Response to the Department of Innovation, Industry, Science and Research (DIISR) 2011 Strategic Roadmap for Australian Research Infrastructure Discussion Paper

Council of Australian Biological Collections (CABC)

4 May 2011

Background

The Council of Australian Biological Collections (CABC) welcomes the opportunity to comment on DIISR's 2011 Strategic Roadmap for Australian Research Infrastructure Discussion Paper.

CABC is a peak body representing biological collections in Australia. It currently comprises the executives of the following collections councils:

- Council of Heads of Australasian Herbaria Incorporated (CHAH, representing all State, Commonwealth and University herbaria in Australia and New Zealand);
- Council of Heads of Australian Faunal Collections Incorporated (CHAFC, representing Australian local government, state and territory natural history museums, and Commonwealth zoological collections);
- Council of Heads of Australian Entomological Collections (CHAEC, representing Australia's insect collections);
- Australian Biological Resources Study (ABRS).

Representatives of the Council of Heads of Australian Collections of Microorganisms (CHACM) have also been invited to join CABC, with acceptance yet to be confirmed. From time to time the council appoints *ex officio* observers; the Atlas of Living Australia (ALA) is currently an observer.

CABC represents the vast majority of biological collections in Australia. It provides an authoritative, coordinated body representing biological collections and their research institutions to government, industry and the broader community. The aims of the Council are to:

- coordinate and develop common approaches to biological collections issues;
- raise awareness and appreciation of the importance of evidence-based research focussed on Australia's collections of our natural heritage, particularly through the sciences of taxonomy, systematics, biogeography, molecular genetics and other related biodiversity research issues;
- respond to government needs on biodiversity;
- provide advice to government on national and international biodiversity issues.

Together, members of CABC are custodians of a distributed national biodiversity collection holding in excess of 36 million specimens of Australian plants, algae, fungi, insects, animals and microbes (ABRS, 2006), in over 15 leading institutions. The combined value of the collections is estimated to be in excess of $4.7 billion; it is, however, irreplaceable, representing an enormous fraction of the sum knowledge of the biodiversity of Australia and its territorial waters.

Together, the collections (here called the Distributed National Biodiversity Collections Research Infrastructure) constitute critical national infrastructure, and support research directly related to five of the six National Research Priority areas.

Collections as infrastructure

Infrastructure comprises the physical, organisational and capital structures required for the operation of an enterprise, in this case biodiversity research. Infrastructure is often built up over long time periods, is difficult and costly to replace, requires ongoing maintenance investment, and is required for the growth of knowledge, investment and economic return.

The Distributed National Biodiversity Collections Research Infrastructure represents investment and effort over several centuries. It comprises three main components: biological specimens, information about the specimens, and the human resources that build, manage and create knowledge based on the specimens for the government, the Australian and international research communities, and the public.

Collections have been identified as critical infrastructure by the European Union in the *European Roadmap for Research Infrastructures: Roadmap 2008*, and mission-critical national research infrastructure by the *USA Interagency Working Group on Scientific Collections*. An important strength of collections is in the high quality, verifiable information that collections document. Specimens in collections voucher scientific endeavour over long time spans, providing quality assurance and future-proofing of data when compared to observational records. Vouchers can be revisited in the face of new information and new technologies, generating, validating and regenerating new information about species through time. Adequate vouchering of data is an area of growing concern in the scientific community; there is an urgent need to ensure that data generated by new technologies such as genomics and proteomics can be linked back to specimens and to species, as this enables studies to be verified and revised as needed. Similarly, vouchers provide a mechanism for complying with increasingly tight data accounting standards imposed by state audit agencies. Without vouchers in collections, biodiversity data degrades rapidly over time in the face of new knowledge.

