AUSTRALIA'S DATA-ENABLED RESEARCH FUTURE:

TECHNOLOGY & ENGINEERING

A collaboration between ARDC, ACOLA and Australia's five Learned academies

Australian Academy of Technological Sciences & Engineering

Sponsored by

Australian Research Data Commons

About this report series

This project is the result of a partnership between the ARDC, Australia's five Learned Academies and ACOLA to ensure Australia can undertake excellent data-enabled research across all fields of research.

Notably the project sought to help build a more coherent data policy and strategic data planning environment to uplift national data infrastructure.

Five domain reports were developed and a synthesis report focused on common themes and multidisciplinary opportunities and needs.

We hope that this project will transition into an ongoing national data policy and strategic planning capability.

Report prepared by the Australian Academy of Technological Sciences & Engineering.

This project received investment from the Australian Research Data Commons (ARDC).

The ARDC is supported by the National Collaborative Research Infrastructure Strategy (NCRIS).

Date of publication
June 2022

Publisher
Australian Council of Learned Academies
9 Gordon Street
Acton ACT 2601
Australia
www.acola.org

Suggested citation
Australian Academy of Technological Sciences & Engineering (2022) *Australia's Data-Enabled Research Future: Technology & Engineering*

Australian Academy of Technological Sciences & Engineering

Contents

Acknowledgments

The Australian Academy of Technological Sciences and Engineering (ATSE) gratefully acknowledges funding provided by Australian Research Data Commons (ARDC). The ARDC is supported by the National Collaborative Research Infrastructure Strategy (NCRIS).

ATSE is very grateful for the contributions of the project Steering Committee, external reviewers and contributors. Sincere thanks to the participants in dialogues, roundtables and individual consultations conducted throughout the course of the study. The opinions expressed herein are attributed to ATSE and may not reflect the views of supporting organisations.

Project Steering Group

Professor Saeid Nahavandi FTSE
Chair of the Steering Committee
Pro Vice Chancellor (Defence Technologies)
Deakin University

Professor Shazia Sadiq FTSE
University of Queensland School of Information Technology and Electrical Engineering

Dr Susan Pond AM FTSE
Chair
New South Wales Smart Sensing Network

Dr Helen Cleugh FTSE
Atmospheric scientist
Inaugural Director
CSIRO Climate Science Centre

Francesca Ferrazza FTSE
Senior Vice President – Decarbonization & Environmental R&D
Eni
Italy

Consultation Group

James Ball
Sue Barell AO FTSE
Ian Opperman FTSE
Brett C Phillips FTSE
Tom Hatton FTSE
Jackie Craig FTSE
Michael C Frater FTSE

Project Team

Suryodeep Mondal
Peter Derbyshire
Dr Harry Rolf

Climate change is one of the most significant challenges facing the global community.

Without action to mitigate the effects of global warming, there will be even more significant impacts threatening biodiversity, ecosystem and human health, water resources, and food security that in turn have dire effects on all sectors of Australia's economy. Though the general trends and risks of a warming climate are well-known, important questions relating to food security, water resources, biodiversity, and other socio-economic issues need more research. How extreme will the weather changes be? How will these changes impact diverse ecosystems? Answering these critical questions requires new approaches that can help Australians better respond to changing climate conditions.

To mitigate the worst effects of climate change, Australia must develop resilience against climate related disasters that have already begun to threaten our built environments, economies, and health systems. We must also ensure the sustainability and security of our resources to meet domestic and global demand for water, food, energy, minerals, and technology.

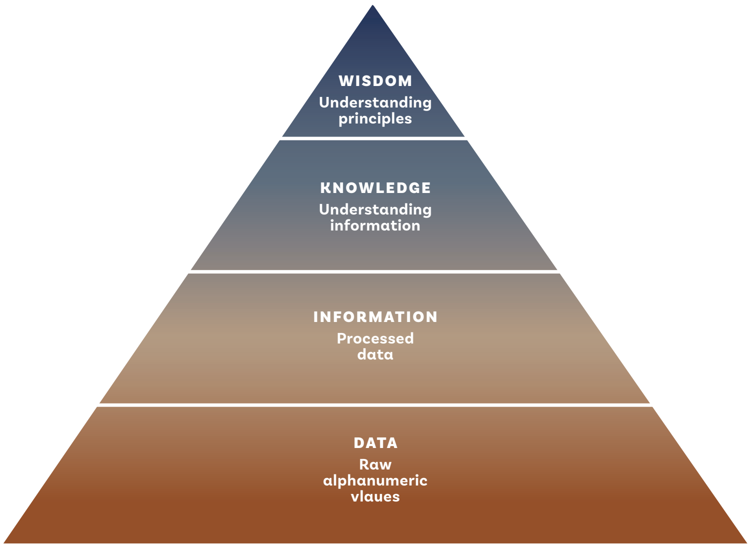

FIGURE 1
The wisdom hierarchy: representations of the DIKW hierarchy[1]

Evidence based research is the bedrock of all such activities and data plays a critical role in helping to solve these major issues. Data is the lifeblood of decision-making and the raw material for accountability in the modern world. Although data management research has a long-standing history of outstanding contributions[2], the new setting characterised by the volume, variety, and velocity of data creation and use presents new challenges in the management and effective use of data assets[3]. Under a scenario where Australia successfully achieves resilience and adapts to the impacts of climate change, what data capability for research and decision making did we need to get there? Do we have this capability now or does it need to be created?

