%	Level 3
Area type	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Pedestrian crossing facilities (inspected road/side road)	F	All	Sp.	NSW & WA, Australia	≧90%	Level 3
Pedestrian fencing	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Sidewalk (driver side/passenger side)	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
School zone warning	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3
Roadworks	F	All	Sp.	NSW & WA, Australia	≧95%	Level 3

3.4 Percent of travel mapping

UN Member States have agreed on Global Road Safety Performance Targets to drive action across the world. Among the 12 new targets, Target 4 aims that more than 75% of travel on existing roads is on roads that meet technical standards for all road users (equivalent to 3-star or better) that take into account road safety.

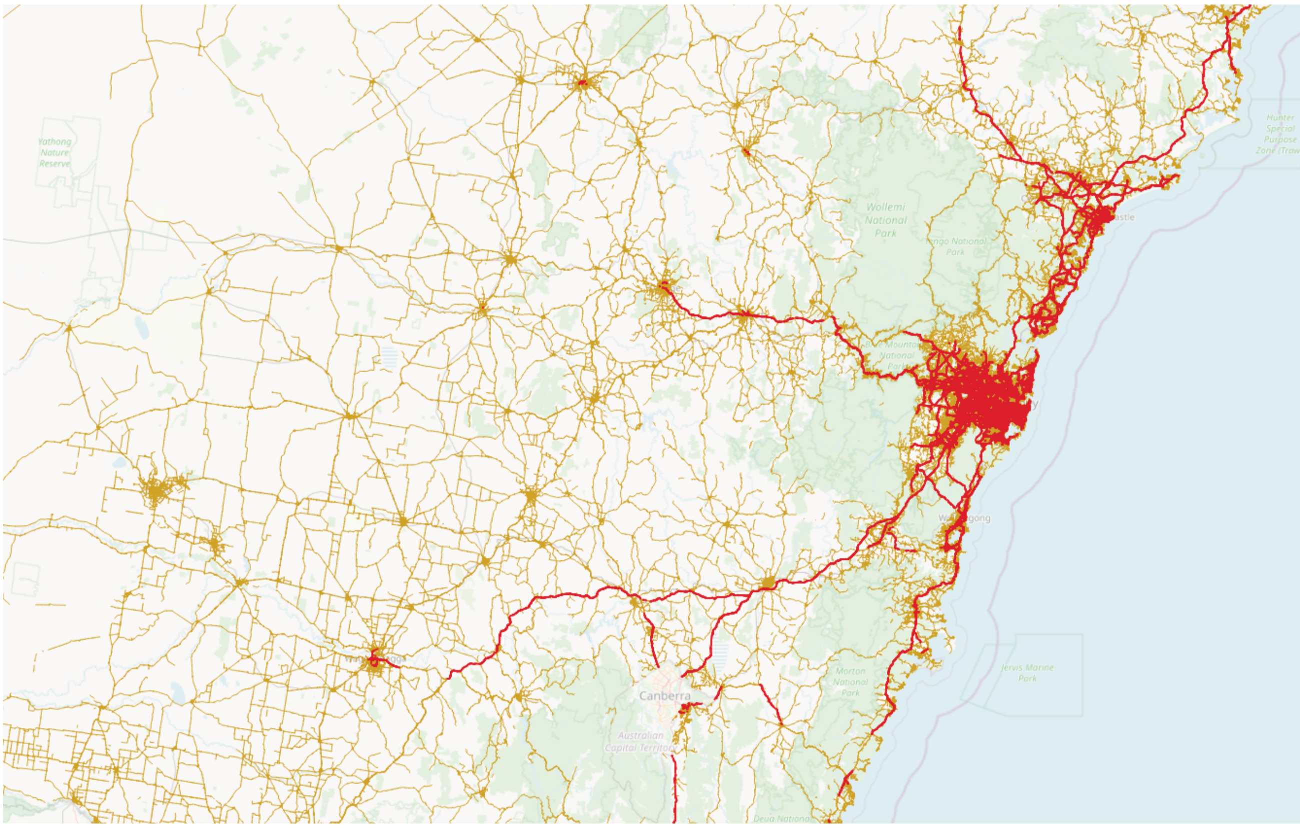
Using TomTom Traffic Stats data, roads carrying 75% of vehicular traffic were mapped for NSW.

The analysis estimates that 7.3% of the road network (9,035km) of the total NSW road network carries 75% of vehicular travel.

As part of the iMOVE project, iRAP agreed to validate TomTom's approach to identifying the roads carrying percentage of travel using the New South Wales data, so the approach can be replicated in other locations in Australia and around the world where TomTom has data coverage. This report is available at www.irap.org/75percent.

Recently announced National and NSW State targets are for 80 percent of travel roads to meet road safety targets. The approach used can be adjusted accordingly.

Figure 12 Where 75% of vehicle travel occurs in the region surrounding Sydney, NSW (yellow – total road network; red – 75% traffic road)



3.5 Limitations identified

In the course of the project, some limitations were identified. First, those infrastructure-related attributes that required advanced or complex data processing were not able to be accredited based on the available data sources for the project or processing capability. These included the 'qualitative' attributes: quality of curve, intersection quality and pedestrian crossing quality. Others such as service roads and bicycle facilities would also require further exploration of data sources and development of detection algorithms.

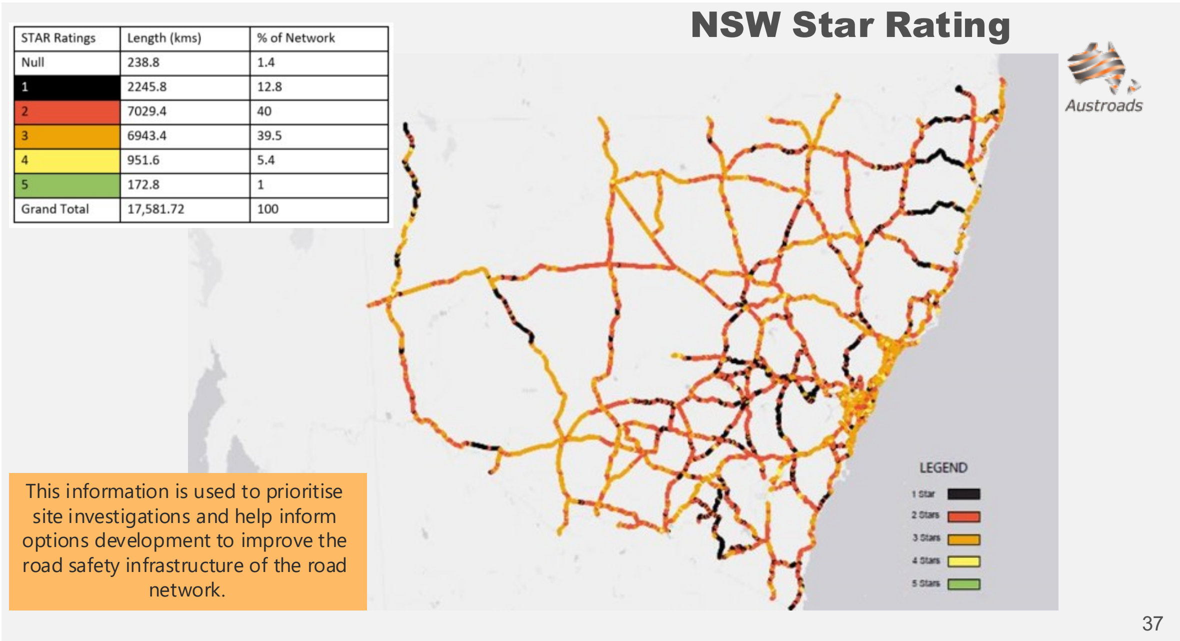
Second, due to the limited range of data sources included in the initial project scope, it was not possible to deliver data on operating speeds or road user flows. During the project, these were identified as being high priority for TfNSW and efforts were made to identify potential data sources. These included:

- Discussing with engineering consultancy company WSP about validating their AADT algorithms using TomTom's Traffic Stats data for the NSW network building on similar work being undertaken in New Zealand. Despite looking promising early in the project, it was not possible for WSP to do this as an in-kind contribution within the project timeframe. This should be explored as part of a possible second phase of the project. TomTom's Traffic Stats data also has operating speed data which could be applied in any subsequent phase of the project.
- Alternative telematics data providers, such as Compass IOT, Ericsson, Otonomo or HERE, could also be considered in subsequent project phases to provide speed and flow data. Harsh braking data also has potential for detecting qualitative attributes (see the AiRAP data specification for more details). Possible sources of other road user flow data (motorcyclists, bicyclists and pedestrians) should be an active consideration in evaluating alternative data sources.

4 NEEDS AND USES OF THE DATA

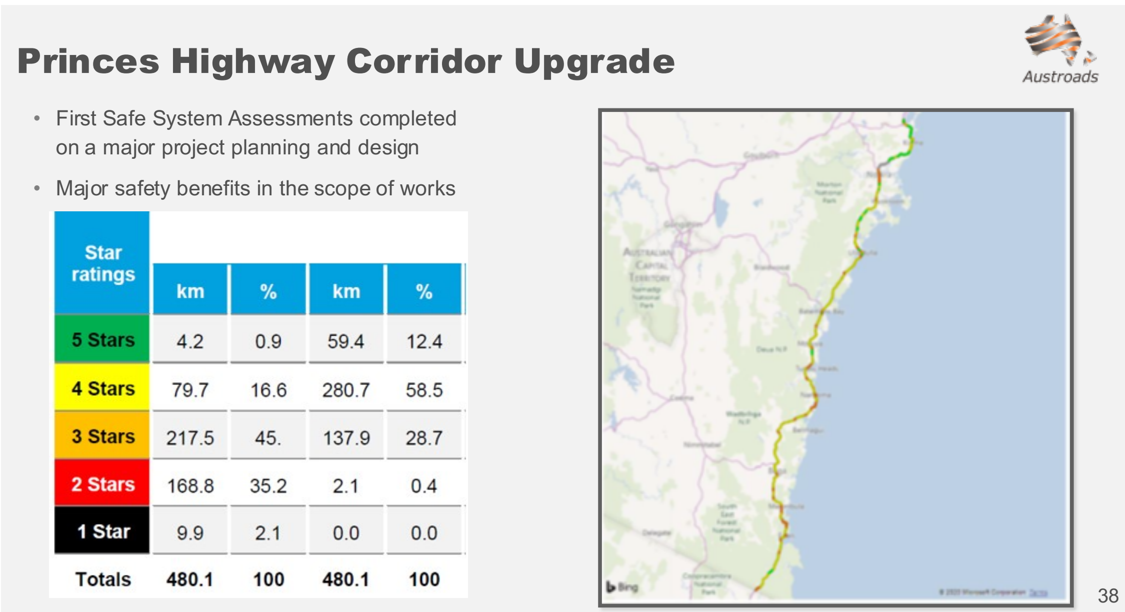
AusRAP data is currently used to inform policy, investment and design decisions. TfNSW, as the manager of the State's 18,000km road network, uses this data to support its overall effectiveness in prioritizing sections of the 18,000km State road network for investigation and investment.

