SAFE

A GUIDE TO A SHARED ANALYTIC FRAMEWORK FOR THE ENVIRONMENT

The Western Australian Biodiversity Science Institute

Western Australian Marine Science Institution

The Western Australian Biodiversity Science Institute

ACKNOWLEDGEMENTS:

We acknowledge our many partners across industry, government and research for their ongoing contribution and support. Our thanks to Dr Greg Terrill, Rob Freeth and Chris Gentle in the development of the framework highlighted in this guide.

Prepared by The Western Australian Biodiversity Science Institute:

ISBN 978-0-646-83607-2

Images courtesy: DBCA and Lochman Transparencies (inset)

Cover images courtesy: WAMSI, Megan Hele and Lochman Transparencies

CONTENTS

Background and Aim 4

What is SAFE? 5

Environmental Assessments: a vision 6

SAFE Envisioned 8

SAFE Explored 10

Images above courtesy: Megan Hele

BACKGROUND AND AIM

The Western Australian Biodiversity Science Institute (WABSI) and the Western Australian Marine Science Institution (WAMSI) have been working together to enhance access, aggregation, interpretation and management of biodiversity information collected in Western Australia.

In 2019, in conjunction with a working group and advisory committee, WABSI and WAMSI published the report *Digitally Transforming Environmental Assessment: Leveraging information to streamline environmental assessment and approvals*,[1] which made the following recommendation:

> ***"A shared analytic framework for the environment: Develop digital analytic tools to assist environmental impact assessment, including identifying trends and predicting impacts of multiple activities in a region over time, to improve confidence in decisions made and to reduce the need to rely on the precautionary principle."***

WABSI and WAMSI are now developing a strategy to implement this recommendation and this guide has been developed as a basis for discussion and collaborative refinement. It outlines the rationale for and approach to a shared analytic framework for the environment. The next stage will include the presentation of case studies and a roadmap to be published in a second report, Dynamically Transforming Environmental Assessments, due for release in 2021.

We welcome your feedback on this guide. Please contact:
chris.gentle@wabsi.org.au

[1] Refer to the report *Digitally Transforming Environmental Assessment*. ISBN 978-0-646-81505-3.

Image courtesy: Andrew Halsall, WA Museum

WHAT IS SAFE?

A shared analytic framework for the environment (SAFE) depicts the capabilities – the building blocks – which work together across the information and analytic supply chain to provide input decision-support and reporting tools for environmental assessments. It is a management tool, providing a framework and language to:

- Facilitate a consistent view of the capabilities and their interdependencies;
- Help align effort and prioritise investment across these capabilities.

SAFE has been developed to accelerate the move to devolved robust, repeatable and transparent decision making for environmental assessments. This will:

- Reduce risk for investors, as they will be better able to understand the impact of, and to develop mitigation strategies for, activities that they propose to undertake;
- Remove duplication between regulators at different levels of government;
- Provide public reassurance about the quality of decisions.

SAFE can help individual projects determine the capabilities that they need, as well as help prioritise effort across the information and analytic supply chain that supports national decision making.

SAFE has been developed by WABSI, WAMSI and many others. It is based upon the Global Biodiversity Information (GBIO) Outlook[2].

[2] Based on the report *Delivering Biodiversity Knowledge in the Information Age*. Available at: https://doi.org/10.15468/6jxa-yb44.

ENVIRONMENTAL ASSESSMENTS:

A VISION

A major review of national environmental legislation, the *Report of the Independent Review of the EPBC Act*[3], has proposed a vision of improved environmental outcomes, combined with transparency around faster and lower-cost decisions. Elements include:

- **Single-touch environmental approvals**
  - Underpinned by legally enforceable National Environmental Standards and subject to rigorous, transparent oversight (Recommendation 14).
- **National Environmental Standards**
  - National Environmental Standards that would be legally enforceable, and evolve (Recommendations 3, 4).
  - Including "A Standard for Data and Information to set clear requirements for providing best available evidence" (Recommendation 31).
- **National supply chain of information**
  - "A national supply chain of information will deliver the right information at the right time to those who need it. This supply chain should be an easily accessible, authoritative source that the public, proponents and governments can rely on. A clear strategy to deliver an efficient supply chain is needed so that each investment made contributes to building and improving the system" (Recommendation 32, p22).
- **Improvements to the environmental information system**, including:
  - An interim supply chain Custodian to oversee the improvements to information and data;
  - A set of national environment information assets to ensure essential information streams are available; and
  - Expanding existing work with jurisdictions on the digital transformation of environmental assessments and ensuring it is aligned with implementation of the national environmental information supply chain.

Proponents, regulators, the technical community and the broader public do not generally have an agreed touchstone of information and analyses informing a given proposal. This can become the basis of appeal or challenge, and under any realisation of delegated approvals, becomes crucial to Commonwealth assurance of state decisions.