Methodology for the Environmental Scan

The objective of the project is to identify, through a series of scenarios, the essential data needs and capabilities to reach these potential scenarios and the current gaps that need to be addressed.

The Australian Academy of Technological Sciences and Engineering (ATSE) is a Learned Academy of independent, non-political experts helping Australians understand and use technology to solve complex problems. Bringing together Australia's leading thinkers in applied science, technology, and engineering, ATSE provides impartial, practical, and evidence-based advice on how to achieve sustainable solutions and advance prosperity.

As a Learned Academy of around 900 independent experts, ATSE has conducted an environmental scan of the data assets that are in place in the three scenarios presented. The development of this environmental scan has been led by the Steering Committee comprising of six interdisciplinary Academy Fellows with expertise in data-enabled research. This was followed with three separate roundtable discussions with Fellows, that resulted in the identification of the challenges and opportunities associated with the data needs in these scenarios.

Guiding questions for the environmental scan

Understanding the question of data capability under these scenarios, the following questions were used a guide:

Q1 What are the technologies present for data collection, management, and use?

Q2 What is the current state of data infrastructure (technical and social) in the scenario?

Q3 Is there data available for discovery, understanding, and prediction?

Q4 What are the key gaps in the data needs?

1. Water resources — security and sustainability

Sustainable management of the natural environment is essential with continued environmental degradation potentially leading to worsened health outcomes across generations.

The availability of high-quality data has become a key element in guiding good decision making to help sustainably manage the natural environment. Advances in science and technology have resulted in an extensive array of powerful sensors and other observation tools that provide a wealth of data that is processed for analysis. Processing this massive amount of data requires high-performance computing. The emergence of big data has been used to discover, analyse, and understand environmental changes at micro to global scale. So far, it has revealed a stark picture on the state of our environment, but it has also supported the identification of a way forward through harnessing technologies to mitigate the challenges of climate change.

1.1. Using data to create sustainable management of water resources

Water remains one of the most important and multifaceted resources, and its sustainable management is critical. Creating water management strategies confidently with information derived from sufficiently robust data is difficult with the high temporal variability of Australian hydrology (which is often quoted as the highest in the world)[4]. Additionally, Australia has a highly complex system of water management regulation, with different responsibilities sitting across different levels of government.

Water management is difficult, and at times contentious, in an ecologically diverse continent like Australia. Australia is a world leader in water management, but it faces significant, well-documented challenges that will only worsen as the climate continues to change.

Knowing how much water is in the system, where it is, where it goes, and what value it brings is critical to water resource management. We also need to use historical data to forecast how much water will be in the system in the future. Monitoring and forecasting are becoming increasingly difficult in large, complex catchments with multiple points of managed and unmanaged losses or withdrawals, as well as competing upstream and downstream demands. Inadequate or incomplete knowledge limits our opportunities to improve system efficiency, reliability, fairness of distribution, and compliance.

The difficulties in monitoring and managing water vary depending on the goal and the landscape. Groundwater is critical to agriculture, the environment, mineral and energy resource development, and the well-being of regional communities in many parts of Australia. To properly manage groundwater, we need to better understand it. Understanding the link between groundwater and surface water systems, as well as minimising the impact of development on groundwater supply and quality, are critical to our water security and regional development. Good information derived from accurate and timely data is required to support informed decision making, but the challenges of capturing data and presenting it as useful information vary depending on the information's 'end user.'

1.1.1. Technologies for data collection:

While advances in sensor technologies have made some progress towards more efficient data collection, there is still some work to be done in terms of how emerging technologies can improve understanding of that data to create useful information for a variety of end users. Specifically, when using sensors, the metadata that would be collected by people in the field using traditional sampling methods is frequently lost. This metadata, and the context it provides for data analysis and interpretation, is critical in developing information and outputs that are appropriate for various end users. Continuous datasets are crucial to complete a comprehensive analysis.

The infrastructure for collecting water data is ageing, and the upkeep and expense of new and replacement equipment, as well as the calibration and accuracy of monitoring equipment, are important issues with current water monitoring infrastructure and technology[5]. This drawback in data collection infrastructure leads to insufficient data, compromising the derivation of insightful information from data for decision making. For example, in New South Wales there is a lack of metering data for unregulated rivers, and a lack of meteorological data to support modelling and the water monitoring network is inadequate to calculate the flow metric[6].

A key observation in many data-driven decision-making systems is that there are conflicts between the roles of data collection and data analysis. This is due to two factors: the first is vulnerability to "data selection bias," which can result in poor (irreproducible) analysis and inefficient decision-making[7]; secondly, decision-makers may be hesitant to share data with other stakeholders that would enable them to conduct similar or related analysis. Allowing others to use the same data to develop alternative hypotheses and points of view is critical for ensuring decision-making robustness and resilience.