The creation of any infrastructure invites several key questions that should influence initial and ongoing investment. Large scientific research infrastructure in the physical sciences (such as the Square Kilometre Array and Large Hadron Collider) requires both initial investment to build and establish but also ongoing investment to cover maintenance and running costs. Investment may also be required to repurpose existing infrastructure to address new questions and respond to new technologies.

Similar issues arise in the biodiversity collection infrastructure space. Significant initial investment over several centuries has created a nationally and internationally significant research infrastructure, but without maintenance and growth this investment will degrade. The significance of the long time-series data represented by biodiversity collections cannot be underestimated. For example, research at the Swedish Museum of Natural History in the 1960s on specimens collected over the previous hundred years showed trends in accumulation of mercury in animals. Recognizing that collections could be used for this purpose led to the creation of a tissue bank with better taxonomic, geographic and temporal coverage. Today, the tissue bank is used to create long time series data for a range of pollutants in the environment. Specimens in collections have also been used as sources of data on historical change in ratios of stable isotopes in the environment and on morphological adaptation to change, and can thus be used as tools for long-term environmental monitoring. For example, Woodward (1987) demonstrated that plants have responded to changing atmospheric carbon dioxide levels by altering the density of pores for gas exchange on the leaf surface. This work would not have been possible without herbarium collections built up over hundreds of years. Having developed a calibration it was then possible to retrospectively determine atmospheric carbon dioxide levels and to extend the time series of instrumental records back hundreds of years (and even millions once fossil calibrations were developed). This highlights the opportunities to repurpose collections for new and innovative research, and the need for investment in both hard (specimens, equipment) and soft (organisational and human resource) infrastructure

In the biodiversity collections infrastructure area, specimens and their physical housing comprise hard infrastructure, while in-house and external research capacity and resources such as the ALA comprise soft infrastructure that supports the hard infrastructure of the collections. The importance of the human capital that manages large infrastructure such as the Distributed National Biodiversity Collections Research Infrastructure should not be underestimated. Our ability to manage and enhance the collections for the broad research community relies on key skills sets and research capacity within the collections institutions, particularly taxonomic expertise. We strongly endorse the Environmentally Sustainable Australia Working Group's recognition that the continuing decline in taxonomic capacity is an important constraint within their Research Infrastructure Capability Area. Species names and species concepts comprise essential framework knowledge that underpins research within the whole of biological and environmental sciences. Examples given in the *Proceedings of the National Taxonomy Forum*[1] held in 2007 are as relevant today as they were then, and indeed the decline in the taxonomic workforce has not been arrested. CABC strongly endorses an investment approach that reverses the decline in taxonomic capacity, and will assist in whatever way feasible in this endeavour.

Commentary

CABC supports the process of road mapping proposed by DIISR in the 2011 Strategic Roadmap for Australian Research Infrastructure and the tie to the National research Priorities (NRPs). We offer some general observations and suggestions below that we hope will assist the strategic development of research infrastructure and improve capacity for collaborative research both nationally and internationally.

We note that on page 7 of the discussion paper DIISR expresses a concern that the structure of the discussion paper, by being broken down into individual NRPs, risks losing the connections between the NRPs and further siloing research. The Distributed National Biodiversity Collections Research Infrastructure cross-cuts five of the six NRP areas and although we are confining our comments largely to the Environmentally Sustainable Australia section they are equally applicable to the other areas. We request that DIISR makes this paper available to all Working Groups for consideration

The Australian Research Council *Excellence for Research in Australia 2010 National Report* noted that Australia performs well above the global average in plant biology, evolutionary biology, ecology, zoology and geology, all areas that the Distributed National Biodiversity Collections Research Infrastructure underpins. Our knowledge of these disciplines through our international linkages supports this assessment. In order to maintain this competitive advantage Australia needs to continue to invest in these areas. CABC suggests it would be prudent for DIISR to consider the Distributed National Biodiversity Collections Research Infrastructure as a candidate for landmark infrastructure investment, to create a virtual national biological collection and national biodiversity reference data set for the broadest possible research community. Such infrastructure is of a scale and significance that falls well within the scope of the Roadmap.