Figure 13 Transport for NSW Star Rating results as presented at Austroads AusRAP webinar Nov 2021



Corridor-level safe system assessments are also utilized by TfNSW to optimise investment that supports the business case for upgrades and maximises the safety performance of the road project.

Figure 14 Transport for NSW After Star Rating results as presented at Austroads AusRAP webinar (Nov 2021)



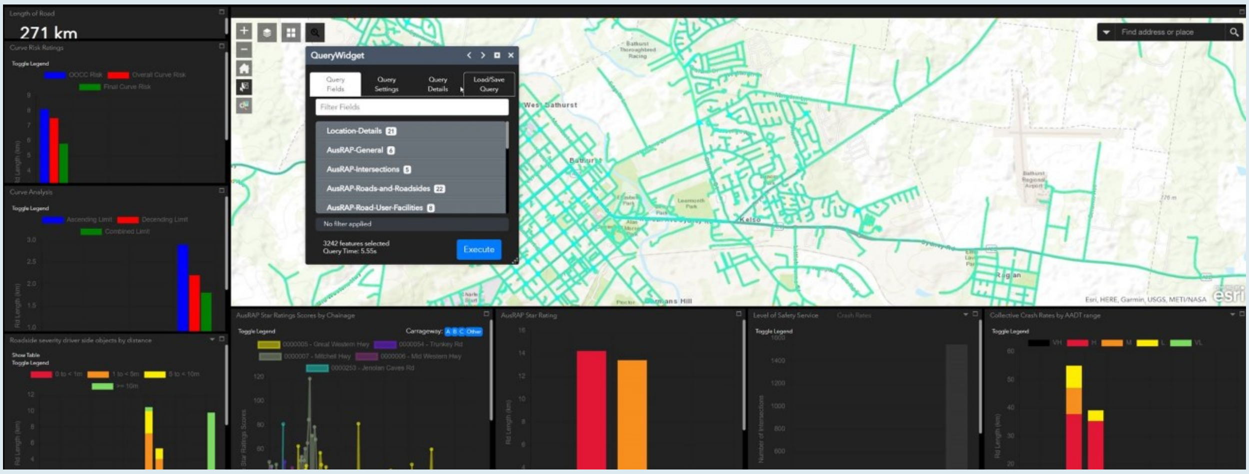
4.1 TfNSW Safer Road Risk Assessment (SRRA) application

The TfNSW Safer Roads Risk Assessment (SRRA) is a spatial application specifically built to provide complex spatial and textual query capabilities in a simple manner.

About SRRA

- A data hub which provides a holistic network level assessment to spatially risk assess roads
- Assists the Safer Roads Task force to perform their task of implementing policies, practices and procedures to reduce the NSW State road toll to zero by 2056.
- Designed to inform decision making for strategic planning, concept development right through to detailed engineering treatment selection.
- Incorporates the pro-active and emerging road infrastructure risk assessment models so that the high-risk sites are identified and rectified before crashes happen.
- Provides spatial analysis of every 100m segment of road in the state of NSW and investment in road safety treatment plans.
- Enables the real time (depending on frequency of data refresh), prioritisation, mitigation and infrastructure risk rating of all hierarchies of roads from State roads to local.

TfNSW's SRRA interface



AusRAP is one of the key models in SRRA. Input datasets go through a number of Extraction, Transfer and Loading (ETL) processes for calculations, manipulation, transformations and curation. This allows data analysis using complex criteria which was previously resource intensive using spatial applications. For example, identification of 1- and 2-star roads with out of context curves with speed limits 80km/h or over.

This would improve the management of the State's road assets by ensuring this data is more reliable and up to date, and ensure safety can be adequately monitored and managed using AusRAP. Currently, road infrastructure data is not readily available for local government roads. Improved data collection and availability will provide a holistic network level road safety assessment of the road environment that would identify road sections with the highest risk of severe crashes on the local government road network. It is important that agencies are able to understand risk across every road in a jurisdiction.

Tracking against percent of travel will also allow network performance to be monitored on the higher exposure areas and prioritization of potential infrastructure funding towards most needed and highest return on investment places. It is noted that lower volume networks should not be excluded from investment with related crash-based monitoring and other techniques used to ensure low-cost high return on investment infrastructure treatments are deployed in addition to speed management and behavioural initiatives where needed.

4.2 National Service Level Standards

The national-level Transport and Infrastructure Council agreed to develop nationally consistent Service Level Standards for all roads.[9]

The framework includes a range of service level standards for:

- Seven primary road categories (national roads through to local streets and significant places)
- Six primary road context attributes (movement counts, place definition, criticality, context and customer types), and
- Eight customer outcome areas (reliable and efficient journeys, safe journeys through to level of amenity provided).

Key indicators included in the National Service Level Standards include:

- Number of fatal and serious injuries annually
- Ability to undertake a safe journey based on the AusRAP Star Rating and the percentage of road segment rated 3-star or more
- Provision of safe areas for customers to stop and rest, and
- Attribute level KPIs related to safe overtaking, sight distance, safe crossing opportunities, safe speeds, access for cyclists etc.

Data is required to monitor these indicators. The methods and sources considered as part of this project to improve the efficiency and cost-effectiveness of road related data would directly meet and/or support the measurement of these National Service Level Standards.

4.3 Local Government roads

Local government roads make up approximately 80% of all roads in NSW. Given their location and purpose, they also have highest levels of pedestrian and cyclist activity.

There is a need for more efficient and cost-effective data to drive road safety decision making on local government roads to ensure road safety is addressed proactively for all road user types.

Extension of the existing AusRAP Star Rating process to a sample of three local council areas in NSW highlighted the lack of consistent asset management data held by the individual local government agencies and the need for a

Figure 15 LiDAR point cloud representation of local government road (Source: Anditi)

[9] The Transport and Infrastructure Council is made up of all the Federal, State and Territory Transport Ministers.

common and repeatable approach that can support rollout across all local council roads in the State.

State and Territory agencies often play a role in building local capacity and assessing safety performance on the local roads of regional significance or similar network definitions.

With 537 local councils across Australia, a viable method to procure and undertake efficient data collection is required to ensure data and evidence-driven investment is targeted to where most deaths and injuries can be avoided.

4.4 Asset management and related purposes

In addition to road safety, road agencies require data for asset management, asset valuation, performance tracking and operational performance. The data required for road safety can also be used for asset management and a range of other purposes, such as:

- Over-height and over-dimension vehicle assessments
- Readiness for advanced driver assistance technologies and driverless vehicles
- Vegetation management
- Environmental monitoring, and
- Sight distance measurement.

4.5 Extended applications of remote sensing data in the road corridor

There is a broad range of potential extended applications for the AiRAP-compliant remote sensing data including aerial and mobile LiDAR; aerial and 360-degree imagery. These have been identified through discussions with industry stakeholders and grouped by subject matter as set out in the table below.

Table 5 Extended applications for AiRAP-compliant and remote sensing data

Application	Description
Road safety assessment – Austroads / AusRAP	Compliance with *National Road Safety Action Plan 2021-2030* requirement – Star Rating of roads 2D and 3D mapping of attributes and attribute categories in GIS, Shape file and CAD formats Accelerated and intelligent coding of safety attributes Accelerated and intelligent auditing of coded attributes Safer Roads Investment Plans (SRIP) Star Rating for Schools (SR4S) National Service Level Standards Visualisation of point cloud and imagery along the road corridor
Road network management	Visualisation of network / road segments Mapping network features and assets Identifying priority locations Informing road investments Development of concept plans for upgrades Safer Road Investment Plans Visualisation and analysis Informing asset management and maintenance Location of poles, streetlights, bus shelters, bridges, culverts, drainage pits Safety barriers (location, length, type) Pavement widths, shoulder width Location, length, width of medians strips and kerbs Location, height and length of retaining walls

Application	Description
	Location of substations and kiosks Line markings (location, length, type and quality) Signs (location, type, quality, obstructions (i.e. vegetation)) Intersection type, merge lanes, channelisation, traffic calming devices etc. Quality of road condition – deformation of road surface, rutting, shoulder or edge of pavement damage Skid resistance Vegetation management Visualisation and analysis using coloured point cloud, 2D and 3D imagery Identification of verge areas for parking during maintenance
Asset and geospatial information	Database/shapefiles of road assets (location, type, length, area, slope/gradient, curvature) Sealed roads, kerbs, shoulders, median strips, unsealed roads, parking space Line markings, pedestrian crossings, channelisation, speed bumps, school zones Signs, poles, streetlights, trees Overhead powerlines, telecommunication cables Safety barriers, pedestrian fencing, retaining walls, bus shelters Bridges, culverts, pedestrian overpasses drainage/gully pits, overland flow paths Retaining walls Kiosk substations and rest areas Vegetation stratification (grass, shrubs, trees) Pervious and impervious areas Validation/correction of existing data Visualisation and analysis using coloured point cloud, 2D and 3D imagery
Development of digital twins of road corridor	3D of road assets (point cloud, voxelization) Classified mobile LiDAR Visualisation and analysis
Road and traffic engineering	Road geometric analysis Vertical alignment Horizontal alignment Continuous sight distance/line of sight Visualisation and analysis
Project planning & development	Concept plans, master plans Roads Cycleways Accessibility for disabled Kerbside space availability – parking Visualisation and analysis
Road safety audits	Baseline information Visualisation and analysis
Heavy vehicle services	Bridge/structure clearance analysis Clash detection/swept path analysis Over size – over-mass vehicle permitting Heavy vehicle specific road safety attributes Identification of potential heavy vehicle temporary parking locations Visualisation of routes
Crash investigation	Base spatial data to assist with investigation Visualisation of crash scene and surrounding context Analysis of attributes such as distance, slope, height, diameter
Connected and autonomous vehicles	Readability of lines and signs Visualisation and analysis High-definition maps of road corridor