1.1.2. Collecting Data fit for purpose:
Across the broad areas of water policy, planning, management and operations, there is a wide variety of uses for water data. The different states or territories are usually responsible for environmental assessments (e.g., State of Environment reports), however, some natural drainage lines cut across more than one jurisdiction. This creates a challenge in aligning environmental assessments by jurisdiction with specific divisions. In addition, because the drainage divisions are large, the condition and trend of the water environment and the pressures on it vary significantly within a division. Water data provides an objective basis upon which to answer important water assessment questions such as:

1. What are the demand and supply needs of water?
2. How have patterns of water use varied over time and how do they differ between regions?
3. Can we use rainfall and runoff data to model for droughts and floods, and better frame policiesfor water allocation and management?
4. Can better water data help us understand the cause of acute water shortages?
5. Can we understand and predict the frequency and severity of flooding in a region?

There are many water quality parameters that can be collected in real time, but others would require a physical collection of samples to look at dynamics, and further analysis which would be more costly to carry out.

1.1.3. Interoperability of data:
The current structure of Australian water data collection and management has each entity or organisation collecting and holding their own water data to meet their own specific needs. For example, an urban water utility may be concerned with gathering data on how water is distributed, whereas the State or Territory department in charge of water management will be concerned with the overall input and outflow of water across the region's water systems. Each entity also advances and updates their own technology systems on their own timetables, resulting in a somewhat uncoordinated technological change across the country.

Because various jurisdictions employ different data storage technologies, systems, and formats, data must be 'translated' before it can be passed on to other jurisdictions. This is typically a difficult procedure since the underlying coding or quality of the data points is not consistent – each institution or organisation gathers their own data for their own benefit – therefore personal intervention is necessary to standardise data formats and quality (data is used for different reasons, resulting from the different ways it is collected and stored). The issue of interoperability can also be compounded at times due to a lack of agreements on access arrangements.

On a national scale, the speed of technology advancement used in water management depends on the slowest adopter, which means that implementation of leading-edge technology is a barrier for the sector. Practical translation and implementation of any new technology should be equal across the jurisdictions working with each other to allow these challenges to be addressed collectively and with the end users in mind, to allow Australia's entire water sector to advance.

The Bureau of Meteorology (BoM) holds the role as the national water information provider, is a major repository for water data for research, and has the authority to issue National Water Information Standards. It holds this provision under the Water Act 2007, but it has chosen not to do so yet. Instead, it works collaboratively with the domestic water industry and international organisations to develop and promote water information standards and guidelines[8]. Private industry in Australia collects a wide variety of water data, but this data remains inaccessible for research. This water data could be very useful if a water data sharing arrangement framework was created.

1.1.4. Access to Data

Data availability and quality is a major barrier to how water data can be used. The effectiveness and efficiency of all water-related decision-making, both in government and by external stakeholders, would be substantially improved and made more resilient if water data was made 'open' – freely and publicly available to all in some format (not necessarily providing raw data). Data is often described as either being 'open', meaning it can be accessed by anyone with few (or no) restrictions, or 'closed', meaning that specific restrictions must be placed on the access to the data and the use of the data, including use of insights generated from the data. 'Open data' is globally now a commonplace policy in many situations that impact public good and that require many different experts to contribute to decision-making. Open data encourages collaboration and transparency and drives efficiency and resilience of decision-making.

1.2. Key Takeaways

Opportunity

- Utilising and deploying current technologies (especially in regional Australia) to collect more water data for analysis.
- Fit for purpose data collection design', making sure data collected is representative both spatially and relative to its environmental complexity.
- Creating a framework for sharing water data collected by industry.
- Designing modelling technology that can handle imperfect data, in real time, to provide clear information in usable outputs should be a priority for future technology development.
- Ability to collect dynamic data.
- Collecting and storing a uniform format of metadata from water data collection to help in solving the problem of interoperability.
- Developing and implementing technologies to streamline data management and transfer – including in real or short timeframes allowing water managers to efficiently analyse quality data and use that information to make the best possible evidence-based decisions.
- Archiving water data and water samples used for analysis.
- Harnessing Indigenous knowledge to collect better water data.

Challenge

- Economic cost required for the maintenance of data technologies in the field and the cost incurred to setup innovative technological advances to collect water data.
- Setting up more water data collection sites in both urban and regional areas with reliable instrumentation.
- Data security.
- Multiple technologies, systems, and formats of data storage in use make it difficult to manage and share water data for use in analysis to inform nation-wide (or multi-jurisdiction) planning and forecasting.

- Clear standards of coding of water data with the fragmented collection process and holdings.
- Knowledge continuity - understanding what to do with the vast amount of data that will be collected through different devices, setting up relevant fit for purpose parameters, interpreting historical data and dealing with previous manipulation of data.
- Ensuring that knowledge about data holdings and capabilities are shared.
- Accessibility to archived water data, metadata, and water samples.

2. Data from earth observation systems

Climate change is one of the biggest challenges facing the world. It is one of the challenges in which use of earth observations (EO) can make a difference, as EO has the capability to capture environmental and socio-economic data over a range of spatial, spectral and temporal resolutions.

EO refers to a wide range of activities that collect observations and generate measurements and spatial data to monitor and analyse our planet, its environments, human activities, and infrastructure. EO involves the collection of data about the Earth's physical, chemical, and biological systems using remote sensing technologies. This data is gathered with the help of satellites equipped with imaging devices to monitor, measure, and interpret the state of and changes in, the natural and man-made environments.