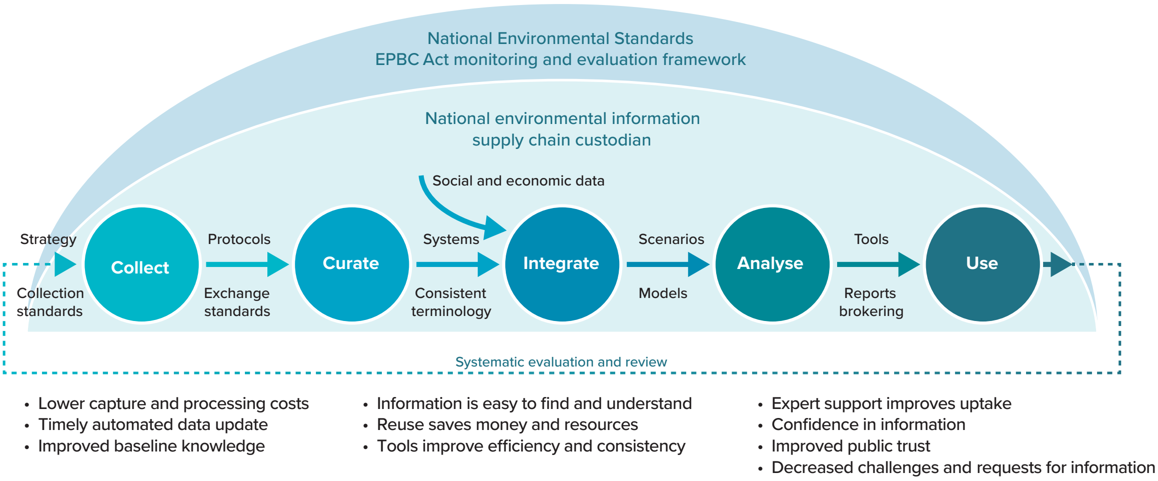
SAFE can help overcome fragmentation and deliver the information infrastructure needed to enable transparent, trusted and devolved decision making. It does this by providing a framework and language to facilitate a consistent view of the capabilities and their interdependencies, and to help align effort and prioritise investment across these capabilities.

[3] Samuel, G 2020, *Independent Review of the EPBC Act—Final Report*, Department of Agriculture, Water and the Environment, Canberra, October 2020. CC BY 4.0

The value of having a shared analytic framework includes:

- **Flexibility:** different organisations and researchers can work within their own area of expertise while contributing to a larger effort to support information and analytic supply chain.
  - The current approach includes many independent actors providing solutions that do not readily integrate and with low collective improvement.
- **Reuse:** Analytics tools developed for one purpose can readily be re-used.
  - Tools developed for decision support can often be reused for reporting or to guide environmental management.
- **Progressive improvement:** once solutions are in place, they can be incrementally improved through a consistent and managed framework as new knowledge and methods become available.

The value of SAFE is illustrated by the future state depiction contained in the *Final Independent Review Report*.



Future state of the national environmental information supply chain[3]

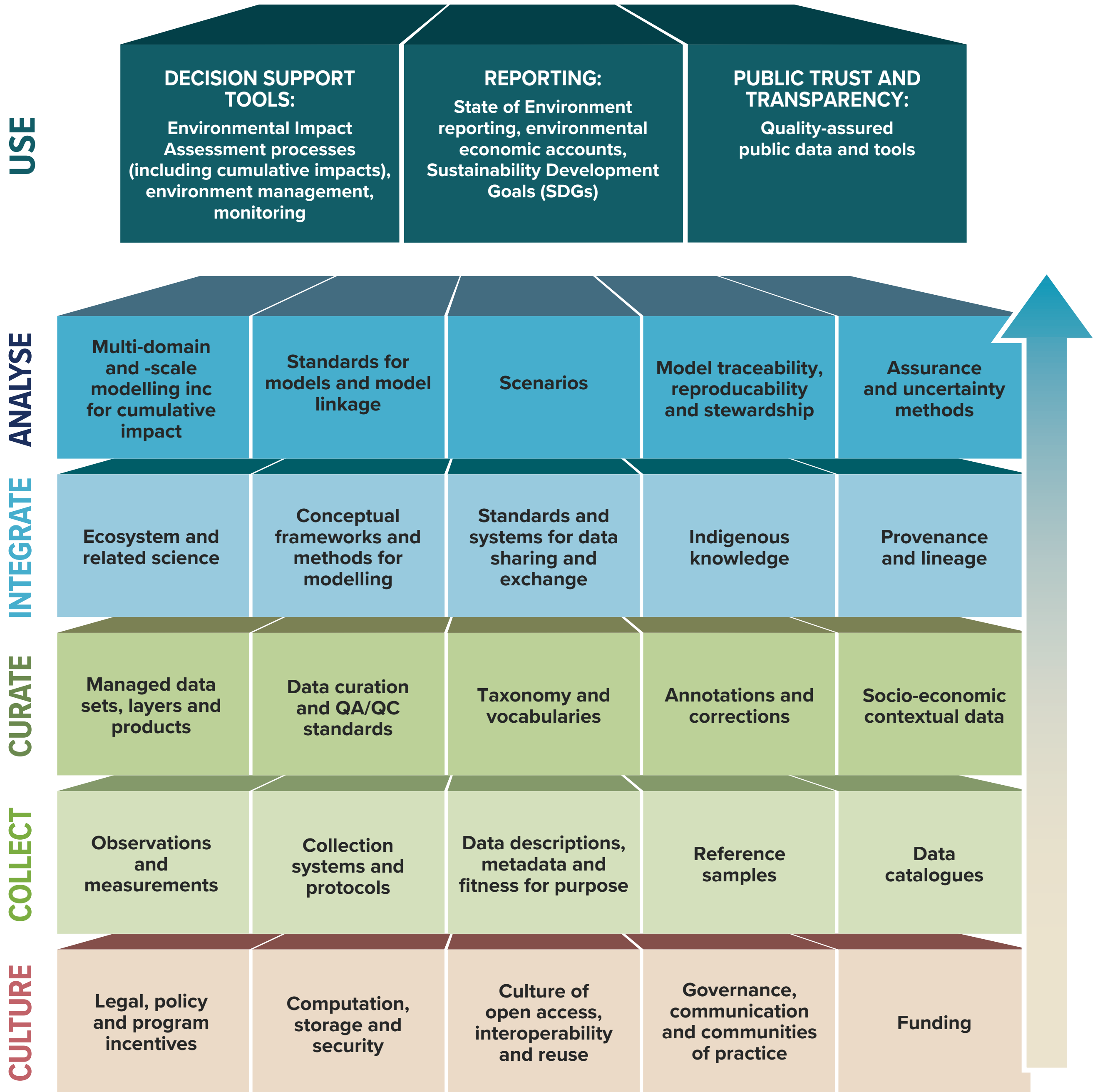
Aligned capabilities could support, for any region in Australia, a collection of knowledge products supporting environmental decision-making and reporting. The knowledge products would take cumulative impacts and offsets into account, and be linked across regions to enable:

- Rapid assessment of the current state of the regional environment and likely future states under various stressors and scenarios;
  - Automated or digitally assisted preliminary assessment of proposals presenting low environmental risk;
- Faster assessment of impacts of an activity on a region, taking into account cumulative impacts, identification of offsets, monitoring and reporting;
- Easier and more consistent reporting of the state of the environment, environmental economic accounts, sustainable development goals, and more;
- Improved public trust and transparency through public visibility of the above tools; and
- Collective learning and continuous improvement.

SAFE ENVISIONED

SAFE has five tiers, each of which describes key capabilities that support decision-making and reporting for environmental assessments. Each tier has several core components, and all tiers interconnect and add value to each other. The tiers and components are conceptual, and in many cases a particular organisation may provide services across more than one tier. At this stage organisations are not individually identified in the framework, although illustrative examples are provided later in the document.

The diagram illustrates the framework, showing the tiers and the capabilities within each tier:



SAFE EXPLORED

The following section outlines each layer and capability in more detail, providing some illustrative institutional examples.

CULTURE

The Culture tier comprises the fundamental approaches and capabilities needed to enable all elements of SAFE to interact effectively.

LEGAL, POLICY AND PROGRAM INCENTIVES

Australia has some world-leading institutions for data collection and curation, for integration and for modelling. While sharing is frequent and widespread, a culture of sharing is not universal. There remain concerns that sharing data may lead to unwelcome scrutiny, lack of data control and confidentiality breaches. Policy, legislative and financial frameworks often include requirements to share, although these do not fully align across an institutional landscape with many players.

Public policies, legislation and funding initiatives need to reinforce open access, interoperability and reuse of data and knowledge products, as well as the application of standards to support interoperability and reuse. This applies to the system, as well as within individual programs, capabilities and projects. The challenges extend beyond legal frameworks to include culture.

COMPUTATION, STORAGE AND SECURITY

Security is a challenge with large scale federated systems. Ensuring that data which need to be kept secure remains so is a critical requirement to support decision-making and reporting for environmental assessments.

Long term storage for assurance and reuse is a common problem with data and data products developed by shorter term activities (eg surveys for environmental assessments, or research projects), or where incentives are weak. Too often, data and model resources disappear, go offline or change protocols, making any systems built on them unreliable and costly to maintain. Australia has national environmental data repositories, but they are not capable of storing all of the large data streams now available. There are not yet comparable national repositories for model code and outputs, though private and open source options exist.

CULTURE OF OPEN ACCESS, INTEROPERABILITY AND REUSE

One commonly used framework for data is the FAIR principles – data should be Findable, Accessible, Interoperable and Reusable. Data are often expensive to collect and curate, and this cost should be incurred only once, while the benefits are derived many times.

Open access is best practice, with some limited exceptions to protect sensitive data or knowledge products. Licencing arrangements, attribution protocols and authentication controls, which enable tracing of what a user has viewed, altered or copied, can reduce the risks associated with open access.

Interoperability requires participants to adopt agreed approaches and standards so that data, models or knowledge products can readily be used by other systems and people. This is particularly important when integrating elements from a range of sources, as is necessary in Australia's current fragmented information landscape.

Reuse is a transparency and efficiency principle. Ensuring that reuse can occur is both cultural and technical, and dependent upon clear standards and frameworks at all levels.

GOVERNANCE, COMMUNICATION AND COMMUNITIES OF PRACTICE

SAFE encompasses a large range of capabilities and institutions, both public and private. No single governance approach exists to cover all aspects; a network of governance approaches is therefore needed. Governance, as a supporting capability, will apply within each tier of SAFE (eg various naming standards and definitions related to data), and across the SAFE tiers (eg routine means for data and data products to link to models, as well as for models to link to each other). Governance needs to incorporate ethics, privacy and consent considerations.

Communication is a dimension of almost all components. From the user's perspective, it is critical to have confidence in the knowledge products being used to support decisions. This involves communication about the data, assumptions and methods used to generate a result, as well as the uncertainty that surrounds it. This involves communication products, as well as provenance and assurance processes.

Communities of practice exist within individual capabilities and tiers, as well as across tiers. Some communities are mature (eg for many data capabilities), while others are evolving (eg to assure decision support tools, or multi domain model development and integration).

FUNDING

While there are many existing funding sources for individual elements of SAFE, it will be necessary to work together to achieve the overall potential of the system. A key argument for SAFE is that it should be possible to gain significantly greater benefit from existing funding using shared approaches.

