Submission

2016 National Research Infrastructure Roadmap Capability Issues Paper

Name	Daryl Lee Belbin
Title/role	Director
Organisation	Blatant Fabrications Pty Ltd

Questions

Question 1:   Are there other capability areas that should be considered?

No, as the capability areas cover the domains of significance.

Question 2:   Are these governance characteristics appropriate and are there other factors that should be considered for optimal governance for national research infrastructure.

The governance characteristics are appropriate, but I would emphasize the priority toward open access models.

Question 3:   Should national research infrastructure investment assist with access to international facilities?

Most certainly as Australia must not be isolated from Best Current Practice (and in some cases generating BCP). In my domain of the environment, an excellent example is that the Atlas of Living Australia is the Australian node of the Global Biodiversity Information Facility (GBIF). Agencies such as GBIF and the international standards group – Biodiversity Information Standards (http://www.tdwg.org) set the scene for biodiversity data information exchange and interoperability. Australia must maintain a consistent technical and politically aware presence with both of those agencies.

Question 4:   What are the conditions or scenarios where access to international facilities should be prioritised over developing national facilities?

In my domain of biodiversity data, there is no justification for international facilities having priority over national facilities. Unlike geoscience, ecology is local-regional and the resources required to provide infrastructure support to those researching and managing areas on which all life depends.

Question 5:   Should research workforce skills be considered a research infrastructure issue?

There are two issues.

There needs to be at least one person within the national research infrastructure facilities that has the research experience and skills of the domain. In the capability area of Environment and natural resource management, there is an ongoing difficulty in ensuring effective communication between those with a background in the environment and those with IT skills to build and support

research infrastructures. This is a significant and long-standing issue for international agencies mentioned above (GBIF and TDWG).

Awareness of national research infrastructure facilities by the research community is far from comprehensive, leading to less than efficient research. Experience suggests that it would be a wise investment to have resources at NCRIS to work with the facilities on outreach to all levels of education, particularly secondary schools and undergraduates.

Question 6:   How can national research infrastructure assist in training and skills development?

The national research infrastructures provide some of the best and most efficient platforms for education, bringing the discovery of, and ready simple and open access to data. The Atlas of Living Australia is the NCRIS facility that I am most aware of, and it provides simple and effective access to Australia's biological and related environmental resources. For example, within seconds, anyone can find out where species occur and what species occur in an area. Associations of species with environmental factors then allow evaluations to be made. The facility is used by primary schools to post-doctoral scientists because it contains the largest central store of biological observations (62 million records) and associated environmental data (~500 layers), is easy to use and has extensive documentation and case studies relevant to all levels.

Question 7:   What responsibility should research institutions have in supporting the development of infrastructure ready researchers and technical specialists?

As (6), research institutions need firstly to be aware of the infrastructures, evaluate their utility and provide feedback to those infrastructures. Those are their responsibilities. If the infrastructures genuinely support research, then 'infrastructure ready researchers and technical specialists' will be a natural outcome.

Question 8:   What principles should be applied for access to national research infrastructure, and are there situations when these should not apply?

As a default position, access to data/information in the research infrastructures should be open and free to all, but due acknowledgement must be an automated aspect of use.

Question 9:   What should the criteria and funding arrangements for defunding or decommissioning look like?

In the capability area 'Environment and natural resource management', I cannot foresee anything other than Federal Government funding for research infrastructures for the 'public good'. At a rational level, it is difficult to believe that a facility such as the Atlas of Living Australia being decommissioned. All the work that has gone into the gathering, integration and delivery (in an effective manner) of biological observations of the Australian region would be lost, in some cases forever. Such facilities require a commitment to ongoing funding at least at maintenance level to be able to service research, education, environmental managers, environmental consultants, NGOs and the public.

Question 10:   What financing models should the Government consider to support investment in national research infrastructure?

As (9). While I can see some one-off funding from industry and education, this needs to be targeted at the development of specific value-add projects. For example, mining and oil companies and consultants could provide one-off funding for the development of a mechanism in the Atlas of Living Australia for automated preliminary reporting for environmental impact assessment.

Question 11: When should capabilities be expected to address standard and accreditation requirements?

Standards: Effective international standards are the foundation for efficient research infrastructures. It is imperative, as noted above, that national research infrastructures have a high profile on relevant international standards groups. In the case of 'Environment and natural resource management', and specifically facilities such as the Atlas of Living Australia and TERN, this is largely Biodiversity Information Standards.