The ability to access complete and high-quality data is crucial to creating reliable predictions of extreme weather events on short and long-time scales, and to reduce potential risks and damages that result from them (Seneviratne et al., 2012[9]) (IPCC 2013[10]). Understanding, modelling, and predicting weather and climate extremes is identified as a major area necessitating further progress in climate research. It can be organised around the four overarching research phases:

1. Document (focusing on observational requirements),
2. Understand (focusing on the relative roles of different spatial scales and their interactions),
3. Simulate (focusing on model reliability and improvement), and
4. Attribute (focusing unravelling the contributors to extreme events)[11].

Through this scenario we explore sourcing and distribution of data from earth observation systems (with a focus on climate data). We explore the data policies in place for sharing and storing and the future challenges and opportunities in this space.

2.1. Using Earth Observation Satellites to collect data

EO creates an essential data value chain in the economy. The data and knowledge derived from EO helps governments and other stakeholders at regional, national and sub-national levels to respond in many areas, including mitigation, adaptation and other specific provisions that are agreed upon in international agreements (e.g., Paris agreement).

Earth observations provide near real-time data on greenhouse gas (GHG) concentrations and emissions for carbon accounting in relation to mitigation responses. When EO data is combined with other critical socioeconomic information at the local scale and over extended timescales, efforts to monitor progress on adaptation responses can all be enhanced in addition to impact, vulnerability and risk assessments and the development of measures to increase resilience.

2.1.1. Data Reception from earth observation

Data from EO satellites provide public and private actors with insights which are then applied to areas such as:

- Weather forecasting
- Wildlife conservation
- Agriculture
- Resource management
- Natural disaster response
- Climate science
- Infrastructure
- Urban planning
- Financial services.

The BoM assimilates data from over 30 satellites[12] into weather, ocean and hydrology prediction and visualisation systems every day. It is the major repository for climate data that is used for both operational and research purposes.

In Australia, most EO data is used by government (defence and civil) and researchers. Weather forecasting, disaster management, bushfire management and environmental management are examples of public sector uses of EO data. In 2014, more than 60 per cent of Australia's public Earth observation programs relied on US satellites, with around a third of Earth observation programs relying on Landsat (a joint NASA / USGS program, and the longest-running enterprise for satellite imagery of Earth) alone. At the same time, 25 per cent of Australia's public Earth observation programs relied on satellite data provided by Europe, and 11 per cent relied on satellite data provided by Japan[13]. Almost all these programs rely on open access data agreements.[14]

2.1.2. Maintenance of Increasing Data Volumes

Public organisations such as BoM, CSIRO and Geoscience Australia have a high reliance on open access data for all aspects of their EO activities. Furthermore, many government agencies require the ability to track changes over time, which is offered by the major satellite operators via open access data (e.g., Landsat).

An expected (but largely unplanned for) consequence of the increasing availability of EO under free, full, and open data licensing arrangements such as the Sentinel data access policy[15] is the volume of data that Australian government agencies will be required to store and manage. In 2011, the CEODA-Op's[16] report projected the annual storage volume required for all data types, including three levels of processing, would be approximately 1.2 PB per year by 2015. Downlinking, storing, processing, and analysing this data will be costly, and urgent action is required to put in place policies and systems for effective and efficient data management to fully realise its benefits[17]. EO storage, management, processing and analytics are an ever-increasing component of the data value chain.

Utilising data fabrics could be a potential solution to this problem. At its core, data fabric is a way to manage and utilise data that essentially means the data models would be brought to the data repository for analysis rather than the other way around. It works by the orchestration of numerous data services that support data discovery, consumption, metadata management, governance, and delivery to the required endpoint (or endpoints) in an analytics-ready format. Analytics can then be performed by artificial intelligence, or by human operators, to derive insights and inform action.

A data fabric architecture is also agnostic to data environments, data processes, data use, and geography. This would make it a fit for purpose solution in the global, open-data environments generated by EO satellites.

Ground stations are another critical asset that form the core of global satellite networks. These stations help with on-boarding satellites, scheduling satellite contact, command, control and downlink data, and receive, process and distribute data. Weather conditions (e.g., the number of clear skies exceeding 250 days in Alice Springs and remote parts of Australia) present a great advantage for ground station activity. Australia is continuing to invest in this capability, for example, Australia's first Indigenous ground station came online in 2020 based in Alice Springs[18]. The station is one of nine Earth observation ground stations managed by the Australian National Ground Segment Technical Team (ANGSTT) and is part of a global network of ground stations operated by European agency, ViaSat.

With the increasing volumes of all types of data from EO satellites, 'ground truths' of data also becomes an important issue. Ground-truthing of data allows the images data collected from EO satellites to be related to confirmed features and materials on the ground. The validation of prediction and analysis by 'ground truths' enable calibration of remote-sensing data by EO (leads to getting better spectral data), and aids in the interpretation and analysis of what is being studied.

2.1.3. Data Policies

Data Sharing is a pre-requisite for building an effective Global EO system of systems. Presently, Australia's EO data supply is entirely derived from internationally owned (public or private) satellite infrastructure. Except for commercial organisations, all satellite data used by the EO industry is obtained from international partners. To secure satellite data, Australia has several individual partnerships with international space agencies. Australia is a member of three international Earth observation organisations: the Group on Earth Observations (GEO), the Committee on Earth Observation Satellites (CEOS), and the World Meteorological Organization (WMO).