FURTHER SOURCES: CULTURE

LEGAL, POLICY AND PROGRAM INCENTIVES:

EPBC Act – www.environment.gov.au/epbc

National Collaborative Research Infrastructure Strategy – www.education.gov.au/national-collaborative-research-infrastructure-strategy-ncris

National Environmental Science Program – www.environment.gov.au/science/nesp

COMPUTATION, STORAGE AND SECURITY:

National Computational Infrastructure – nci.org.au

Pawsey Supercomputing Centre – pawsey.org.au

Australian Signals Directorate – www.cyber.gov.au/acsc/view-all-content/ism

CULTURE OF OPEN ACCESS, INTEROPERABILITY AND REUSE:

FAIR principles – ardc.edu.au/resources/working-with-data/fair-data/

National Data Commissioner – www.datacommissioner.gov.au

Digital Transformation Agency – www.dta.gov.au

Australian Research Data Commons – www.ardc.edu.au

GOVERNANCE AND COMMUNITIES OF PRACTICE:

Australian Research Council – www.arc.gov.au/policies-strategies/strategy/research-data-management

Data Governance Australia – www.datagovernanceaus.com.au

Data Management Association Australia – www.dama.org.au

Maiam nayri Wingara Aboriginal and Torres Strait Islander Data Sovereignty Collective – www.maiamnayriwingara.org

Image courtesy: Bayden Smith, Greening Australia

COLLECT

The Collect tier includes the capabilities to generate multiple types of data, from existing sources to new fieldwork observations and automated sensors.

OBSERVATIONS AND MEASUREMENTS

This capability includes the primary data from observations made on the physical world. Often these are outputs from the many providers of ecosystem data, including:

- Marine and coastal
- Geospatial, satellite, drone and other remote earth observation technologies
- Soil and geomorphology
- Hydrology
- Atmospheric, including meteorological and climate
- Landscape and terrestrial, including ecology, biology, and genetics, through samples and sensors.

Many of these observations and measurements are made in Australia. Some will be sourced internationally. There are rapidly emerging technologies for more automatic collection and recognition of environmental characteristics, including environmental DNA sampling, automated species recognition software, and environmental sensors.

COLLECTION SYSTEMS AND PROTOCOLS

Data collection is often expensive and there are many ways to collect data, including traditional ecological field assessments, as well as newer technology involving field assessment mobile apps, remote capture through cameras and other sensors. Systems are rapidly evolving but the need to collect and curate data remains common – including interoperability across collection platforms. Even if sensors or methods change, they may be measuring the same phenomenon, and so ways to ensure comparability need to be continually updated.

There are often trade-offs between survey methods, survey detail, and the range of subsequent uses. Different approaches and solutions to these trade-offs may cause a lack of comparability among data sets, in particular in relation to national and regional data sets. In addition, it can be difficult to compare data collected at site level with that at larger scales.

Data sets sometimes extend beyond offering evidence that a species was present at a given location and date, also making it possible to assess community composition for broader taxonomic groups or the abundance of species at multiple times and places. These quantitative or sampling-event data sets typically derive from standard protocols for measuring and monitoring biodiversity such as vegetation transects, animal or bird censuses and freshwater or marine sampling. Through indicating the methods, events and relative abundance of species recorded in a sample, these data sets improve comparisons with data collected using the same protocols at different times and places.

DATA DESCRIPTIONS, METADATA AND FITNESS FOR PURPOSE

Data quality refers to data being fit for purpose. In the case of environmental assessments, the purposes for which data contributes relate to investment planning, regulatory decisions and public trust, and these impose high requirements for data quality.

There is no simple agreed definition of data quality, and it differs depending upon whether the data are from samples, observations and measurements, or is statistical or derived from modelling. Data quality starts at the point of creation. This applies to the data itself, as well as the metadata used to describe it. Clear information about data quality enables decisions to be made about how different forms of uncertainty can be propagated through analytics to provide the end user with overall estimates of uncertainty.

Metadata is an essential component of data quality and is important to enable decisions regarding fitness for further use. There are many metadata schemas, or structures for metadata about particular domains. Metadata schemas specify a set of metadata concepts or terms, as well as their definitions and relationships. They also clarify how the data were created, scale of the data, the areas of interest, any cleaning or validation processes and whether there are any restrictions that apply to the data. Reference data is a form of metadata, used to classify or categorize other data, e.g. in relation to units of measurement or calendar structures, and typically changes only slowly over time.

Image courtesy: Brooke Gibbens, UWA/CSIRO

REFERENCE SAMPLES

Reference specimens maintained in biological collections include the material samples on which new species are described – the type specimens – and also additional specimens that represent the variety and variability that support species identification. Collections are essential for taxonomic and systematic research, identification and naming.

Reference samples may be housed in physical collections, such as museums, herbaria and other collection institutions.

DATA CATALOGUES

There are a number of international and Australian biodiversity and ecosystem data catalogues. The Catalogue of Life is the most comprehensive and authoritative global index of species, holding information on the names, relationships and distributions of over 1.8 million species.

The longevity of access to data sources is an issue. The Biodiversity Information Projects of the World (BIPW) attempted to organise the web's biodiversity databases into an indexed list. A recent investigation found that of 600 databases from BIPW, only half were accessible in early 2020.[4] Many of the other databases remain available, but are difficult to access because of broken web links, etc, putting the information they contain in danger of being lost.

In Australia, there are policies requiring data created using public funds to be made public. While this is often the case, and there are some excellent national and state level data aggregators, much research data is at best difficult to find.