Application	Description
	Provision of Star Rating of Roads 4G/5G Network accessibility
Telecommunications	Location, height and extent of existing road corridor powerlines, telecommunication cables and other overhead features Locations of street side buildings, fences, poles, signs and other structures Location of footpaths, kerbs, property access points Viewshed/obstacles for 4G/5G transmissions Distance, length, slope measurements Visualisation and analysis
Environmental assessments	Roadside trees and vegetation coverage Drainage lines, overland flow paths and water bodies Visual impacts Shading analysis Viewshed analysis
Compliance	Star Rating for road safety Strategic targets Road Environmental Asset management and maintenance
Strategic drivers	Increased use of digital information Reduced vehicle kilometres travelled (VKT) Reduced carbon footprint Democratisation of data

4.6 Potential benefits of using remote sensing data in the road corridor

A number of benefits of utilising and accessing remote sensing data (such as mobile LiDAR and 360-degree imagery) in terms of cost, resources and time to deliver projects have been identified. These include:

- **Health, safety and environment benefits** through reduced exposure in data collection and site investigations together with less travel and emissions
- Increased staff productivity through easier access to data – right data/right decision
- **Multiple uses of data that has been captured, curated, quality controlled and made accessible**. This results in less overall cost in data acquisition, data processing and data storage and improved accessibility of data.
- **Greater accessibility and usability of data** – access to cloud-based data from anywhere via web-based portal - data democratisation
- **Ability to check and augment existing data** – comparison of feature location, identification and extent in captured point cloud and imagery data with existing information:
    - Improved investment decisions
    - Higher granularity/resolution of data
    - Greater accessibility to all stakeholders
    - Better and more timely decisions
- **Greater accuracy of data compared to video assessment** – ability to measure, position and analyse features in LiDAR point cloud data to less than 0.1 m relative accuracy
- **Lower costs** of acquiring data per project

- Reduced whole-of-agency tendering and procurement costs
- Lower survey costs
  - Targeted surveys
  - Untargeted surveys
  - Controlled versus uncontrolled surveys
  - Ability to mesh or fuse with other data to improve overall accuracy, resolution and clarity of data
- Project management costs
  - Ability to rapidly visualise and assess project data
  - Greater ability to measure and analyse information provided
  - Greater contract control
  - Reduced time to undertake projects

- **Feasibility Studies, Concept and Master Planning** – rich mobile LiDAR and 360-degree imagery information for the road corridor that can be used in its own right or combined with other data sources (aerial LIDAR, aerial imagery, satellite imagery, ground survey, GIS information etc) to undertake feasibility studies, concept plans and master plans – including retrospective analyses and before/after studies.
  - Reduced need for site inspections
    - Reduced WHS risk
    - Reduced travel time
    - Reduced traffic management costs
    - Reduced disruption to traffic flow
    - Reduced vehicle kilometres travelled (VKT)
    - Reduced wear and tear on road network
    - Reduced congestion
    - Reduced crashes
    - Reduced CO_2 emissions
  - Reduced feature extraction and analytics costs
    - Ability to automate feature extraction
    - Ability to combine with other data sets (satellite and aerial imagery, aerial LiDAR, driver behaviour stats)
    - Greater range of features that can be extracted
    - Higher accuracy (size, height, location) of extracted features
    - Greater consistency of extracted features
    - Ability to automate feature extraction quality control
    - Improved repeatability
  - Ability to generate multiple products for multiple purposes
    - Road Safety Audit (baseline information/ 3D visualisation)

        - Road condition assessments (lines, signs, deformation, skid resistance)
        - Star Rating for roads – reduced road deaths and serious injuries
        - Safer Road Improvement Plans (SRIP)
        - Star Rating for Schools (SR4S)
        - National Service Level Standards
        - 2D and 3D strings
        - 2D and 3D GIS/CAD layers for asset management
        - 3D visualisations
        - Feasibility studies, concept and master plans
        - Readability of roads for connected and autonomous vehicles
        - Bridge clearance measurements
        - Swept path analysis for heavy vehicles
        - HD maps of road system
    - Reduced Compliance Costs
    - Longer term benefits
        - Reference data against which new data can be reviewed and assessed
        - Higher and more accurate levels of automation of data cleansing, classifying and feature extraction
        - Accessible reference data for master planning and concept planning
        - Ability to track change in detail and predict/plan maintenance and upgrade needs
    - Reduced Reputational Risk
        - Better data accessibility and transparency
        - Faster response time – right data for right decisions
    - Enhanced Community Profile and Awareness
        - Roads
        - Environment
        - Community
        - Democratisation of data
        - Data visualisation

4.7 Light Star Rating applications

In parallel to the iMOVE project, iRAP successfully developed a statistically valid light methodology. This enables Star Ratings and FSI estimates to be produced using reduced input data. Its application could potentially reduce costs associated for large-scale Star Ratings or FSI estimations for 'network-scanning' purposes, or enable adaptation of asset data and AiRAP data for producing Star Rating estimates prior to more detailed assessments of high-priority road lengths. The light methodology offers an affordable method for initial network-level safety assessments of the local government road network. Detail on the methodology is provided in the next section.

5 LIGHT STAR RATING METHODOLOGY

In the course of the project, iRAP developed a light Star Rating methodology. This enables statistically valid light Star Ratings to be produced using a reduced set of attributes.

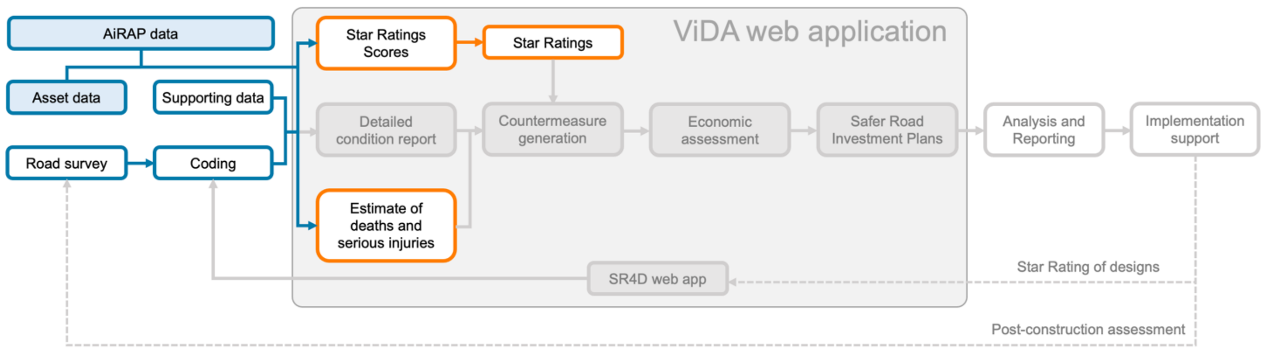
The light data methodology uses a minimum number of road features to predict remaining attribute values to calculate 'light' Star Ratings and FSI estimates. Its application is intended to reduce costs associated for large-scale Star Ratings for 'network-scanning' purposes, or enable adaptation of asset data and AiRAP data for producing Star Rating estimates prior to more detailed assessments of high-priority road lengths.

It has been designed to:

- Be compatible with iRAP's existing software which can produce Star Ratings and FSI Estimation maps.
- Be flexible and scalable, depending on what data is available and its application. The light method may be used with conventional (manual) data collection, AiRAP methods or existing asset management data. It can also accommodate a range of data availability from the minimum attributes required up to where only a small number of attributes are missing. This means estimated Star Ratings can still be produced where data normally required to produce Star Ratings is only partially available.
- Offer a standardised approach with known confidence levels by crash type and road type. The calibration process provides an understanding of how the model will perform for different road and crash types, which can be considered when viewing and using results to ensure limitations are well understood and transparently presented.
- Be used as an interim step to produce full Star Ratings/FSI Estimates and Safer Road Investment Plans. The method allows the available data used for the light Star Rating to be retained and extended upon to produce full Star Ratings.

The light model has more limited functionality compared to full Star Rating assessments. As shown in the diagram below, it **can** be used to produce Star Ratings and FSI Estimations for road networks using the ViDA web application. It **cannot** be used for detailed condition reports, countermeasure generation, economic assessment and Safer Road Investment Plans that require a full level of confidence in actual attribute condition to be effective.

The 'light' iRAP Star Rating and Safer Roads Investment Plan process



5.1 Minimum data inputs for the light data methodology

The light data method uses the same land use, area type and speed and flow attributes as used by AusRAP and a simplified version (present/not present) of 12 others.

Table 6 Minimum data inputs for the light data methodology

Attribute	Specification	Exceptions
Land use	As per iRAP specification (*iRAP Coding Manual*)	
Area type	As per iRAP specification (*iRAP Coding Manual*)	
Vehicle flow data		
AADT	As per iRAP specification (*iRAP Star Rating and Investment Plan Manual*)	
% motorcycles	As per iRAP specification (*iRAP Coding Manual*)	
Intersecting road volume	As per iRAP specification (*iRAP Coding Manual*)	
Speed data		
Speed limit	As per iRAP specification (*iRAP Coding Manual*)	Light Star Ratings may be produced based on Speed limit or Operating speeds only.
Operating speed 85th %ile	As per iRAP specification (*iRAP Star Rating and Investment Plan Manual*)	Light Star Ratings may be produced based on Speed limit or Operating speeds only.
Operating speed mean	As per iRAP specification (*iRAP Star Rating and Investment Plan Manual*)	Required only for FSI estimation only.
Simplified road feature data		
Pedestrian crossing	Recorded as present or not present	Required only for pedestrian Star Ratings.
Sidewalk	Recorded as present or not present	Required only for pedestrian Star Ratings.
Bicycle lane	Recorded as present or not present	Required only for bicyclist Star Ratings.
Motorcycle lane	Recorded as present or not present	Required only for motorcyclist Star Ratings where dedicated motorcyclist lanes are used.
Pavement marking	Recorded as present or not present	Required for all.
>1 lane each direction	Recorded as present or not present	Required for all.
Median island or barrier	Recorded as present or not present	Required for all.
Intersection	Recorded as present or not present	Required for all.
Safety barrier	Recorded as present or not present	Required for all.
Curve	Recorded as present or not present	Required for all.
Good surface condition	Recorded as present or not present	Required for all.
Streetlighting	Recorded as present or not present	Required for all.