Research infrastructure policy issues

It is encouraging to note (p. 7) that DIISR recognises the need for ongoing maintenance of research infrastructure, and that future research infrastructure should be broadly targeted to support all aspects of research infrastructure including operations and maintenance. In our view, this will add significantly to the effectiveness of NCRIS infrastructure programs. In particular, CABC is very concerned that the life

---

[1] See http://www.environment.gov.au/biodiversity/abrs/funding-and-research/workshops-and-forums/pubs/ntf-proceedings.pdf

span of existing funded infrastructure (e.g. the ALA) be adequately addressed, and that discussions commence on integration of successful aspects of multiple programs (e.g. ALA, TERN, IMOS). In addition, capabilities such as BioPlatforms Australia (BPA) should examine how they interface with the Distributed National Biodiversity Collections Research Infrastructure and how they could integrate with the ALA, especially in the face of an increased need to link the various 'omics capabilities to vouchered collections through the ALA. CABC strongly endorses the comments in the discussion paper about the need for ongoing maintenance of the infrastructure (both hard and soft) and would be pleased to be involved in future discussions on these issues.

Contributions of biological collections to NRPs

Environmentally Sustainable Australia

A challenge in improved understanding of the impacts of climate variability and prediction is having long time series of observations. Establishing long-term observation programs now may not be effective as it may be too late to respond once results are in. An alternative approach is to construct long time series records from archives. There are good examples of long time series reconstructions of atmospheric carbon dioxide proxies from collections (see e.g. the Woodward study noted above), a collection use case that was unanticipated at the time the collections were made. Similarly, estimates of past and future climates in Australia have used insect and plant distribution data based on the National Biodiversity Collections Research Infrastructure to estimate past (pre-Pleistocene) environments in southern Australia (Porch et al. 2009).

Past, current and future comprehensive, large scale biological collections from extensive surveys provide opportunities for as-yet undreamt of reconstructions and analyses in both the marine and terrestrial realms. For example, collections acquired during the $7m Great Barrier Reef Seabed Biodiversity Project (GBRSBD) (Pitcher et al., 2007) include 1,200 sediment samples, 140,000 records of 5,300 species of invertebrates, plants and fishes sorted from 14,000 benthic sample lots, and 4,000 fish sample lots. An estimated 300,000 samples from the project are stored in a variety of national collections. These samples potentially contain many hundreds of new species, as well as providing a critical snapshot of the biota and of marine biological material at this time point. Adequately stored, curated and made accessible to researchers, they will remain an invaluable resource indefinitely.

In the terrestrial realm, the Distributed National Biodiversity Collections Research Infrastructure is engaged in *Bush Blitz* (http://www.bushblitz.org.au/), a program of surveys of the National Reserve System (NRS) funded through a consortium of SEWPaC (ABRS and Caring for Country), BHP-Billiton, and Earthwatch. The globally unique program is Australia's largest nature discovery project and has already identified hundreds of new species and provided thousands of new, authoritative distribution records of plants and animals in the NRS. These records will provide the knowledge base for modelling the proportion of our biodiversity protected in the NRS, as well as allowing quantification of the benefits of different acquisition strategies and investments. The Distributed National Biodiversity Collections Research Infrastructure is critical to this endeavour at all stages, from the expertise required for survey design to curation, taxonomic identification, species resolution, vouchering and electronic capture of data records.

The *Proceedings of the National Taxonomy Forum* (2007)[2] sponsored by the Federation of Australian Scientific and Technological Societies lists a wide range of taxonomic projects that have had a direct and

[2] see http://www.environment.gov.au/biodiversity/abrs/funding-and-research/workshops-and-forums/pubs/ntf-proceedings.pdf

significant impact on environmental sustainability, ranging from the discovery of new species of snakes, frogs, disease organisms and fish to the clarification of the taxonomy and hence control methods for significant weed species. Several of the examples given have also had direct economic benefits saving the Australian economy millions of dollars. None of these projects would have been possible without the infrastructure provided by the National Biodiversity Collections Research Infrastructure, including their specimen base, corresponding electronic information, and the expertise embodied in their researchers.