[4] Blair J, Gwiazdowski R, Borrelli A, Hotchkiss M, Park C, Perrett G, Hanner R (2020) Towards a catalogue of biodiversity databases: An ontological case study. *Biodiversity Data Journal* 8: e32765. https://doi.org/10.3897/BDJ.8.e32765

Image courtesy: WAMSI

FURTHER SOURCES: COLLECT

OBSERVATIONS AND MEASUREMENTS:

Geoscience Australia – www.ga.gov.au

Bureau of Meteorology – www.bom.gov.au

Atlas of Living Australia (ALA) – ala.org.au

Terrestrial Ecosystem Research Network (TERN) – www.tern.org.au

Integrated Marine Observing System (IMOS) – www.imos.org.au

CSIRO – www.csiro.org.au, adaptnrm.csiro.au/biodiversity-options/ (and much more)

Global Biodiversity Information Facility (GBIF) – www.gbif.org

Australian Government national authority on measurement – www.industry.gov.au/policies-and-initiatives/national-measurement-institute

COLLECTION SYSTEMS AND PROTOCOLS:

Terrestrial Ecosystem Research Network (TERN) – www.tern.org.au

Index of Biodiversity Surveys for Assessments (IBSA) – www.wa.gov.au/service/environment/environmental-impact-assessment/program-index-of-biodiversity-surveys-assessments

Index of Marine Surveys for Assessments (IMSA) – www.epa.wa.gov.au/forms-templates/instructions-for-preparing-data-packages-for-the-index-of-marine-surveys-for-assessments-imsa

DATA DESCRIPTIONS, METADATA AND FITNESS FOR PURPOSE:

ARDC – ardc.edu.au/resources/working-with-data/metadata/

Bureau of Meteorology – www.bom.gov.au (including National Environmental Information Infrastructure)

Geoscience Australia – www.ga.gov.au/about/facilities/geophysical-network/data-quality-control-data-availability-and-statistics

Terrestrial Ecosystem Research Network – www.tern.org.au

National Vegetation Information System – via: www.environment.gov.au

Australian Bureau of Statistics – www.abs.gov.au/websitedbs/D3310114.nsf/home/Quality:+The+ABS+Data+Quality+Framework

5* Open Data – 5stardata.info/en/

REFERENCE SAMPLES:

Australasian Virtual Herbarium – avh.chah.org.au

Geoscience Australia – www.ga.gov.au

National soil archive – www.csiro.au/en/Do-business/Services/Enviro/Soil-archive and www.asris.csiro.au

DATA CATALOGUES:

Data.gov.au – www.data.gov.au

CSIRO Knowledge Network – via: www.csiro.org.au

Global Biodiversity Information Facility – www.gbif.org

Catalogue of Life – www.catalogueoflife.org

CURATE

Managed data sets, layers and products	Data curation and QA/QC standards	Taxonomy and vocabularies	Annotations and corrections	Socio-economic contextual data

Curate level is the engine room where data are processed to make it fit for purpose, complete and interoperable. Data curation is an active and ongoing process that covers the full data lifecycle.

The result is often a dataset that differs in structure and form from the original data. Organizing data in forms that can support analysis and modelling should increasingly take place through automated processing, based upon naming frameworks, structures and standards.

There are varying views of 'big data'. For some types of data there are powerful means emerging to automate curation, ingesting large amounts of data of different types, and enabling it to be used even if poorly structured or defined. Other types of data, including species observations data, have to date proved less amenable to such treatment and require manual, often expert, curation. Both can contribute. In some circumstances, standardised, long term data sets are critical to validating and calibrating approaches that use large, unstructured data sets; and large unstructured data sets can sometimes extend conclusions that might be drawn from standardised, longer term data sets.

MANAGED DATA SETS, LAYERS AND PRODUCTS

Managed data sets, layers and products are primary inputs for further analysis. The quality and the accuracy of the records in the data are an integral part in both the selection of the data and in preparing it for subsequent analysis. Some data may need filtering before it is fit for use. The method of analysis to be undertaken will determine the degree to which the data may need filtering. Data may not always cover the areas of concern and some form of modelling may be required to extrapolate into those areas where the data are inadequate. There is often an extensive process involved in preparing data sets, layers and products for further use.

DATA CURATION AND QA/QC STANDARDS

Data quality management is a process where protocols and methods are employed to ensure that data are properly collected, handled, processed, used, and maintained at all stages of the data lifecycle.

It is critical that the approaches used to transform raw data into managed data sets, layers and products is transparent.

The increasing availability of digitized biodiversity data, provided by an increasing number of institutions and researchers, and the growing use of those data for a variety of purposes have raised concerns related to the "fitness for use" of such data and the impact of data quality on the outcomes of analyses, reports, and decisions. A consistent approach to assess and manage data quality is currently critical for biodiversity data users. However, achieving this goal has been particularly challenging because of idiosyncrasies inherent in the concept of quality.

TAXONOMY AND VOCABULARIES

A vocabulary sets out the common language a discipline has agreed to use to refer to concepts of interest. One of the current challenges for biodiversity data is that while there are broadly shared vocabularies, there are many exceptions. There are also important gaps, for example in relation to descriptions of pressures or threats to the environment.

Vocabularies, taxonomies and other knowledge organisation systems ensure that both machines and humans can interpret and use data arising from multiple sources. Agreed vocabularies are important to enable efficient collaboration to occur.

ANNOTATIONS AND CORRECTIONS

Data quality can be improved through annotations and corrections, using human and automated tools to correct and annotate individual data elements, so that annotations become visible to researchers who subsequently access the data. Annotations often take the form of metadata, whereas corrections modify the original data record. Annotations should be tied to the original data.

Annotations need to be transparent, and their provenance traceable. Tools have now been developed for online annotations and corrections to be associated with digital records.