Accreditation: I would associate accreditation of individuals and industries to <u>effective</u> usage of the data/information available. This is turn depends on the data and metadata. While I have been training research scientists in the use of data from national research infrastructures, accreditation for the effective use of data from the infrastructure is a new concept that is appealing. I can see great value in for example, for an environmental consultant to display an accreditation with a report for an environmental impact assessment. Accreditation and research seems currently like strange bedfellows.

Question 12: Are there international or global models that represent best practice for national research infrastructure that could be considered?

In the case of "Environment and natural resource management", I believe that the Atlas of Living Australia (Atlas) has surpassed the Global Biodiversity Information Facility (http://www.gbif.org) in providing Best Current Practice in this domain. Evidence: With (moral) support from GBIF, the infrastructure of the Atlas has been, or is in the process of being deployed as national facilities to Argentina, Belgium, Brazil, Costa Rica, Denmark, England, Finland, France, Germany, Ireland, Italy, New Zealand, Norway, Scotland, Spain, Sweden and Wales. Australia is therefore acknowledged as leading the world with its biodiversity data research infrastructure facilities.

Question 13: In considering whole of life investment including decommissioning or defunding for national research infrastructure are there examples domestic or international that should be examined?

Not beyond what I have said above.

Question 14: Are there alternative financing options, including international models that the Government could consider to support investment in national research infrastructure?

Health and Medical Sciences

Question 15: Are the identified emerging directions and research infrastructure capabilities for Health and Medical Sciences right? Are there any missing or additional needed?

Question 16: Are there any international research infrastructure collaborations or emerging projects that Australia should engage in over the next ten years and beyond?

Question 17: Is there anything else that needs to be included or considered in the 2016 Roadmap for the Health and Medical Sciences capability area?

Environment and Natural Resource Management

Question 18: Are the identified emerging directions and research infrastructure capabilities for Environment and Natural Resource Management right? Are there any missing or additional needed?

The effective role that the Atlas of Living Australia fulfilling is demonstrated by international replication and by the usage figures of the Atlas – as at 2/7/2016, 7.7 billion records have been downloaded, 4.2 billion records for research, 0.6 billion for 'conservation management and planning' and 0.15 billion for 'environmental assessment'.

That said, there is a significant gap that was flagged as long ago as 1979 by Professor Henry Nix, then of CSIRO and now retired when he proposed an Australian Environmental Geographic Information System in 1986. That significant gap remains even though there was funding in the 2010-2011 Federal budget for a National Environmental Information System.

I made it a part of my responsibility with the development of the 'research portal' to the Atlas of Living Australia (http://spatial.ala.org.au), to gather environmental data that related in some way to species. Over the past 5 years, 500 such layers have been collected from over 70 different Australian agencies (see http://spatial.ala.org.au/layers). As far as I am aware, this is the only repository of its type in Australia or elsewhere.

These environmental data layers enable anyone to relate the biological observations to environmental information to permit an extremely wide range of analyses. For example, the data in the Atlas can be used to determine where a species may be found given a climate change scenario. It also allows the evaluation of how any species, species richness, endemicity and diversity are partitioned across parks and reserves, States and Territories, Local Government Areas, catchments, ecoregions and much more.

Why do I say that then that there a gap? While the Atlas of Living Australia currently integrates and makes this environmental data available, it is not the logical or ideal home. Environmental data isn't the priority of the Atlas and it does represent ~5 times the overhead of the biological observation data.

What is needed, ideally, is a small separate national facility who job it is to a) facilitate the discovery of Australia's environmental data and 2) work with the agencies producing the data so as to provide open and standard web services for access. We do not want or need another central repository if such web services could be established. The Atlas provides an example of such a service to intersect any one or more locations with any of the 500 environmental layers - http://spatial.ala.org.au/ws/intersect/cl22/-23.1/149.1 - which returns the State or Territory while http://spatial.ala.org.au/ws/intersect/el874/-23.2/149.2 will return the mean annual temperature at the location in the web address.

The environmental data gets updated regularly and the management effort associated with detecting such changes and updating the information is considerable for the Atlas. It would make far more sense for each of the agencies who created the data to provide, transparently, the updated data via their own, but nationally standardized web service. I have asked Geoscience Australia for such a service to determine which State/Territory/Marine jurisdiction a coordinate is in – and they have agreed that this would be the ideal way to provide the data, rather than 'publishing' a complete new dataset as the unit of work and in a form that is less useful. A small national agency in NCRIS that begins with a focus on coordinating standardized web access to bio-environmental data/metadata would be an investment resulting in considerable cost-savings. How would costs be saved?