5.2 Light data method

The method uses a selection of this input data (area type, speed limit, whether the road is divided and the number of lanes) to organise the data into road types.

These road types are then used as the basis for a series of primary and secondary attribute default rules. Primary attribute rules convert the simplified road feature data into iRAP codes, and secondary rules predict attributes for which data is not collected. The rules can be adjusted for road types, depending on predominant road features for that road type.

5.3 Method development and testing

The method was developed based on reviews of early light data trials and road stereotype approaches that utilise the RAP specifications, data and/or philosophies including:

- Austroads road stereotypes (Australia)
- ThaiRAP pilot (Thailand)
- World Bank RSSAT tool
- TRL for Highways England (UK)
- MRWA road asset data (Australia)
- IRR (New Zealand)
- iRAP simple demonstrator tool (Philippines).

The primary and secondary attribute rules were initially derived from analysis of the seven LMIC datasets: Barbados, Pakistan, Philippines, Ho Chi Minh City (Vietnam), Saint Lucia, Phuket and Hua Hin (Thailand). It was initially aimed that a standard set of 'global' rules could be used to create a single light model. The model testing was then extended to datasets from Australia, Brazil, Mexico, Netherlands and Portugal. These datasets are from HIC and were sample size controlled.

The approach recognised that the prevailing features of a road were strongly dependent on the type of road. To address this, eight road types were defined based on the simplified data. These are:

- Area type (urban or rural)
- Median (single or dual carriageway)
- Number of lanes (one or multiple lanes in one direction), and
- Speed limit

To understand how much the light model results align with those of the full model, correlation relationship analysis between the simple model with the full model was completed using existing datasets available in ViDA.

Calibration test results

The model clearly performed better for the LMIC datasets for which the primary and secondary rules had initially been calibrated. The key conclusions based on the analysis are:

- The differences observed in results from the LMIC and HIC datasets show that it would be very difficult to establish a uniform set of default value assumptions (by way of predictive rules) that would be applicable globally.
- Calibrating the rules according to existing datasets for a given country or region would likely substantially improve correlation of the light method results to that of the full model. The road type functionality in the model is a good framework for this.
- The analysis used in this report allows for a standardised approach by which future datasets can be calibrated based on existing data.

As part of this work, an automated analysis reporting program has been established which allows any dataset to be quickly and easily converted and analysed. It is proposed this is used as a method of calibration.

Based on the findings of the initial development and testing, iRAP then proceeded to calibrate the light model for Australia based on Australian datasets.

6 SCALABILITY REVIEW

The purpose of the scalability review was to:

1. Assess the scalability and repeatability of the pilot approach adopted in the *AiRAP Automation for Australian Road Safety* project, including how amenable the current approach is to deliver accelerated and intelligent Star Rating data on an expanded range of road types and across a larger geographic region than investigated in the pilot. The review is based on sample data and methodological information provided by project partners.
2. Scope future research directions that can leverage or extend the foundation established through the project.

This part of the project benefitted from an in-depth investigation by UTS on whether the data mining methods and approaches trialled in this project could be extended other road networks, other Australian State jurisdictions or even internationally countries. The main benefits that automated road Star Rating method would be a significant reduction of manual road surveys of features in places and areas where it would take a significant amount of time to reach.

UTS conducted first a deep dive into the data specifications and procedures to handle these, which consisted of examining the data flows, investigating the transformation processes and performance of systems associated with the automated road feature extraction as well as making a projection of a future extension of this approach to other local roads in NSW, other states or other countries.

The technical and commercial aspects of scalability are discussed in the following sections. UTS' findings are summarised in Appendix B.

6.1 Technical considerations

Numerous factors have potential to impact on the ability to use AI and other accelerated technologies to extract Star Rating safety attributes consistently and accurately from mobile LiDAR and 360-degree imagery in a scalable manner. These are primarily the availability, accessibility and quality of the data.

The range of attributes and attribute types to be detected in Star Rating a road is significant. Not all attributes and attribute categories are suitable to be detected using mobile LiDAR and 360-degree imagery.

Of those that can be detected, some are more reliably detected using a coloured point cloud or intensity-of-return information, some using imagery and some needing a combination of all seven characteristics available from mobile LiDAR and imagery (x, y, z and intensity-of-return derived from LiDAR and red, green, blue bands derived from imagery). The most reliable technique or combination of techniques also varies depending on data quality and extent of coverage. For some attributes to be reliably detected, other data sources such as aerial imagery, satellite imagery and aerial LiDAR may also be required.

6.1.1 Data availability and cost

There are many companies that capture mobile LiDAR and 360-degree imagery across the globe. These range from those who provide "off the shelf" LiDAR (such as TomTom's MoMa data which is available globally), through to survey controlled custom capture that is captured on a project-by-project basis.

Mobile LiDAR ranges in its accuracy, capture width and level of pre-processing from uncontrolled single pass LiDAR through to survey-controlled LiDAR derived from multiple passes of a capture vehicle. Single pass LiDAR is significantly cheaper than survey grade LiDAR but can be impacted by proximity of vehicles in adjoining lanes and satellite availability for GPS during capture.

The cost of mobile LiDAR and 360-degree imagery ranges from approximately $30/km to $7500/km with data from varying sources, varying levels of accuracy and capture processes requiring different types and levels of processing.

There are numerous suppliers of 'off the shelf' and custom capture mobile LiDAR and 360-degree imagery across Australia and globally making this type of data widely available.

6.1.2 Data accessibility

Data is accessible through a range of means with some providers making data available via an API, others via the cloud or other forms of file sharing. Data can be available via cloud-based marketplaces or directly from the supplier on request. Access times can vary from hours to weeks/months depending on:

- How much data is required
- Whether it is already available or if it requires further processing before delivery, or
- If it still needs to be captured.

6.1.3 Quality control of input data

Consistent quality of input data is essential to the scalability of the approach.

Although data can be accessed from many sources and suppliers, there are many occasions where the choice of data suppliers or data availability will be limited, or the cost of alternate data sources will be prohibitive. In these circumstances, it may be necessary to use a range of technologies and processes to firstly rectify the shortcomings of the input data as far as possible and then process the data using techniques that work best with the quality of data that is available. This can include AI, accelerated, semi-automated and manual techniques.

Outputs from quality controls help inform the types of technologies and processes that need to be used to successfully extract Star Rating safety attributes from the available data.

Data quality can vary significantly depending on a broad range of factors such as:

- Scanner and IMU quality and setup
- Capture conditions (GPS satellite access, weather, light, traffic congestion)
- Driver behaviour
- Survey control utilised
- Number of passes of capture vehicle used to generate point cloud, and
- Type and level of pre- and post-processing of data.

Other factors also need to be considered such as whether mobile LiDAR and 360-degree imagery includes road segments that are not part of the road being data Star Rated. These segments need to be detected and removed before further analysis.

6.1.4 Processing data at scale

Extracting Star Rating attributes from mobile LiDAR and imagery in a consistent and repeatable manner requires the use of a range of technologies and approaches and the ability to modify or adapt these approaches as and when required. The need to readily modify and adapt these technologies and approaches is driven by the rapid changes that are occurring in capture technology and the range of varied road conditions, signage and road markings that are used across the globe.

Assuming the UN target for 75% of travel of the world's roads to be on the equivalent of 3-star or better roads requires assessment of approximately 10 million kilometres of road and approximately 15% of the world's road network is modified each year, it is estimated that approximately 25 million kilometres of road will need to be targeted by 2030 to achieve UN global targets for road safety.

Accessing, handling and processing Mobile LiDAR and 360-degree imagery in its raw form typically accesses approximately 1.25 TB of data per 1000 kilometres. Processing of attributes and generating products that can be utilised globally add another 10% to 20% to this volume of data if the data is transformed and managed as efficiently as possible.

On this basis, assessing and Star Rating 25 million kilometres of road by 2030 will require the handling, assessment, delivery and sharing of over 40 Petabytes of data. Processing and assessing Star Rating of roads using mobile LiDAR and 360-degree imagery can require that the data is processed at several scales using several different techniques. To meet iRAP specifications for Star Rating processing capabilities, data needs to be assessed using 100m (and possibly 10m in the future) segments. It also must be assessed at a larger scale to determine attributes such as the start and finish of an intersection, merge lanes, school zones or the length of a lane or carriageway which may extend over several consecutive 100m segments.

As a result, processing and assessment techniques used need to cater for these various requirements. Given the volume of data that needs to be onboarded, handled, processed and delivered, the techniques used need to be computationally efficient and fast. The final product (i.e. coded attributes) needs to be as small as possible to enable it to be shared and used globally or made available via a cloud-based marketplace.

6.1.5 Scalability of project outcomes

This section considers the scalability and repeatability of the specific approaches and data sources piloted in the project, as well as the project outcomes to support this.

To conduct the current scalability report, UTS explored the current set of data and processes employed by project partners in producing a final set of unified road features. This presents an analysis of the available data sets and underlying processing frameworks and draws on this analysis to assess potential scaling of the approach to a greater diversity of roads within NSW and to jurisdictions beyond NSW (including internationally).

A summary of their findings is provided in Appendix B.

Available data sources

The proof of concept used in the iMOVE research project included a baseline of 23,652km of road network that was provided by TfNSW. This covered all major motorways, freeways, highways, arterial roads and sub-arterial roads. This represents approximately 13% of the total NSW road network.

Apart from the baseline road network from TfNSW, TomTom has provided access to the entire Multinet-R (MN-R) database, which covers 99% of the top 8 road classes in Australia. This facilitates a straightforward, large-scale feature extraction from MN-R data, provided that the road segmentation to be used is compatibly aligned prior to feature extraction.

Data definitions

To enable automated and accelerated techniques to be used as part of AiRAP coding of road safety attributes, more detailed definitions and specifications for many of the 220 attribute categories were required. As part of this process, an AiRAP data specification providing additional detail for attributes and attribute categories was developed by iRAP in consultation with Anditi. The definitions have been compiled into a Data Dictionary that will be published and refined on an ongoing basis.

Development of more detailed definitions and specifications for road safety attributes and attribute categories will require sensitivity testing and historic comparisons between manual iRAP coding and AiRAP coding taking these changes in definition and specification into account.