SOCIO-ECONOMIC CONTEXTUAL DATA

Decision support tools may need information beyond that of the biophysical world. Data about social and economic factors need to be capable of being integrated with biophysical data. There are many rich sources of social and economic data, often highly structured and readily accessed.

FURTHER SOURCES: CURATE

MANAGED DATA SETS, LAYERS AND PRODUCTS:

Geoscience Australia – www.ga.gov.au

Bureau of Meteorology – www.bom.gov.au

Atlas of Living Australia – ala.org.au

Terrestrial Ecosystem Research Network – www.tern.org.au

Integrated Marine Observing System – www.imos.org.au

ABARES – www.agriculture.gov.au/abares

CSIRO – www.csiro.org.au, adaptnrm.csiro.au/biodiversity-options/ (and much more)

Global Biodiversity Information Facility (GBIF) – www.gbif.org

DATA CURATION AND QA/QC STANDARDS:

Australian National Data Service (now subsumed into ARDC) **–** www.ands.org.au/guides/curation-continuum

Australian Bureau of Statistics – www.abs.gov.au/websitedbs/D3310114.nsf/home/Quality:+The+ABS+Data+Quality+Framework

National Archives of Australia – www.naa.gov.au/information-management/building-interoperability/interoperability-development-phases/data-governance-and-management/data-quality

TAXONOMY AND VOCABULARIES:

Australian Biological Resources Study (ABRS) – www.environment.gov.au/science/abrs

Taxonomy Australia – www.taxonomyaustralia.org.au/about-taxonomy-australia

Australian Research Data Commons – ardc.edu.au/services/research-vocabularies-australia/

National Environmental Information Infrastructure – www.neii.gov.au

ANNOTATIONS AND CORRECTIONS:

Atlas of Living Australia – www.ala.org.au/blogs-news/annotations-alerts-about-new-annotations-and-annotations-of-interest/

SOCIO-ECONOMIC CONTEXTUAL DATA:

Numerous sources, including:

ABS – www.abs.gov.au

data.gov.au – www.data.gov.au

INTEGRATE

The Integration tier takes data and curated data products and links them to other data products in preparation for being used in analytic and modelling tools. It also identifies the key characteristics necessary to ensure their continued integrity, and the scientific basis for their integration.

ECOSYSTEM AND RELATED SCIENCE

The Convention on Biological Diversity defines an ecosystem as a dynamic complex of plant, animal and micro-organism communities and their non-living environment interacting as a functional unit.[6] Some aspects of ecosystems, such as individual species distributions and community-level modelling of species, have good data streams and modelling tools, and the challenges are well known. How different species interact and communities and ecosystems function is less understood.

CONCEPTUAL FRAMEWORKS AND METHODS FOR MODELLING

Conceptual models of the environment – from the molecular level to whole ecosystems – remain immature overall. There is considerable scientific work needed to build conceptual models to improve understanding of biological systems and integrate that knowledge into other models, from geology to economics.

One commonly applied classification of regions is the Interim Biogeographic Regionalisation for Australia. IBRA classifies Australia's landscapes into 89 large geographically distinct bioregions based on common climate, geology, landform, native vegetation and species information. The 89 bioregions are further refined to form 419 subregions which are more localised and homogenous geomorphological units in each bioregion. IBRA is a more detailed subset of the global ecoregions defined by WWF.

One conceptual approach, the Australian Ecosystem Models Framework, captures knowledge of ecosystem dynamics in a set of dynamic ecosystem models which describe the dynamic characteristics and drivers of Australian ecosystems. The models have the potential to provide an architecture for natural resource management prioritisation, including environmental assessments, as well as monitoring and evaluation.

Causal networks or causal pathways provide an approach to estimating cumulative impacts, through identifying pressures, and assessing links between pressures and endpoints. Risks to ecological or other values are assigned by likelihood and severity, and protection and management processes that might avoid or mitigate potential impacts can be considered. Causal networks can offer a transparent logic linking pressures, states and responses, common factors included in environmental assessments.

[6] https://www.cbd.int/convention/articles/?a=cbd-02

STANDARDS AND SYSTEMS FOR DATA SHARING AND EXCHANGE

Standards for data exchange ensure that information carries consistent meaning across various transformations and as it is fed into analytic and modelling systems. Standards achieve this by using a dictionary of agreed terms, definitions, relationships and formats regardless of how the information is stored. These standards are built into APIs, and may enable provision of data as a service (DaaS). DaaS takes observation and derived data and makes it model-ready in relation to format, structure, temporal and spatial up/downscaling, translation, harmonisation, quality control and uncertainties.

Data sharing often takes place according to legal agreements between a custodian and a recipient. These can be effective means to manage the risk of unwanted release, though may also be time-consuming to finalise. Systems to track data lineage, covered below, can also help mitigate risk.

INDIGENOUS KNOWLEDGE

Irreplaceable knowledge is held, and continually developed, by Aboriginal and Torres Strait Islander people. Where appropriate, some of this has been integrated into management practice, for instance through Indigenous ranger and Caring for Country programs. Best practice has Indigenous engagement and knowledge built into management approaches, based upon a clear sense of value to the Indigenous people involved, as well as long term ongoing engagement.

Indigenous people will often have both direct interests and deep expertise that will benefit decision making. Engagement with traditional owners and Indigenous communities is needed to ensure culturally appropriate governance processes, supply and use of traditional knowledge, and engagement with data, data products, model structures and modelled outputs. Engagement involves Indigenous people being involved in decision making at all levels.