Ownership and rights that are attached to data are complex. A major international policy that allows the international exchange of Earth system data is the WMO unified policy for the international exchange of earth system data. It is the apex global body for collecting climate data through earth observation systems, it is a United Nations body with countries being its members (Australia is represented in this through BoM and CSIRO). The policy agrees that the WMO will '*commit itself to broadening and enhancing the free and unrestricted (free and unrestricted means available for use, re-use and sharing without charge and with no conditions on use) international exchange of Earth system data*'. This policy is a step away from the earlier iteration which involved only sharing the data to the various meteorological agencies of the world . This policy makes data availability uniform to both public and private entities and lays a great precedent for better data sharing.

GEO is another international earth observation organisation that recognizes that the societal benefits arising from Earth observations can only be fully achieved through the sharing of data, information, knowledge, products and services. GEO Principals endorsed a new set of Data Sharing Principles, which promote 'Open Data by Default', in Mexico City at the dawn of the second decade of GEO (2016-2025)[20]. These policies reinforce the key role of data to inform decision-making in addressing environmental challenges.

2.2. Key Takeaways

Opportunity

- Open data policies and international data sharing agreements.
- Utilise data fabric architectures to handle the increasing data complexity and volume.
- Create a common standard for metadata collection.
- Ground truthing of data to validate data collected by EO satellites, ensuring its indicative of the reality on the ground and is appropriate for subsequent modelling and analysis.
- Better historical data availability to enable the longitudinal evaluation of climate change, and application of new and emerging modelling to historical data sets. This would enable enhanced predictive capabilities.

Challenge

- Creating a standardised metadata collection and reaching international consensus on the type of data that is made available.
- Perception that the implementation of the Open Data policy could pose challenges to the development for data providers, resulting in limited revenue.
- A significant investment would be required to establish earth observation systems; however, it would become an essential piece of Australia's data infrastructure.
- Availability of ground-truths data when data is collected from foreign sources.
- Life-cycle data management with increasing volumes of data.
- Creating better access to the data already collected and available for sharing.

3. Infrastructure Resilience

Australia's physical infrastructure needs to be prepared and capable of adapting to shocks (climate-driven threats). Given the hyper-connectivity of core infrastructure assets, as well as the impact of digital transformation, data plays a critical role in decision making to strengthen critical infrastructure resilience. Data is at the forefront of reducing the risk of disruptions in these critical services and increasing the capacity to quickly recover after a shock.

Within the Sendai Framework for Disaster Risk Reduction 2015-2030, the United Nations defines resilience as:

> *"The ability of a system, community or society exposed to hazards to resist, absorb, accommodate, adapt to, transform and recover from the effects of a hazard in a timely and efficient manner, including through the preservation and restoration of its essential basic structures and functions through risk management."*

For infrastructure to be resilient it should be able to withstand disruption, operate in crisis, and deal with and adapt to shocks and stresses.

Access to real time data is essential to resilience. The Minderoo Foundation initiative has emphasised the importance of early, accurate, and monitored real-time knowledge in any natural disaster to ensure an appropriate response, including through their aim statement "We Rise Together – Lifting Australia to be the Global Leader in Fire and Flood Resilience by 2025" . Communication infrastructure is critical to ensure continuity of service in an emergency, but it has a high proclivity to fail. A quantitative assessment of infrastructure resilience using data acquisition and artificial intelligence has the potential to achieve a rapid response to emergency response based on real-time knowledge.

Resilience and risk are fundamental concepts for critical infrastructure protection, but it is complex to assess them. If we model critical infrastructure interdependency, that tells us about the resilience and risk associated with the infrastructure network, but not necessarily the individual asset. The operation of critical infrastructure represents a significant proportion of economic activity. The challenge is determining where and how resilience integrates into risk assessment, as risk is a feature of threats and hazards, weaknesses, and consequences.

3.1. Creating Resilient Infrastructure Networks

Australia has seen a 40% increase in bushfire frequency between 2011 and 2016. The 2019-20 bushfires caused up to $100 billion in damage, including $5 billion worth of road damage[22]. Rainfall is increasingly volatile, with extreme drought and flooding in parts of the country[23]. By 2050, the annual economic cost of natural disasters in Australia is expected to more than double – from an average of $18 billion per year to more than $39 billion per year. In New South Wales, the expected total economic costs of natural disasters are projected to increase from $5.1 billion in 2020-21 to between $15.8 billion and $17.2 billion (real 2019-20 dollars) per year by 2061[24].

The relevance of resilience has been underlined in numerous reports including the recently released 2021 Australian Infrastructure Plan, Protecting Critical Infrastructure, National Climate Resilience and Adaptation Strategy, and the UNEP Adaptation Gap Report 2021. These plans all link sustainability to resilience.