6.2 Commercial considerations

6.2.1 Data suppliers and market

Use of remote sensing data to help reduce road deaths and serious injuries is an emerging field. Likewise, the use of this data for the development, assessment, management and maintenance of road networks is relatively new. As these emerging fields develop and many of the current challenges are addressed, the costs, resources and time required to undertake projects will significantly reduce.

A range of new suppliers have already demonstrated the potential to generate AiRAP data through a series of pilot studies worldwide. New suppliers are also undertaking their own internal investment to generate AiRAP compliant attributes.

At the global level, a range of suppliers have expressed interest and/or are in the process of accreditation using the new AiRAP Framework developed in this project. As a harmonised global standard for data coding the advances by one supplier in one country should be adaptable to Australia and vice-versa providing benefits to both data suppliers and data consumers. Examples of potential and existing AiRAP data suppliers include but are not limited to:

- Agilisys
- Anditi
- Google
- Here Technology
- Mobileye
- Navinfo
- NCTech
- Retina Visions
- The Floow
- TomTom
- University of Washington
- University of Zagreb
- Main Roads Western Australia
- Zenseact

While the costs of the data are expected to reduce over time as the market scales and competition increases, traditional single purpose and sporadic road agency use is insufficient to create the necessary market demand. Expansion of the road agency use of data to include some of the potential applications and benefits discussed in Section 4 will assist in generating demand.

Alternative demand for infrastructure data is therefore needed to maximise the potential of AiRAP compliant data to scale on a local and global level and encourage technology companies to invest in the generation of the data to the iRAP global standard.

iRAP has been engaging with commercial companies with a demonstrated interest and demand for some, or all of the iRAP attributes and associated metrics including Star Ratings and fatality estimations collected to the same global standard regardless of country. Greatest potential exists in fleet management, insurance markets and navigation companies with a range of potential customers identified who value the global harmonised iRAP infrastructure data. This commercial demand for iRAP compliant data creates the necessary market demand for suppliers to continue investment in their source data and AI efforts.

To support this initiative a self-sustaining single data marketplace is being developed to support the commercial companies who can benefit from the iRAP attribute data. This approach will also help the wide variety of AiRAP data suppliers who may not have the resources or risk appetite to take on the essential marketing costs to generate sales. The single marketplace will focus on being the honest broker between the myriad of AiRAP data suppliers and the many AiRAP commercial data consumers and in doing so remove the barriers and cost to entry for all market participants.

6.2.2 Cost savings and efficiency gains

The increased commercial company demand for AiRAP compliant data will ensure sufficient market opportunities to encourage new AiRAP data sources and processed data to be created. This will ultimately lead to a greater scale and coverage of data and a lower cost per unit of data. This will benefit traditional road safety practitioners, accredited service providers, accredited suppliers, government and development partners of iRAP.

Potential also exists to facilitate global reporting of metrics for the public good as the volume of data increases (e.g. satellite or telematics-based attributes at the global level) that can be shared through the UN, World Health Organisation, Regional Road Safety Observatories and national road safety data hubs. This will provide greater awareness of the opportunity to save lives and contribute to the achievement of the UN Sustainable Development Goals and UN Decade of Action Plan for Road Safety.

Based on the pilot projects to date, it is estimated that the use of remote sensing data and AiRAP techniques will initially result in savings in the order of 10% to 30% compared to manual methods. However, it is expected that AiRAP techniques will significantly improve the accuracy, reliability and repeatability of the attribute coding. It will also, with the development of specialised quality control tools, significantly enhance the ability to review and audit coded attribute data.

Further reductions in cost, time and resources will be realised as AiRAP techniques are refined. Further reductions will be achieved as the quality and coverage of available remote sensing data increases and techniques are developed to make the data more accessible. These cost reductions will be for both data collection and coding associated with Star Rating assessments specifically. It will improve the breadth and cost effectiveness of remote sensing data collection (such as LiDAR and 360-degree imagery) which has broader road corridor and road network applications. The multiple use of coded data by both public and private sector stakeholders should also see scale, frequency and cost benefits for all stakeholders.

It is envisaged that the cost of Star Rating roads to a high and consistent standard will reduce by over 50% within 5 years. At present, the cost of Star Rating of roads using video ranges from about A$30/km to A$50/km for capture and about A$50/km to A$75/km for coding and quality assurance. Costs typically vary according to rural or urban area types and the complexity and consistency of the road environment. Analysis and reporting needs are often road agency specific and additional to collection and coding costs.

Significant reductions in cost, time and resources are expected for remote sensing data where the multiple uses of the data are maximized. At present, project specific LiDAR survey costs currently range from A$1000/km for uncontrolled surveys and A$7500/km for controlled surveys. These are expected to reduce to less than A$150/km. Similarly, based on available costings, specific applications such as bridge clearance surveys (which currently cost approximately A$1500 per bridge) will be able to be done using mobile LIDAR and 360-degree imagery cost reductions of approximately 60% to 70% anticipated.

Further examples of these types of cost savings are currently being compiled through consultation with a range of road authorities and stakeholders across Australia for:

- iRAP Safety Star Rating of roads
- Analysing line markings, sign locations, readability of lines and signs
- Development of Safer Road Investment Plans
- Concept and Master Planning in the road corridor
- Bridge clearance surveys
- Swept path analysis and oversize vehicle permitting
- Road asset assessment and inventory studies, and
- Parking analysis.

7 CONCLUSIONS AND RECOMMENDATIONS

7.1 Summary of project outcomes

The methods and mechanisms developed as part of this project represent a world-first for the systematic extraction and accreditation of road infrastructure data from readily available data sources. Furthermore, with the use of the iRAP global data specifications, this sets the precedent for consistent and universal application of this data to improve road safety outcomes.

Upon completion, the project has successfully:

- ✓ Delivered 20,000km of attribute data (7 attributes) using automated data extraction methods from existing data sources
- ✓ Proved viability of extracting 52 Star Rating attributes using 'off-the-shelf' data with 34 of these being AiRAP accredited for further use
- ✓ Established an accreditation framework for doing the same for other data sources and suppliers
- ✓ Mapped roads carrying 75% of vehicle travel for the State of NSW
- ✓ Established a 'light' Star Rating method for network scanning for local government roads, and
- ✓ Mapped potential data providers and estimated cost savings and efficiency gains of using new data sources and extraction methods.

This project has pioneered concepts and methods which have long-term implications for TfNSW, other Australian jurisdictions and internationally. The project comes at the beginning of the UN's Second Decade of Action for Road Safety, and coincided with the release of Australia's National Road Safety Strategy and NSW's Road Safety Plan. All three have road risk ratings as a critical component for road safety management and reporting.

Data is needed to drive these road risk ratings on an unprecedented scale. This project is the first of its kind globally, and its activities and outcomes are of interest to stakeholders across Australia and around the world.

The table below highlights progress made across a number of areas since the project commenced.

Table 7 Advance of the 'state of the art' resulting from this project

Situation pre-project (2020)	Situation upon project completion (2022)
No systematic method to convert or verify data from other (non-manual) sources.	The AiRAP accreditation framework provides a standardised process and validation for the conversion and validation of road infrastructure, speed and flow data.
Limited knowledge about possible data providers and sources.	Understanding of current and potential data providers and range of data sources available across Australia and internationally.
Approximately 13 attributes able to be automatically extracted from off the shelf LiDAR data.	Automatic or accelerated data extraction for **all** AusRAP infrastructure attributes has been achieved. 34 are now AiRAP accredited with an accuracy of >95%.
Specifications for manual coding processes only.	Data definitions for automatic coding of infrastructure data established and documented.
Precision data, such as LiDAR, is expensive. The cost of bespoke LiDAR surveys approx. $250 per km or more.	Pre-collected LiDAR data from existing sources proven to meet road data needs at 1/10th of price (approx. $27 per km).
Costs of automatic data extraction unknown.	Costs and potential efficiency savings known.

AiRAP aims to capture advances in artificial intelligence, machine learning, vision systems, LIDAR, telematics and other data sources to deliver critical information on road safety. The accelerated and intelligent collection and coding of road attribute data has the potential to reduce the **time, cost and effort** required to undertake road safety assessments and improve the frequency and accuracy of data.

With the *Global Plan for the Decade of Action for Road Safety* and national policy targets to increase the percent of travel on 3-star or better roads, road agencies have highlighted their need for supplier independent data collected to the iRAP global standard. To support these needs, iRAP is committed to creating an open marketplace for the data that reduces the costs for all those who can benefit from using the data.

This project was established to explore the potential for accelerated and intelligent approaches to be used for gathering data for AusRAP road safety assessments and documenting consistent global specifications for the accreditation of data. Improving the efficiency and accuracy of data collection, and putting such data to multiple uses, has the potential to deliver substantial cost and efficiency savings to TfNSW.

A primary outcome of the project is the development of the AiRAP Framework to support the use of iRAP-compliant data from pre-collected sources. The framework has been proven using the initial capabilities of TomTom and Anditi and is available and applicable to all prospective data suppliers within an open data marketplace.

The table below shows the coverage of the MN-R and MoMa data sources investigated during this project. They have been demonstrated to be capable of capturing most AusRAP data attributes using either fully automated or accelerated (partially automated) methods. These have now been accredited by iRAP for ongoing application on Australian roads. Those which are applicable to the light Star Rating methodology are shown in blue.