PROVENANCE AND LINEAGE

The ability to track unit level data from the point of creation through curation and to integration into data products and systems for exchange, to be analysed and modelled, is critical to building confidence in decision-support tools. This is often referred to as recording provenance or lineage. It facilitates data sharing through providing reassurance to the data custodian; it also helps tracks intellectual property. It can aid analysis of results based upon dependencies upon particular data or other inputs, as well as error-detection, auditing and compliance investigation. Capturing provenance and lineage may require considerable metadata documentation, including of data transformations. It is often challenging to track which data or data products contributed to model results.

Assurance of integration processes is needed to ensure that the inputs to be analysed and modelled are fit for purpose. Persistent identifiers are important to ensure the longer term stability of references to particular data or knowledge products or model versions. Digital object identifiers (DOIs) are commonly used unique identifiers.

Image courtesy: Monique Grol

FURTHER SOURCES:

ECOSYSTEM AND RELATED SCIENCE:

Many sources, including:

Ecosystem Science Council – ecosystemscience.org.au

Ecological Society of Australia – www.ecolsoc.org.au

CONCEPTUAL FRAMEWORKS AND METHODS FOR MODELLING:

Australia Ecosystem Model Frameworks – research.csiro.au/biodiversity-knowledge/projects/models-framework/

Australian bioregions – www.environment.gov.au/land/nrs/science/ibra

Pressure state response model – www.epa.wa.gov.au/state-environment-reporting

Causal pathways – www.bioregionalassessments.gov.au/methods/developing-conceptual-model-causal-pathways

STANDARDS AND SYSTEMS FOR DATA SHARING AND EXCHANGE:

Australian Research Data Commons – ardc.edu.au/services/research-data-australia/

National Archives of Australia – www.naa.gov.au/information-management/building-interoperability/interoperability-development-phases/implementation/data-exchange

National Data Commissioner – www.datacommissioner.gov.au/resources/draft-data-sharing-agreement-template

INDIGENOUS KNOWLEDGE:

National Indigenous Australians Agency – www.niaa.gov.au/indigenous-affairs/environment

AIATSIS – aiatsis.gov.au

IP Australia – www.ipaustralia.gov.au/understanding-ip/getting-started-ip/indigenous-knowledge

CSIRO – www.csiro.au/en/Research/LWF/Areas/Pathways/Sustainable-Indigenous/Our-Knowledge-Our-Way

Bureau of Meteorology – www.bom.gov.au/iwk/

CSIRO – www.csiro.au/en/Indigenous-engagement/Indigenous-engagement

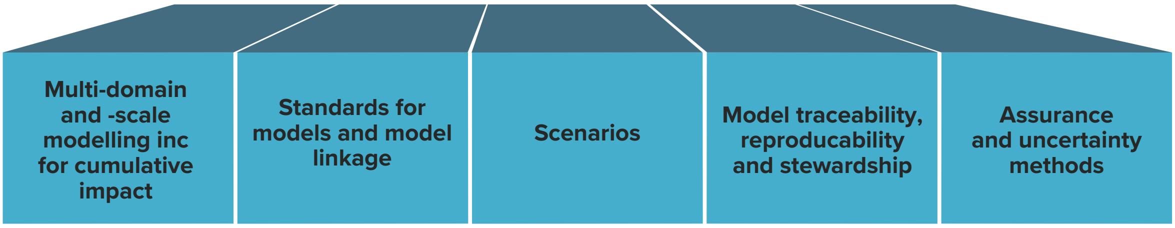
Australian Institute of Health and Welfare – www.aihw.gov.au/reports/indigenous-australians/engagement-with-indigenous-communities-in-key-sect/contents/table-of-contents

PROVENANCE AND LINEAGE:

Digital object identifier system – www.doi.org

Australian National Data Service – www.ands.org.au/working-with-data/publishing-and-reusing-data/data-provenance

ANALYSE



The Analysis tier identifies the analytic and modelling capabilities that underpin decision support tools.

MULTI-DOMAIN AND MULTI-SCALE MODELS, INCLUDING FOR CUMULATIVE IMPACT

Domain specific models are available for environmental characteristics including climate, land surface, species, hydrology and habitat extent and condition. There are several such models available, peer reviewed and curated, often with reliable data sources, though particularly at local scales data may need to be supplemented.

Cross-domain modelling can take place through the loose coupling of specialised models. This has the advantage that the specific strengths of each model are retained, though limited information is generally exchanged between coupled models, and often in only one direction, with an accompanying lack of feedback between the modelled components. There is also the risk of inconsistencies in representations of the same phenomenon in the different models.

Scale refers to the extent and resolution of models. Ecological patterns and processes change at different scales. Ecosystems have different features and structures that influence inter-relationships between interacting species. The scale-dependence of these relationships is not always apparent because of variations in methodological reliability as well as data availability and accuracy.

Integrated assessment models embed different model representations of the system in a consistent manner. The inclusion of feedback and interaction between the different modules is generally stronger and there is more likely to be consistent representation of variables across the different modules. Such models have inherent complexity, which reduces the applicability and transparency of the models.

A challenge is to make models dynamic, able to readily update based on new flows of data.

Simple models of ecosystems have been developed using Bayesian or causal networks to organise disparate information in a consistent framework and incorporate some of the uncertainties inherent in natural systems. Causal networks have potential to contribute more significantly to environmental impact assessment, as they can concisely document cause-and-effect relationships. These models can be attractive due to their high transparency, the possibility to combine empirical data with expert knowledge and their explicit treatment of uncertainties. Sensitivity analysis tools allow characterisation of uncertainties so that key causal factors and knowledge gaps can be identified. Such models tend not to represent dynamic processes, as continuous probability distributions require conversion into discrete figures for calculation.