3.1.1. Probabilistic assessment

Increasing extreme weather events and a changing climate represent a risk to major network infrastructure. The probability of an extreme event occurring can be better assessed through accurate and complete climate data as discussed in the previous scenario. This climate data gives us an essential source to build infrastructure that can stand resilient to all these extreme events. In addition, the global shift to a low carbon future will impact infrastructure design, delivery, and the supply chain. This will result in increasing costs and will require a paradigm shift in infrastructure planning. Therefore, our deterministic approaches need to give way to probabilistic assessment

What is Probabilistic Modelling?
Probabilistic (stochastic) assessment is used to simulate uncertainty and partial knowledge. It offers a means for measure of uncertainty that may be estimated using field, laboratory, or historical data, as well as complex computer simulation models giving a better understanding of the real process or systems.

While deterministic approaches can assist in determining past or current vulnerabilities and resiliency, it has very little predictive capability if the network or system changes. It also fails to accommodate increases in demand and shifting community demographics, or if infrastructure degrades over time, or if the likelihood or severity of natural disasters increases due to climate change.

As a result, probabilistic assessment carried out with the help of complex computer simulation models reinforced by real-time and other performance data can anticipate how infrastructure will function in the future because of these time-dependent or geographical changes in hazard, vulnerability, or resiliency. These stochastic approaches quantify risk, such as the likelihood and magnitude of severe occurrences, infrastructure damage, and recovery time. This risk-informed decision-making methodology is known as a Probabilistic Risk Assessment, Quantified Risk Assessment, or Probabilistic Systems Modelling.

Why would Probabilistic Modelling be required?
Probabilistic modelling of infrastructure helps in understanding how resilient current infrastructure setups are as well as give us a better understanding to create more resilient infrastructure for the future. Some of the outcomes of probabilistic risk analysis also include:

a. Calculating the likelihood and extent of infrastructure damage and potential loss to the owner, users, community, and other stakeholders, including the loss of life.
b. Influence that infrastructure resiliency has on the time to renewal and follow-on consequences and losses.
c. Calculating the effect risk mitigating measures have on predicted damage and losses.
d. Getting a more accurate cost estimate.

Every dollar spent reducing risk prior to an event replaces $4 that would need to be spent on recovery and response[25]. This provides a major opportunity to act on the paradigm shift to probabilistic assessment.

What are the data challenges of Probabilistic Modelling?
Obtaining data to determine the uncertainty is usually difficult to get hold of because:

1. The data has never been collected in the past.
2. The data is too expensive to obtain.
3. Past data is no longer relevant; and/or
4. The data is sparse/infrequent (i.e., not continuous)

The lack of data means that subjective estimates must be made regarding the uncertainty of the variables within the model. If there are risk events in a model, the probability of each one occurring will have to be assessed[26].

3.1.1.1. Transport Networks
Typically, transport networks are considered major contributors to climate change. However, these networks are also vulnerable to extreme weather caused by climate change, for example many road networks and bus terminals are in flood zones. Extreme weather increases the probability of a failure of these networks.

If we take aviation as an example, we can see it is probably the most affected by climate change in the transport sector. For a start, no one likes to be on a turbulent flight, and climate change increases unstable weather. Large airports are typically built where there are favourable wind conditions. With hailstorms, floods and damaging winds all increasing with climate change, airports are extremely exposed and are already experiencing severe disruptions.

With rising temperatures some large aircraft at airports can no longer take off[27]. The aviation industry is of course aware of this and has not only started to reduce its carbon footprint but is beginning to implement countermeasures such as new airports being built well above current sea levels.

Regional and remote airport infrastructure is crucial to both bushfire response and disaster relief operations. Most people associate air travel with tourism, however, its role in commerce and trade is preeminent: airports are the backbone of global freight and supply chains, any further change in climate will not only severely disrupt the aviation industry but also the global economy.

A probability-based safety assessment of these transport networks can help better understand the issues with the current infrastructure. Carrying out these assessments now and in the future would help come up with solutions to mitigate these problems and in turn plan the upcoming infrastructure to be more resilient. For example, a probabilistic assessment of existing bridges may be used for bridges that have failed a traditional deterministic assessment. This is a higher tier assessment including variables like weigh-in-motion data to develop probabilistic traffic load models. This model includes variables like vehicle mass, flow rates and proportions, vehicles transverse location, lane importance factors, vehicle impact factors, and model uncertainty. This bridge-specific stochastic modelling avoids the conservatism present in deterministic assessment codes.

A structural reliability analysis of this nature that also considers the variability and uncertainty in 'bridge capacity' was used to estimate the probability of bridge failure. Compared with acceptable levels of structural safety in Australia, it found that about 75% of bridges that failed a deterministic assessment were shown to have a reliability higher than the acceptable level of safety – leading to the conclusion that "probability-based bridge assessment frameworks can provide more fit-for-purpose assessments, better utilising bridge assets, and thereby facilitating an optimally productive road transport network"[28]. There is a need for future studies to consider the broader societal viewpoint of bridge damage, including the need to look at the likelihood, duration, and consequences of impact on a community or industry.

3.1.2. Interoperability of Data

The increasing intensity and frequency of adverse weather events due to a changing climate poses an episodic but extreme risk to developing and existing infrastructure. From an infrastructure perspective, the focus needs to be on adaptation and resilience building. Climate resilient infrastructure can be built by incorporating various climate/environmental data in developing structural adaptations, such as changing the composition of road surfaces, so they do not deform in high temperatures. However, the open access to datasets and general availability required to provide insights to this issue are not well established. There is need for standardised, centralised, and transparent data to support planning of these infrastructure projects and ongoing management of these assets. Standardised information on shocks and stresses and consistent data gathering and organising methods can improve data sharing and enable informed decisions.