Table 8 Attribute coverage resulting from the project

iRAP attributes	TomTom MN-R data	Anditi MoMA data	
Speed limit			Accelerated
Differential speed			Auto
Speed management			Auto
Number of lanes			Auto
Lane width			Auto
Curvature			Auto
Quality of curve			Qualitative attribute – See AiRAP data specification for advice on source data and attribute detection.
Median type			Auto
Skid resistance			MN-R Level 1 (sealed/unsealed only) MoMa Level 3 Accelerated
Road condition			Accelerated
Vehicle parking			Auto
Grade			Auto
Sight distance			Accelerated
Delineation			Accelerated
Street lighting			Accelerated

	TomTom	Anditi	
iRAP attributes	**MN-R data**	**MoMA data**	
Service road			More work required on automated techniques for attribute detection. Expected to be available from MN-R
Centre line rumble strips			Auto
Roadside severity – distance (DS/PS)			Accelerated
Roadside severity – object (DS/PS)			Accelerated
Shoulder rumble strips			Auto
Paved shoulder (DS/PS)			Auto
Intersection type			Accelerated
Intersection quality			Qualitative attribute – See AiRAP data specification for advice on source data and attribute detection.
Intersection channelization			Auto
Property access points			Auto
Intersecting road volume			Additional data source required
Land use			Accelerated
Area type			Accelerated
Pedestrian crossing facilities			Accelerated
Pedestrian crossing facilities quality			Qualitative attribute – See AiRAP data specification for advice on source data and attribute detection.
Pedestrian fencing			Accelerated
Sidewalk (DS/PS)			Accelerated
Facilities for motorcycles			Low priority - not used in Australian road designs
Facilities for bicycles			More work required on automated techniques for attribute detection.
School zone warning			MN-R Level 2 (signs/not applicable only) MoMa Accelerated
School zone crossing supervisor			Additional data source (from schools) required.
Roadworks			Accelerated
Vehicle flow (AADT)			Additional data source required
Motorcycle %			Additional data source required
Pedestrian peak hour flow across the road			Additional data source required
Pedestrian peak hour flow along the road driver-side			Additional data source required
Pedestrian peak hour flow along the road passenger-side			Additional data source required

iRAP attributes	TomTom MN-R data	Anditi MoMA data	
Bicycle peak hour flows			Additional data source required
Operating speed – 85th percentile			Available from TomTom Traffic Stats
Operating speed – mean			Available from TomTom Traffic Stats

	Colour key
No accredited data source	
Accredited data for full attribute – fully automated	Level 3 (auto)
Accredited data for full attribute – partially automated	Level 3 (accelerated)
Accredited data for partial attribute (only some sub-attributes available)	Level 2
Accredited data for a simple version of the attribute (may not have an equivalent sub-attribute code)	Level 1
Attributes required for light star rating. **Bold** font indicates the full attribute, not bold indicates a 'simple' version only is required.	Blue text

7.2 Next steps

This project was conceptualised as an initial phase of research and development to establish the framework needed to utilise 'off-the-shelf' data sources with a view to improving the efficiency and cost of data collection for road safety assessments. With the AiRAP framework now established and the concepts proven, the outcomes of the project provide a solid basis for a second phase of research and development to further this work and address remaining priorities.

Road user data remains a priority for all road agencies and road safety stakeholders worldwide. In particular, improved data sources for vehicle flow (AADT), as well as other modes such as that for pedestrians, bicyclists and other light mobility users and public transportation.

The need to extend data-driven and evidence-based decision to local government road networks is also a priority. Efforts to streamline data collection and analysis techniques that fully integrate with road agency systems have yielded results as part of this iMOVE project with further advances needed in the future.

7.3 Phase 2 recommendations

With view to building on the success of the first *AiRAP Automation for Australian Road Safety* iMOVE partnership (referred to here as 'Phase 1'), recommendations are proposed for further collaborative research.

A second phase is proposed to pursue a number of high priority actions. These are:

1. Further develop capability for attribute data, especially speed and flow attributes

To focus on the generation of high value, low-cost attributes that can be generated across the state and local government road networks with associated confidence levels. Priorities will include AiRAP-compliant data associated with road user flow (e.g. AADT, pedestrian and cyclist flows) that build on global efforts to generate this data as well as speed related data (e.g. 85th percentile speed, mean speed).

This work would involve the identification and evaluation of available sources of AADT and speed data, and the potential of this data to further contribute to automated extraction of AusRAP data from existing sources. For example, where harsh braking data is available, how this can be used in conjunction with the data already available to produce the additional qualitative attributes (intersection, curve and pedestrian crossing quality).

Phase 1 identified a number of potential suppliers of this data. Comparison and evaluation of the alternatives would consider:

- Data coverage and quality (penetration rates)
- Costs of data purchasing and processing
- Potential for broader application to contribute to the completeness of AusRAP data, and
- How to present data in the data marketplace that captures limitations and potential use

A second part of this activity would be the delivery of this data for the full (or subset of) NSW road network, including local government roads.

2. Pilot light Star Ratings for local government road networks

By drawing on the existing data sources in Phase 1, Phase 2 would focus on further development to improve the efficiency and scale of the data extraction for simplified road attribute data across the local road network. That data, together with the speed and flow data, would then be used to pilot light star ratings for the local government road network.

This would build on the findings and learnings of Phase 1 to further develop capability in Machine Learning and AI methods for data collection and analysis. This could also involve further R&D work by UTS which would aim to address the complexity of manually filtering and choosing which features can be extracted to inform the road star rating by using several Machine Learning (ML) models that can learn from all feature information and data sets and make predictions of a Star Rating when a new road dataset becomes available with new sources that have not been explored before.

In conjunction with the AusRAP Steering Committee, program implementation guides and associated on-line training resources would also be developed to support the appropriate use of scanning tools and detailed assessment techniques that meet the needs of road agencies and local governments.

3. Continue refinement of AiRAP ready attribute and category definitions and extraction techniques

Phase 2 would aim to continue the development of, and build on, the accelerated and intelligent extraction of road safety attributes, such as those that have been demonstrated by Anditi as part of Phase 1. To maximise the utility of the data and potential cost savings, this should include the concurrent identification and extraction of other road corridor attributes and applications.

This work would build on the work undertaken during Phase 1 with the aim of generating robust digital definitions for all iRAP attributes and categories. Additional research should be undertaken to explore the impact that any changes to attribute specifications and definitions may have on Star Ratings so that like for like comparisons can be made.

4. Use a data-driven method to model vulnerable road user flows

Phase 2 would also aim to improve AusRAP functionality and use through modelling vulnerable road user (pedestrians/ motorcyclists/ cyclists) flows from available data sources and ML methods.

Other data sources, such as mobile phone and land use data, can provide a detailed and good source of information with regards to pedestrian movement across any city in Australia. It can be used for aggregating people’s movements in a specific dedicated walkway, schools, childcare, etc. Sources such as mobile phone data can provide preferences for pathway utilisation in the city and can be used in a heatmap accessibility analysis.

This would allow for Star Rating, fatality estimation and safer road investment plan models for each road user group, which is particularly important for local government road networks.

5. Further develop the data marketplace

This will include assessment of price sensitivities and the efficiencies of single data generation activities across multi-agency needs. Investigate need for and mechanisms for geospatial alignment of data and the feasibility of API integration with AusRAP related processing of data within ViDA.

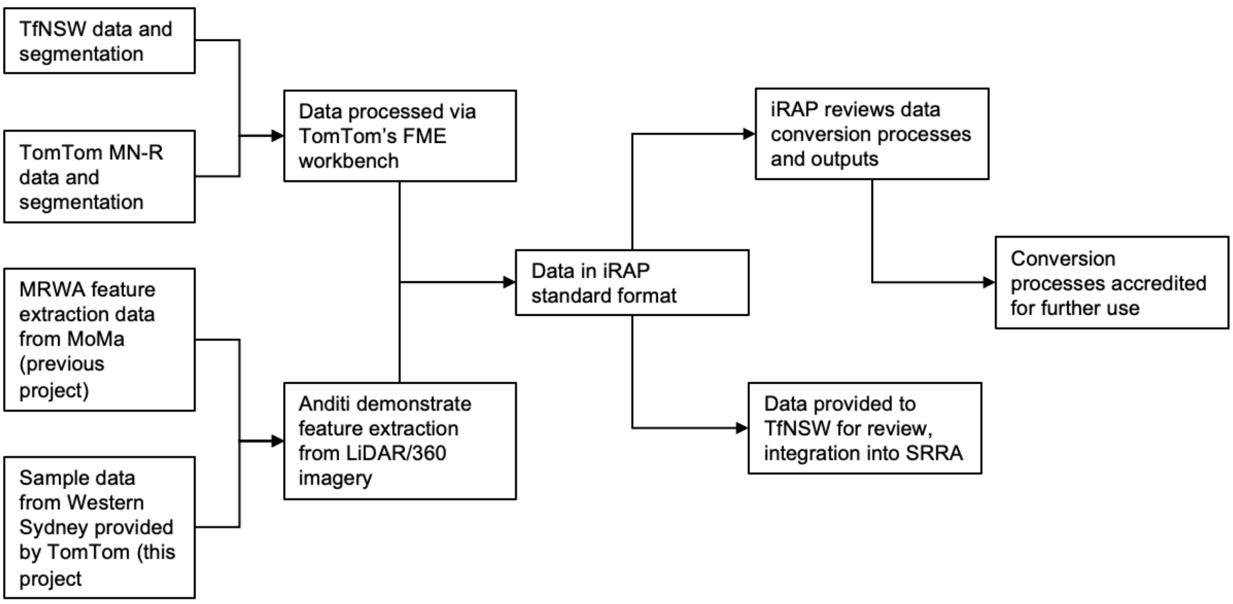
APPENDIX A

Source data conversion: Methodology and Analysis

The major data inputs of this project which are TfNSW data, TomTom MN-R data, Main Roads WA Lidar project outputs conducted by Anditi, and TomTom MOMA data.

TfNSW, TomTom, and Anditi processed and refined the datasets respectively in order to achieve the final data outputs. Major data processing procedures include TomTom FME workbench, Anditi's merging of the iRAP and TfNSW data, Anditi's feature extraction from MoMa data, and TfNSW integration MN-R-iRAP data with RAF database and compare results.

Figure 16 Overview of data processing activities



Method of conversion used for MN-R data

The TomTom MN-R data is captured using globally consistent specifications. All changes made to the MN-R data are done so through their transactional map ecosystem. This map-making platform can introduce over one billion changes monthly to the world map repository.

To ensure accuracy, TomTom's data is subject to checks against strict quality rules before going live. Currently, TomTom have defined around 3,000 business quality checks. Every change is traceable in terms of operations performed, as well as accountability.

MN-R data does not have correlating image data to enable manual review, however, TomTom does capture MoMa data on a continuous basis to ensure there are data sources to support the values provided through the MN-R.

The conversion process requires TomTom's MNR2iRAP FME workbench that is run on request converting the MN-R data into the AiRAP attributes.

In this project, the MN-R data was supplied to Anditi by TomTom in CSV format. The data was plotted in QGIS using the Add Layer Delimited Text (CSV) method. The data was supplied in PSG:4326 – WGS 84 – Geographic and needed to be converted to EPSG:3308 – GDA94/NSW Lambert – Projected to enable accurate spatial analysis with the TfNSW data. This was done using standard GIS conversion processes.