Cumulative impact remains a challenge, as it is often not additive but multi-factorial with feedback loops and relationships between different factors that are not always well understood.

STANDARDS FOR MODELS AND MODEL LINKAGE

Standards for models and model linkage can help assess the reliability of model results, ensure transparency and consistency in the translation of scientific results into decision support tools, and focus on where improvements might be most needed in the underlying science.

There are no overall standards for ecosystem models. There are however means to select models and tools for analysis, and standardised QA/QC procedures for risk characterisation and peer review.

SCENARIOS

Scenarios are representations of possible futures for one or more components of a system, particularly for pressures or drivers of change. They often incorporate alternative intervention, policy or management options. They offer a key opportunity for end user engagement and interaction with the modelling system and a component of communication and education to understand the basis of the system being modelled.

Scenarios and models play complementary roles, with scenarios describing possible futures for drivers of change or policy interventions and models translating those scenarios into projected consequences for nature.

Some decision support tools – especially at more local levels of decision making – will require consideration of intervention options that are not necessarily known in advance, but arise dynamically. Analysis and modelling in such situations requires the modification of intervention scenarios informed by feedback on the modelled consequences of these options.

The spatial extent and resolution of scenarios and models needs to be aligned with the scale of interest. Temporal scales – ranging from changes made over a few years, through to those focused on achieving longer-term change over several decades – have implications for any scenarios and models.

MODEL TRACEABILITY, REPRODUCIBILITY AND STEWARDSHIP

How models are specified can have considerable impact upon the results that they produce. It can be difficult to gain visibility of model parameters, or the impact of any choices and changes. Platforms are available that offer models with default parameters, which can be altered while creating an auditable record of change.

Stewardship is a concept commonly applied to the activities that preserve and improve the information content, accessibility, and usability of data and metadata. The same concept is important for models. Stewardship activities are a critical support for assurance and reuse, as well as long-term preservation.

A recent survey by the scientific journal *Nature* found that "more than 70% of researchers were unable to reproduce research by others, and 50% were not even able to reproduce their own results."[5] Metadata standards are one response to this, as are standardised data sets, models and model parameters (together with customisation and DOIs) such as offered by EcoCommons for species distribution and other modelling. The ability to reproduce modelled results is central to public trust and assurance of any decision-support tools.

ASSURANCE AND UNCERTAINTY METHODS

Assurance includes setting standards for best practices, using model-data and model-model inter-comparisons to provide robust and transparent evaluations of uncertainty and encouraging new research into methods of measuring and communicating uncertainty and its impact on decision-making. It includes QA/QC approaches.

Uncertainty in scenarios and models arises from a variety of sources, including insufficient or erroneous data used to construct and test models; lack of understanding or inadequate representation of underlying processes; and low predictability or random behaviour in a system. Biodiversity and ecosystem models currently available provide a range of options to assist policymakers in understanding relationships between drivers and impacts, and in evaluating interventions.

For knowledge products to be used by proponents to shape investment proposals, for regulators to make decisions, and for public trust, models must be both of high standards and known to be so.

An example of model assurance in practice is through the scientific oversight built into EcoCommons. An expert committee provides assurance over 100s of curated data sets, 17+ peer reviewed species models, default model parameters and more. While users can introduce new data and vary parameters, DOIs can be minted for all analytic results creating a permanent record of all data, model parameters, etc, and offering an audit and reproducibility trail.

[5] Feng, X., Park, D.S., Walker, C. et al. 'A checklist for maximizing reproducibility of ecological niche models'. Nat Ecol Evol 3, 1382–1395 (2019). https://doi.org/10.1038/s41559-019-0972-5

FURTHER SOURCES: ANALYSE

MULTI-DOMAIN AND -SCALE MODELLING, INCLUDING FOR CUMULATIVE IMPACT:

CSIRO bushfire modelling – research.csiro.au/spark/

CSIRO – www.csiro.org.au

Habitat Condition Assessment System – research.csiro.au/biodiversity-knowledge/projects/hcas/

Australian Urban Research Infrastructure Network (AURIN) – aurin.org.au

STANDARDS FOR MODELS AND MODEL LINKAGE:

National Environmental Information Infrastructure [BOM] – www.neii.gov.au/about

Foundation Spatial Data Framework – fsdf.org.au

SCENARIOS:

Intergenerational Report – treasury.gov.au/intergenerational-report

CSIRO Australian National Outlook – via: www.csiro.au

Intergovernmental Panel on Climate Change emissions scenarios – via: www.ipcc.ch

Nature Futures Scenarios – ipbes.net/scenarios-models

MODEL TRACEABILITY, REPRODUCIBILITY AND STEWARDSHIP:

EcoCommons – ecocommons.org.au

Model reproducibility standard – Feng, X., Park, D.S., Walker, C. et al. 'A checklist for maximizing reproducibility of ecological niche models'. Nat Ecol Evol 3, 1382–1395 (2019). doi.org/10.1038/s41559-019-0972-5

ASSURANCE AND UNCERTAINTY METHODS

EcoCommons – ardc.edu.au/project/ecocommons-australia/

A number of reports discuss approaches to uncertainty in their field, eg. ***Great Barrier Reef Outlook Report 2019*** – www.gbrmpa.gov.au/our-work/outlook-report-2019

wabsi.org.au

wamsi.org.au

The Western Australian Biodiversity Science Institute

Western Australian Marine Science Institution

Government of Western Australia

WABSI and WAMSI are supported by the State Government of Western Australia, through the Department of Jobs, Tourism, Science and Innovation