Promoting and maintaining good health

Taxonomy and collections play a crucial role in maintaining human and animal health. An example is provided by a paper published in *The Lancet* reporting the first case of infection with the parasite *Trichinella pseudospiralis* in humans (two cases from Tasmania). Medical diagnosis was based on the fact that the parasite occurred within muscle cells, an aetiology known at the time to occur only in the genus *Trichinella*. Standard steroid treatment for *Trichinella* was given, but the patients deteriorated. Careful examination of the parasite by a taxonomist and comparison with museum samples revealed that it was not *T. pseudospiralis*, but a completely new genus and species, *Haycocknema perplexum.* Understanding the taxonomy of the parasite proved vital for correct treatment.

Recent recognition that human teeth collected in archaeological excavations, and teeth from anthropological collections, could be used for DNA analysis offered the possibility to repurpose collections and use them to examine the history of diseases and the evolution of disease agents. This emerging field of research (Zink et al 2002) has "great potential to provide information about the origin, evolution and transmission of disease through time" (Roberts and Ingham 2008, p 600). While sample protocols and experimental design still remain controversial, collections can provide real benefit in this and similar fields.

Frontier Technologies

Recognition and application of new technologies has been a hallmark of biological investigation, with many areas being early adopters of new technologies, quickly harnessing their potential to enhance research capabilities. A recent program at the Australian National Insect Collection has resulted in the rapid databasing, imaging and barcode sequencing of *c*. 28,000 specimens of *c*. 10,000 moth species by a small team in 10 weeks. This frontier technology provides a cost-effective way to develop a key reference data set for all Australian species. Barcoding projects such as this, if properly designed and implemented, will play a large role in the future in biodiversity research and may provide the key link between authoritative, vouchered collections and emerging technologies such as metagenomics and ecogenomics.

The CERF funded *Taxonomy Research & Information Network*[3] is another good case of early adoption and implementation. This Network developed tools and applied methods that assisted researchers in speeding up the process of gathering taxonomic information. In particular, these tools seamlessly integrate mobile device technology for field data capture with collection databases and federated data systems. This has substantially reduced bottlenecks in the data life cycle, which typically sees data being rekeyed several times before it is readily available. These tools thus accelerate the timely delivery of data for reporting and analysis in the national and international informatics framework.

[3] See http://www.taxonomy.org.au/

Safeguarding Australia

Safeguarding Australia's aquacultural, agricultural and horticultural industries from pests that are currently absent in Australia give our industries a competitive advantage. A critical part of the control process is having adequate infrastructure to detect and identify pests before they become problems. This involves actions both at border control and post border control. The Distributed National Biological Collections Infrastructure has an important role to play here. Two examples exemplify this role.

Bee mites of the genus *Varroa* are major pests and cause disease in Western honeybees. They do not occur in Australia, and are a major biosecurity threat to Australia's beekeeping and pollination industries. Following evidence that called into question the taxonomy of species within the genus, taxonomists at the Australian National Insect Collection clarified the taxonomic status of *Varroa* mites on their primary host, the Eastern honeybee in the Oriental region. They discovered that two genotypes of a previously undetected species, *Varroa destructor*, had switched host to the Western honeybee, causing millions of dollars of damage each year worldwide. All other species and genotypes of *Varroa* mites were found to be harmless to Western honeybees, because they lacked the ability to reproduce on honeybee brood. These results were only made possible using reference collections and expertise in a number of Australian and overseas research collections.