The MN-R data attributes were then joined to the TfNSW geometry using the 'Join attribute by nearest' processing tool. The maximum nearest neighbours were set to 1 and the maximum distance was set to 10 metres. The result was output as SQLite layer with the geometry of the 100 metre segments (as supplied in the TfNSW AusRAP data) and the attributes of the MN-R data and the TfNSW data.

Review of the data indicates that this joining method does produce some errors and it cannot be fully guaranteed that the point on the 100 m segment used to join the data is the intended point. The data supplied by TomTom has an oversupply which due to road geometry can include points that relate to the attributes of roads that either join or run parallel to the roads in TfNSW dataset (i.e. those being analysed).

To avoid this in the future, it is recommended that future data extractions be done by setting up the FME workflow to produce data for pre-determined and agreed segmentation, and not by joining the data using a GIS based method post extraction.

Alignment with iRAP attributes

Taking into account the above limitations, the join by nearest approach was considered successful, resulting in all features in the dataset being attributed with data from the MN-R CSV. The analysis of the results focuses on the following nine iRAP attributes:

- Land Use (Drivers side and Passengers side) – 7 Categories
- Operating Speed (mean and 85th percentile) – all categories
- Speed limit – all categories
- Differential Speed – present or not present
- Intersection Channelisation – present or not present
- Number of Lanes – 6 categories
- Curvature – 4 Categories
- Grade – 5 Categories
- Skid resistance/grip – 5 Categories
- School Warning Zone – 4 Categories.

The results were assessed at a feature level and a road class level. The purpose of this exercise was to determine which attributes were appropriate to be accredited by iRAP. Of the attributes analysed, Speed limit, Differential speed, Number of lanes, Curvature, Grade, Skid resistance and School zone warning proceeded to accreditation. Operating speed, Land use, and Intersection channelisation were eliminated due to lack of data.

Applicability and scalability of MN-R data

The UTS review of the MN-R data included overlaying information with historical TfNSW AusRAP assessment data that has been collected using conventional approaches.

The iRAP-compliant MN-R attributes are automatically generated using an FME workbench. The approach is considered repeatable and scalable for extending to other road networks and lengths.

As detailed in the conversion of MN-R data section, the method of joining features post extraction from MN-R can produce some point-matching errors and it cannot be 100% guaranteed that the point on the 100m segment used to join the data is the intended point. For example, there are 78.6% (166,968 sections) of all sections that are mapped to the points within a distance of 0 to 2 metres, 16% within a distance of 2-4 metres and 5.4% of all sections are mapped within a distance greater than 4 metres.

Mismatch errors can be minimised by transport authorities (such as TfNSW) if the co-ordinates for each 100m road segment are provided as a shape file and future data extractions are done by making the required changes to the FME workflow. This location data can then be used to extract the relevant MN-R values and/or identify the road segment for coding of additional attributes using the MoMa data.

The initial supply of shape files by the relevant road agency provides a valuable first step that can be included in the AiRAP conversion process to assist scalability across new road networks in Australia and worldwide. Where these existing shape files are not available, new segmentation can be automatically generated by mapping tools and/or accredited systems such as Anditi's Roadviewer Inspection System.

The data oversupply is a positive issue and ensures that all the TfNSW road network segments have at least one data point to be mapped, however the cleaning of any oversupply is essential. The Anditi output data aligned well with TfNSW data.

For the purpose of this project, Anditi removed any roads where there was no MN-R data available on specific sections of the TfNSW network. This issue impacted a relatively small length of the road network where a few roads were found to be missing in Anditi's MN-R network when comparing to the TfNSW network.

Anditi recommends that the future data extractions are completed by making the required changes to the FME workflow at the commencement of the project and not by joining the data using a GIS-based method post-extraction. Updates to the MNR2iRAP FME workbench are planned to facilitate automatic road matching and feature extraction from the MNR-data base according to TfNSW (and other agency) road segmentation. The UTS review also reinforced this recommendation that will better align data against base maps/networks from each territory and also save time, improve accuracy and limit any point mismatching.

TomTom regularly conducts matching exercises between various state road sources and the MN-R database, however this process will likely need to be rerun for other states that require a fixed 100m segmentation process. This issue can be readily addressed by the parties involved agreeing to road segmentation and alignment that is to be used at the start of the project before any data is extracted from the MN-R database.

Automatic curvature and gradient calculations which are currently provided as part of the MN-R database are also recommended provided the procedure is validated against multiple road networks/areas to avoid potential errors.

Accreditation of MN-R data attribute conversion

In order to validate the data outputs, two sample datasets were quality checked (both ~10km, one urban, one rural). As a result of this check, additional data was requested for where *Differential speed* was recorded as 'present' and where *Skid resistance* was recorded as something other than 'Sealed – adequate' to demonstrate where these sub-attributes had been recorded. Two additional coding files were provided with data extracted from the total 20,000km of coding available. Six samples of coding segments/road sections were reviewed from each and the results documented in the supplementary reviews. These files were also used to do spot checks on other attributes as well.

All data was found to be of high accuracy and applicable to both urban and rural areas. Post processing quality assurance is recommended for all data to help identify any anomalies in the data. This is particularly recommended for Skid resistance, Number of lanes and Speed limit attributes where mapping (such as that shown in MN-R derived attribute maps) can be used to identify (and correct where necessary) changes in coding which extend for less than 4 consecutive coding segments (<400m).

Of the attributes evaluated, all but two covered the full set of attribute features. For skid resistance, it was found that the data could only reliably determine if the road is unsealed or sealed, and not the full range of surface quality that is used in iRAP assessments. For School zone warning, the MN-R data does not record flashing beacons separately to signage. As such, this may not be suitable for application where flashing beacons are frequently used in school zones, or where school zones are not present or are not being assessed.

Method of MoMa data feature extraction

Data is sourced from TomTom's Mobile Mapping (MoMa) program and is collected using vehicle mounted 360-degree Ladybug 5 camera and a Velodyne 32 scanner. Satellite or aerial imagery is used to provide broader context for data as required.

Photo © TomTom

Anditi is a Value-Added Reseller globally for TomTom MoMa data. Anditi also has relationships with other mobile LiDAR and imagery providers globally which can be commissioned to capture data as required.

Data is typically pre-collected and extracted from TomTom's archives as needed. Data is captured as part of a global capture program that results in higher class roads being typically captured every 1 to 2 years.

Data for the same section of road network has typically been captured numerous times and dates back to approximately 2015. This historical data opens up the potential for retrospective performance tracking of any attributes that can be reliably processed from the historical source data. Data can also be specifically commissioned and captured on demand if required.

The TomTom MoMa data is captured using the same capture system and specifications globally and is passed through Anditi's quality assurance checking process prior to being uploaded into Roadviewer portal for client access. Quality checking includes:

- Completeness of coverage of both LiDAR and 360-degree imagery
- Contrast and brightness of imagery.
- Relative vertical accuracy of captured data with data to have a 95% vertical spread for a feature (i.e. road surface) of 5 cm or less
- LiDAR point density at the road surface of greater than 100 pts per m^2 at a distance of 10m from the scanner perpendicular to the direction of travel or demonstrated ability to accurately generate features from fused LiDAR and 360-degree imagery.

The consistent Roadviewer QA process will be applied to other sources of mobile LiDAR and 360-degree imagery allowing for both supplier and geographic scalability as they become available/approved.

Archive MoMa data is stored in TomTom's cloud storage facility. Once a request for data is received through Roadviewer.market and access for the user authorised, requested data is quality checked and uploaded into Roadviewer portal ready for browser access by a user for subsequent coding of iRAP attributes.

The source data, after quality checking, is processed using Anditi's Roadviewer technology which converts it into iRAP attributes for subsequent viewing, checking and editing where required. Roadviewer is then used to export the iRAP coded data into a VIDA compatible format.

Artificial Intelligence and Machine Learning analysis techniques applied

For the past four years, Anditi has been researching and developing automated and accelerated techniques to utilise mobile LiDAR and 360-degree imagery for Star Rating of roads. This research is the foundation of Anditi's Roadviewer technology and has demonstrated that to make Star Rating scalable, repeatable, consistent and accurate, a range of factors including the capture equipment used, conditions at time of capture, coverage, consistency and accessibility of data need to be taken into consideration.

Following the development of the AiRAP Accreditation process as part of the iMOVE project, Anditi has subsequently received AiRAP accreditation for 34 Star Rating attributes and has developed a range of automated, accelerated and manual techniques to extract the remaining Star Rating attributes.

To generate the Star Rating attributes for this project, Anditi used TomTom's MoMa (mobile LiDAR and 360-degree imagery) data and Anditi's Roadviewer technology. The Roadviewer technology uses a range of AI and other technologies and comprises:

a. Tools for efficiently accessing and onboarding data

b. A suite of automated and accelerated tools for quality controlling the input data

c. An analytics platform that enables the road data to be analysed at two scales:

    i. As an entire section of road to extract features such as carriageways and lanes greater than 400 m in length, merge lanes, service roads, intersections, school zones etc.

    ii. In 100 m (or less if required) segments that allow detailed analysis of road severity objects and distance, line markings, rumble strips, lane and lane width, paved shoulders, streetlights, curvature, gradient, pedestrian crossings, pedestrian refuges, footpaths/sidewalks, speed management devices, channelisation, sight distance, etc

d. An AiRAP attribute extraction tool that converts the coded attribute data into a VIDA compatible format

e. A suite of automated, accelerated and manual tools for quality controlling the output data

f. iRAP accredited online Roadviewer Inspection System that enables mobile LiDAR and 360-degree imagery and coded road attributes to be viewed and checked in 100 metre increments. Roadviewer also can be used to view the entire section of road that is being assessed.

This suite of tools continues to be developed and refined and has been designed to enable road safety Star Ratings to be undertaken accurately, efficiently and at scale. They have also been designed to assist in the transition from iRAP based manual coding to AiRAP coding.

Auditing of coded data

Similar to coding data, auditing the quality of coded data requires the same manual process. To help with auditing of coded data, Anditi has developed a GIS based tool that uses a 2D representation of the AiRAP coded road information and enables the coded 2D output to be checked against available imagery and point cloud data.

Accreditation of attribute extraction from MoMa data

As the project did not include producing LiDAR data for a large network, two sample datasets were provided for quality checks (Western Sydney, NSW and Vasse Highway, WA) to validate the feature extraction method. Due to the number of attributes, quality checks of ten of the highest priority attributes was completed first, followed by the remaining 24 attributes for which accreditation was sought.