- **Scientists and others would have one place where relevant data could be discovered**

- **Time save in processing data in different formats, resolutions and inconsistent metadata**

- **Updates would be automatically available**

- **Standard web services would enable maximal re-use of valuable data**

- **The data is stored at the location of its production. There is no current need for a centralized facility where data would have to be transformed, merged and managed.**

- **The data would remain close to the expertise that generated it**

Question 19: Are there any international research infrastructure collaborations or emerging projects that Australia should engage in over the next ten years and beyond?

As noted above, the infrastructure of the Atlas of Living Australia is being exported as follows (at the time of writing)

Implemented-

1. Spain (http://www.gbif.es/en)

2. France (http://portail.gbif.fr/)

3. Costa Rica (http://www.crbio.cr/crbio/?&)

4. Brazil (https://portaldabiodiversidade.icmbio.gov.br/portal/search#tab_simpleSearch)

5. Scotland (http://www.als.scot)

The following countries are at various stages of implementing a site using the ALA infrastructure but aren't live yet:

1. Argentina

2. Portugal

In discussions with the following countries about using ALA infrastructure to develop their own sites:

1. England

2. Wales

3. Germany

4. New Zealand

5. Belgium?

6. Norway

7. Sweden

8. Mexico (CONABIO)

9. Symbiota (Northern Arizona) - looking at some of the spatial components.

Question 20: Is there anything else that needs to be included or considered in the 2016 Roadmap for the Environment and Natural Resource Management capability area?

Not that I am aware of.

Advanced Physics, Chemistry, Mathematics and Materials

Question 21: Are the identified emerging directions and research infrastructure capabilities for Advanced Physics, Chemistry, Mathematics and Materials right? Are there any missing or additional needed?

Question 22: Are there any international research infrastructure collaborations or emerging projects that Australia should engage in over the next ten years and beyond?

Question 23: Is there anything else that needs to be included or considered in the 2016 Roadmap for the Advanced Physics, Chemistry, Mathematics and Materials capability area?

Understanding Cultures and Communities

Question 24: Are the identified emerging directions and research infrastructure capabilities for Understanding Cultures and Communities right? Are there any missing or additional needed?

Question 25: Are there any international research infrastructure collaborations or emerging projects that Australia should engage in over the next ten years and beyond?

Question 26: Is there anything else that needs to be included or considered in the 2016 Roadmap for the Understanding Cultures and Communities capability area?

National Security

Question 27: Are the identified emerging directions and research infrastructure capabilities for National Security right? Are there any missing or additional needed?

Question 28: Are there any international research infrastructure collaborations or emerging projects that Australia should engage in over the next ten years and beyond?

Question 29: Is there anything else that needs to be included or considered in the 2016 Roadmap for the National Security capability area?

Underpinning Research Infrastructure

Question 30: Are the identified emerging directions and research infrastructure capabilities for Underpinning Research Infrastructure right? Are there any missing or additional needed?

The key issue is ongoing financial support for significant national facilities such as the Atlas of Living Australia that has

- **Discovered biological observation and associated data across multiple sources and negotiated open use (CC-BY) licences**

- **Integrated that data into a single repository and made it openly available is an extremely useful form to a wide variety of communities**

- **Been recognised as world Best Current Practice through other countries seeking to use the infrastructure for their national facilities**

Second, the gap identified in 18 above has been long recognized, remains an area where significant cost/benefit can be readily achieved and has not been addressed.

Question 31:   Are there any international research infrastructure collaborations or emerging projects that Australia should engage in over the next ten years and beyond?

As noted in the response to Q19, it is anticipated that the infrastructure developed by the Atlas of Living Australia will be taken up by other nations. Such collaboration would be expected to benefit all involved. For example, improvements on the Atlas infrastructure by one country would be expected to feed through to others if relevant. The shared infrastructure forms an ideal basis to improve international standards and interoperability.

Question 32:   Is there anything else that needs to be included or considered in the 2016 Roadmap for the Underpinning Research Infrastructure capability area?

Not that I am aware of.

Data for Research and Discoverability

Question 33   Are the identified emerging directions and research infrastructure capabilities for Data for Research and Discoverability right? Are there any missing or additional needed?

Yes, except for the gap identified under Q18. The strategy of using standardized web services would benefit from advances in HPC, high-capacity networks, trusted communication, geospatial systems and digitization.

Question 34:   Are there any international research infrastructure collaborations or emerging projects that Australia should engage in over the next ten years and beyond?

Covered above.

Question 35:   Is there anything else that needs to be included or considered in the 2016 Roadmap for the Data for Research and Discoverability capability area?

Other comments

If you believe that there are issues not addressed in this Issues Paper or the associated questions, please provide your comments under this heading noting the overall 20 page limit of submissions.