Nationally consistent, transparent, and comparable data sets require data standards that enhance the capacity to collect and share data[29]. Only with agreed definitions, methodologies and standards is it possible to implement a national data management system of value to infrastructure decision makers and Australian communities. To achieve this, states and territories should adopt a single system with uniform regulation and protocols, and the associated appetite to carry out data reform.

3.2. Key Takeaways

Opportunity

- Collecting data on mitigation and preparedness measures, as well as performance, for past infrastructure projects to provide lessons for future projects.
- Standardising and publishing national climate data and scenarios, including hazard maps and climate risk data.
- Enhancing infrastructure assessment modelling by using probabilistic assessments to provide a more accurate and robust snapshot of future asset performance.

Challenge

- The availability of risk data across the infrastructure pipeline is a major constraint to creating resilient infrastructure setups across sectors and projects.
- Data on failure of infrastructure.
- Nationally significant risk data sets are not federated, with each state and territory having a separate data standard.
- Getting data that is accessible, clear, and fit for purpose to carry out modelling exercises to create resilient infrastructure is cumbersome.

4. Synthesis of findings

Opportunities

- Better calibration of an ageing sensor equipment network or other new monitoring equipment.
- Collecting and storing a uniform format of metadata from data collection to help in solving the problem of interoperability.
- Open data policies and international data sharing agreements.
- Creating a framework for sharing data collected by industry to be utilised for research.

Challenges

- Creating simplified access to the data already collected and available for sharing.
- Getting data that is accessible, clear, and fit for purpose to carry out research.
- Life-cycle data management with increasing volumes of data.
- Economic cost required for the maintenance of data technologies in the field and the cost incurred to setup innovative technological advances to collect data

5. Endnotes

1. Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. Journal of Information Science, 33(2), 163–180. https://doi.org/10.1177/0165551506070706

2. Abadi, D., Ailamaki, A., Andersen, D. G., Bailis, P., Balazinska, M., Bernstein, P. A., Boncz, P. A., Chaudhuri, S., Cheung, A., Doan, A., Dong, L., Franklin, M. J., Freire, J., Halevy, A. Y., Hellerstein, J. M., Idreos, S., Kossmann, D., Kraska, T., Krishnamurthy, S., Markl, V., Melnik, S., Milo, T., Mohan, C., Neumann, T., Ooi, B. C., Ozcan, F., Patel, J., Pavlo, A., Popa, R. A., Ramakrishnan, R., Ré, C., Stonebraker, M., & Suciu, D. The Seattle Report on Database Research. SIGMOD Rec., 48(4):44–53, 2019.

3. Sadiq, S. et al. Information Resilience: the nexus of responsible and agile approaches to information use. The VLDB Journal (2022). https://doi.org/10.1007/s00778-021-00720-2

4. Chiew F and Prosser I (2011) Water and climate. In: Water – Science and Solutions for Australia (Editor: I Prosser), CSIRO Publishing, (ISBN 9780643100527), pp. 29–46.

5. Malone D., Torrible L., Hayes J., 2009, NSW strategic water information and monitoring plan: final report, NSW Office of Water, Sydney. Report available at: https://www.industry.nsw.gov.au/__data/assets/pdf_file/0005/153806/NSW-strategic-water-information-and-monitoring-plan-final-report.pdf

6. Chief Scientist & Engineer, 2020, Review of water-related data collections, data infrastructure and capabilities, Sydney. Report available at: https://www.chiefscientist.nsw.gov.au/__data/assets/pdf_file/0003/337503/Water-Data-Review.pdf

7. Stagge, J., Rosenberg, D., Abdallah, A. et al. Assessing data availability and research reproducibility in hydrology and water resources. Sci Data 6, 190030 (2019). https://doi.org/10.1038/sdata.2019.30

8. Bureau of Meteorology, 2016, National Industry Guideline for water quality metadata. Report available at: http://www.bom.gov.au/water/standards/documents/NI_GL_101.00-2016.pdf

9. S.I. Seneviratne, N. Nicholls, D. Easterling, C.M. Goodess, S. Kanae, J. Kossin, Y. Luo, J. Marengo, K. McInnes, M. Rahimi. "Changes in Climate Extremes and Their Impacts on the Natural Physical Environment, Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation" (2012), pp. 109-230

10. Climate change 2013: the physical science basis. T.F. Stocker, D. Qin, G.K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, P.M. Midgley (Eds.), Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA (2013), p. 1535

11. Xuebin Zhang, Gabriele Hegerl, Sonia Seneviratne, Ronald Stewart, Francis Zwiers, Lisa Alexander, World Climate Research Programme (WCRP) Grand Challenges: Understanding and Predicting Weather and Climate Extremes. Available at: https://www.wcrp-climate.org/images/documents/grand_challenges/GC_Extremes_v2.pdf

12. Analysis and Prediction Operations Bulletins (bom.gov.au) |http://www.bom.gov.au/australia/charts/bulletins/nmoc_bulletin.shtml

13. Bureau of Meteorology, January 2021, Submission to the Standing Committee on Industry, Innovation, Science and Resources, Australian Government.