The limited sample sizes made it challenging to verify the full range of sub-attributes for the attributes being evaluated. Additional data from past assessments (2018) were requested for some attributes, however due to the time elapsed and the updates to Anditi's conversion algorithms, it did not fully align with current independent Streetview imagery used to validate it.

With the information available iRAP was confident to proceed with the accreditation subject to further data being made available for validation when available.

75 percent of travel mapping validation

TomTom provided iRAP with a 90-day license to its MultiNet-R and Traffic Density Product for this project. The objective of the task was to prove the 75% calculation methodology and involved the following components:

a. Examine datasets used in the 75% of travel calculation

b. Develop the validation framework, including the calculation methodology for roads carrying 75% of travel

c. Discuss the relationship between TomTom probe data and AADT.

In the validation process, four datasets were used:

a. TomTom Multinet-R network geometry link

b. TomTom Multinet-R network routing link

c. TomTom Traffic Density, and

d. NSW Traffic Volume Viewer.

The results show that 7.29% of the road network (9,035.45km) of the total NSW road network (123,983.6km) carries 75% of vehicular travel.

The method to identify where the 75% of travel occurs is highly repeatable and can be applied in other regions around the globe. TomTom has identified 60 countries where their existing Traffic Density Product data can be used to produce maps where 75% of vehicular travel occurs across road networks. This valuable data can support achievement of the United Nations Sustainable Development Goal to halve road deaths and injuries and help countries target action towards meeting Global Road Safety Performance Target 4 and other national road safety targets that use percentage of travel as an indicator.

This validation was done before the National Road Safety Strategy's 80% of travel at 3-Star of better target was announced. However, the method can also be easily adjusted according to different targets (e.g. 80% or 90% of travel). It is noted that alternate data sources would be needed to determine where 75% of travel occurs for pedestrians, cyclists and motorcyclists as the TomTom probe data is related to mapping products primarily used by vehicle occupants.

As TomTom only collects data from vehicles carrying TomTom devices or partnered devices the probe data is regarded as a sample out of the total population, the penetration rate may vary across time of the day and locations. To repeat the process outside of NSW, the accuracy of multipliers must be validated with local traffic data. To achieve a higher confidence level of the multiplier accuracy, it is recommended to use data sources generated from the same days in the year, and even every hour, as multipliers could vary at different period of the day, and different days in a year.

APPENDIX B

Summary of scalability review

TomTom's Multinet-R (MN-R) data

The abundance of the TomTom's Multinet-R (MN-R) data and its broad coverage across Australia represents a good starting point for large-scale automation of the current approach, saving manual effort and editing

The database is refreshed weekly and is scheduled for further updates that will include more road types making it very attractive for further investigations. Internationally the coverage is also at high standards for the top eight road classes and includes similar features.

Due to the abundance of information available within the MN-R database, there is an over-supply of features, and not all are proved to improve road safety. To extract the data relevant to road safety, and specifically Star Ratings, one needs to filter, adjust, or select what is important. The MN-R data is refreshed weekly however TomTom's MNR2iRAP FME workbench needs to be rerun each time to extract the updated dated. It is envisaged that in the future this may be undertaken routinely on a quarterly basis or similar. According to TomTom, the MN-R data coverage will extend over the next 3-4 years to multiple road types and areas.

The table below shows the findings and the confidence level of the scalability approach using MN-R data.

Table 9 Positive findings regarding to the scalability

Data source	Sample data completeness	Positive findings	Confidence of scaling to NSW	Confidence of scaling to Australia	Confidence of scaling internationally
TomTom MNR data	99%	Broad coverage across Australia and internationally.	High	High	Medium (Some Middle East and African countries have low percentage of coverage in the road network.)
TfNSW baseline map	54.4%	The 100 metre segmentation length seems appropriate given the amount of data and desired processing accuracy.	High	Medium. (All the rest of states need to provide an official baseline road network which has the same 100-metre segmentation.)	Low (to keep the 100 metre segmentation as the standard map, every country needs to adapt their road network.)
TomTom MN-R iRAP data	96%	This can be automatically processed by TomTom using the FME workbench.	High	High	High where sufficient coverage available

Data source	Sample data completeness	Positive findings	Confidence of scaling to NSW	Confidence of scaling to Australia	Confidence of scaling internationally
Anditi map matching data	95.1%	Capable of matching different segmentations of maps to the standard segmentation provided by TfNSW.	Medium. (Although this step is automatic, it can produce some mis-matching errors.) Confidence is High where the road segmentation and alignment is agreed at the start of the feature extraction process.	Medium. (Although this step is automatic, it can produce some mis-matching errors.) Confidence is High where the road segmentation and alignment is agreed at the start of the feature extraction process.	Low. (Although this step is automatic, it can produce some mis-matching errors.) Observation: According to TomTom, these issues can be removed at the onset if the entity providing the data has a clear segmented shape file or polygon to show project scope. Confidence is high where the road segmentation and alignment is agreed at the start of the feature extraction process.

TomTom's MoMa Data

TomTom's MoMa data is captured using the same equipment and data processing system across NSW, Australia and globally indicating that the results that were achieved as part of this project are scalable and can be replicated across the globe.

Summary of technical considerations

MN-R data is available for over 160 countries and MoMa data used for the iMOVE project is available for over 60 countries across the globe. Similar mobile LiDAR and 360-degree imagery data is also available from other sources globally ranging from other navigation mapping firms to local surveyors.

MN-R data and Anditi's recently accredited Roadviewer iRAP Inspection System both automatically extract curvature and gradient information for road segments.

Where LiDAR data is available, there is an opportunity for more detailed road feature evaluation than is possible through network/MN-R data alone.

The table below summarises a list of issues which may affect the scalability of the current approach, including their impact, severity and potential mitigations. Note that these findings are based on UTS observations of the project approach with ongoing research and development expected to continuously improve scalability and coverage.

Table 10 Summary of technical considerations at time of project activity

Issue	Scalability impacts	Severity of issue	The path forward
Road network segments are represented differently by data providers.	Agreement on data standards across jurisdictions may be non-trivial; translation	**Significant.** Identifying and agreeing on a representation of standards has required	For future projects it is recommended that the precise road segmentation to be used

For example, in the pilot project, there is a mismatch between 100m road segments defined by TomTom and those of TfNSW.	between data standards may introduce approximation errors; new software systems or data-processing workbenches must be developed to enable the translation of baseline roadmap definition and future road segmentation for feature selection. Addressing roundabouts and similar intersection layouts are one area for specific attention. This problem can be readily addressed by agreeing the start and finish of each of the approximately 100 m (or other agreed length) road segmentation at the start of any project. The secondary challenge is that iRAP requires road segment lengths to be 100 m. In some places this may not be consistent with the road segmentation that was previously used	multiple iterations between partners in the current project; translation processes developed for the project took additional processing time; movement to different road segmentation granularities carries processing overheads (in the current project, delivery of 10m segments was ruled out-of-scope due to processing costs). Agreeing the precise road segmentation to be used at the start of the project can minimise this issue. This may involve a 3-way agreement between the road authority/council, iRAP and those extracting and generating the iRAP attributes.	is agreed upfront and the issue will be minimised. Research can also be undertaken to identify road network representation standards utilised by road transport authorities; translation software should be developed to ensure that MN-R and jurisdictional data can be translated for use by any Australian state or territory authority. The MN-R data can be used to extract information based on boundaries of specific shape files (areas) with filters on the road types.
Oversupply of information from the MN-R database	An oversupply of information in the map data pipeline (with TomTom's iRAP features) from TomTom to Anditi requires more processing time and verification of data alignment with the original baseline map provided by TfNSW.	**Minor.** The oversupply of data can be easily filtered from the investigation as long as clear rules are defined on what should be filtered or not provided there is good alignment between the road network and road alignment as defined by the road authority and that utilised by TomTom.	This can be readily addressed at the start of the project by checking the compatibility of the road alignment and agreeing on the alignment to be used. A filtered selection should be applied from the beginning in the FME workbench to align the road network and select only those road segments/ features that align with those requested by the transport agency. This oversupply of data was resolved during the course of the project, using TomTom's road classification to remove this oversupply by 63% from the initial delivery.
LiDAR data not fully integrated for all existing road segments in NSW	LiDAR data and its feature extraction were not available to analyse for all the TfNSW	**Medium impact.** The accredited attribute data will be immediately valuable. Full automated	The project has resulted in Anditi being AiRAP accredited for 34 attributes that can be

	selection of roads for this study. This was not part of the project scope. Demonstration data provided by Anditi for 6 urban road segments covering 11.5 km in Western Sydney and 11.5 km of rural road segments in Western Australia showed very good potential for extension. MoMa data is available for 124,000 km of road in NSW and scalability is expected. In addition, mobile LiDAR data and 360-degree imagery can be sourced from a range of suppliers.	star rating will be incomplete without LiDAR feature extraction and accelerated/manual generation of remaining attributes and should be taken into account for future large-scale adoption of this method to other states or to more granular roads.	directly imported into Roadviewer. The Roadviewer tool facilitates the integration of both traditional and accelerated and intelligent data generation allowing full Star Rating datasets to be completed. The approach shows great potential for NSW LiDAR and 360-degree mapping, feature extraction and quality assurance on top of the MN-R-related features.
Compatibility or transferability between attribute definitions that were used for manual attribute coding compared to AiRAP **digital data definitions**	The transition from manual video-based coding to accelerated and intelligent coding, has required more detailed definitions to be developed for road safety attributes and categories. This has been necessary to enable automated digitally based processes to be used. This includes a range of attributes such as skid resistance, road condition, curvature and quality of curve, delineation etc. that will ultimately support greater consistency in coding.	**Minor impact** Where digital data is utilised, there is potential to review the calculations mathematically and confirm translation to iRAP attribute data is as expected. Greater consistency of coding when compared to human assessments is also expected. The quality assurance requirements for iRAP assessments should also identify and minimise any variability. Digital data can also be more easily generated for historical comparison, quality checking and trend monitoring.	Further development of digital data specifications and efforts to enhance quality assurance of digital data should ultimately increase confidence and repeatability of any coding exercise. To address comparison with historical manually collected datasets a higher degree of data matching and checking should be undertaken for the relevant attributes. The ongoing use of full digital data specifications should increase data consistency and reliability over the longer term.