14. Australian Space Agency 2021, Economic snapshot of the Australian space sector: 2016-17 to 2018-19, Australian Government.

15. Sentinel data access: https://sentinel.esa.int/web/sentinel/sentinel-data-access

16. CSIRO (2012). Continuity of Earth Observation Data for Australia: Research and Development Dependencies to 2020 (2012, CEODA-R&D) G. Australia. Canberra.

17. Symbios Communications (2015). Risks of Data Supply of Earth Observations from Space for Australia. Available at: https://www.crcsi.com.au/assets/Resources/CRCSI-Risks-of-Data-Supply-of-Earth-Observations-from-Space-for-Australia-Final-web.pdf

18. Media Release: Space Australia, 'Indigenous Ground Station Goes Live', 8 July 2020. Available: https://www.spatialsource.com.au/indigenous-satellite-ground-station-is-online/#:~:text=The%20first%20Australian%20Indigenous,owned%20land%20%E2%80%94%20an%20Australian%20first.

19. World Meteorological Organization (Wmo) Unified Policy For The International Exchange Of Earth System Data, 2021. Available at: https://ane4bf-datap1.s3-eu-west-1.amazonaws.com/wmocms/s3fs-public/ckeditor/files/Cg-Ext2021-d04-1-WMO-UNIFIED-POLICY-FOR-THE-INTERNATIONAL-approved_en_0.pdf?4pv38FtU6R4fDNtwqOxjBCndLIfntWeR

20. Group on Earth Observations (GEO), GEO Strategic Plan 2016-2025: Implementing GEOSS. Available at: https://earthobservations.org/documents/open_eo_data/GEO_Strategic_Plan_2016_2025_Implementing_GEOSS.pdf

21. https://cdn.minderoo.org/content/uploads/2020/09/14223039/MFFFR-Blueprint-V1-200915.pdf

22. Dutta R., Das A. and Aryal J., 2016. Big data integration shows Australian bush-fire frequency is increasing significantly. NCBI. 3(2). Available at: 10.1098/rsos.150241

23. Bureau of Meteorology, State of the Climate 2020. Report available at: http://www.bom.gov.au/state-of-the-climate/australias-changing-climate.shtml

24. A Pathway to Infrastructure Resilience, Infrastructure Australia (August 2021). Available at: https://www.infrastructureaustralia.gov.au/sites/default/files/2021-08/Advisory%20Paper%201%20-%20A%20pathway%20to%20Infrastructure%20Resilience%20FINAL.pdf

25. NIBS (National Institute of Building Sciences), 2017, Natural Hazard Mitigation Saves: 2017 Interim Report, Washington DC, USA, p344

26. Probabilistic Contingency Estimation, Department of Infrastructure, Regional Development and Cities, 2018. Report Available at: https://investment.infrastructure.gov.au/files/cost_estimation_guidance/Guidance-Note-3A-Version-1.0.pdf

27. Airplanes may struggle to take off and fly in extreme heat because hot air is less dense than cool air. As the temperature increases, the air expands, making it less dense. In turn, airplanes produce less lift — even when using the same amount of propulsion created by engines — than when flying in cool, denser air.

28. Bridge Assessment Beyond the AS5100 Deterministic Methodology. M.M. Melhem, C.C. Caprani, M.G. Stewart and S. Zhang, Research Report AP-R61720, Austroads,Australia,2020. https://austroads.com.au/__data/assets/pdf_file/0030/343947/WEB-R617-20_Bridge-Assessment-Beyond-the-AS5100-Deterministic-Methodology.pdf

29. Royal Commission into National Natural Disaster Arrangements 2020, Royal Commission into National Natural Disaster Arrangements Report, Australian Government, p. 111, available via: https://naturaldisaster.royalcommission.gov.au/publications/royal-commission-national-naturaldisaster-arrangements-report

APPENDIX

Glossary of Terms

National data infrastructure
Nationally significant data assets, facilities, and services that are scalable and collectively managed and operated for use by research institutions, public and private users across the country.

Data assets
Nationally strategic databases, data systems, web-based resources, or data services.

Data policies
Principles that describe the rules to control the integrity, security, quality, and usage of data at a national level.

Data skills
Expertise and qualifications to analyse, manage, store, and distribute data.

Critical Infrastructure
The Australian Government defines critical infrastructure as 'those physical facilities, supply chains, information technologies and communication networks which, if destroyed, degraded or rendered unavailable for an extended period, would significantly impact the social or economic well-being of the nation or affect Australia's ability to conduct national defense and ensure national security.'

Scenario
A scenario is a coherent, internally consistent, and plausible description of a possible future state of the world.

Shocks
Shocks are sudden, sharp events that have the potential to disrupt the services supplied via infrastructure.

Stresses
Stresses are longer-term, chronic conditions that impact physical assets, organisations or communities. Stresses also include the increasing interdependencies between critical infrastructure that can exacerbate the impact of shock events.

Metadata
The ABS defines Metadata as data about data or information about data.

Spectral resolution data
The sensitivity of a sensor to respond to a specific frequency range, often includes visible light and infra-red (this is only relevant to optical sensors).

Australian Academy of Technological Sciences & Engineering