Mexican Feather Grass is a prohibited weed with potentially major negative impacts in Australia´s pasture lands. A recent incursion of an unidentified species of feather grass was rapidly identified as Mexican Feather Grass through comparison to vouchers held in Australian herbarium collections. This authoritative identification enabled a rapid response to successfully control the outbreak. Without adequate collections in Australia identification would have been delayed substantially, risking establishment of the pest in Australia. As a by-product, molecular identification tools were created against a set of known vouchers to enable rapid future identification of immature and sterile material; grasses are notoriously difficult to identify when immature, the stage at which they must often be collected to control outbreaks.

Understanding cultures and communities

Humans have and will continue to interact with their environment and to utilise the biota with which they co-occur. New and emerging molecular tools are enabling social scientists and biologists to productively interact. For example, a recent research project has used genetic fingerprinting of plant collections to compare populations of plants thought to have been introduced to Australia with plants in populations overseas. This has enabled source regions to be identified and help confirm and refute various human migration hypotheses. Similar work conducted at the Australian National University used molecular tools to track the movement of taro (*Colocasia esculenta*) around the Pacific: taro was carried by Polynesian migrants, so tracking taro helps track the colonization of the Pacific by the Polynesian seafarers. This work relied on herbarium collections, including material originally collected by Joseph Banks on the Cook expedition in Queensland and now housed in Australian herbaria, to identify wild-type varieties of taro in source areas. Studies like these will be a fruitful area for further research in the future.

eResearch Infrastructure

The biological collections community has been quick to adopt new technologies, particularly in the fields of eResearch and ICT. The Australian collections community in particular has been a world leader in this field for many years. Effective management of biological data is in many respects a significant ICT challenge, characterised by dispersed data repositories, a large number of data formats, a high need for standards and data held in often intractable formats such as natural language. These challenges have driven innovation in this sector for many years.

For example, the Herbarium community in Australia was amongst the first in the world to develop a data standard for the exchange of specimen record information. Similarly, globally significant innovations in the handling of biological descriptive data have been driven from Australia with predominant use cases in the collections community. This has resulted in tools which are world-leaders in the field. The ICT challenges of developing eResearch platforms for biodiversity collections on the one hand, and the high level of integration and collaboration between Australian collections institutions in the other, has driven this innovation.

Tools developed for the ALA are set to continue this trend. The geospatial visualisation and analysis tools of the ALA, which allow geocoded specimen records aggregated from all Australian collections to be modelled against a variety of climate and environmental surfaces, show considerable promise for driving discovery of species and ecological patterns. Initial results have already indicated a number of potentially new taxa, discovered merely by looking at existing collections data in a new way.

Researchers within the National Biodiversity Collections Research Infrastructure space have continued to innovate, adapting new technologies and approaches to enhance the information available to the research community. For example, TRIN's adoption of social media, particularly the wiki, provides a collaboration platform facilitating communication and sharing of taxonomic information for research collaboration and delivery of biodiversity information. Further facilitation and incentive for such communication systems should be encouraged.

The Distributed National Biological Collections Research Infrastructure

Many of the capabilities funded through NCRIS (e.g. TERN, the Marine National Facility, Australian Institute of Marine Science) contribute directly to the Distributed National Biological Collections Research Infrastructure through the acquisition and deposition of specimens. Collections in turn support and provide quality assurance for research findings in these areas. Other capabilities (e.g. IMOS) and agencies (e.g. Bureau of Meteorology, Ocean Biogeographic Information System) provide additional observational information that can be integrated with information infrastructure such as the ALA to deliver both data and analysis tools. The ALA has been successful in developing and deploying delivery mechanisms for a wide variety of data about organisms, and the tools that interact with that data.

One critical piece of infrastructure developed by the collections community with support from the ALA is the Australian Biodiversity Census[4], the first agreed national list of all described species of organisms in Australia. No other country of equivalent size and biological richness has achieved this. This fundamental, infrastructural data set is potentially transformative for the business of biodiversity research in Australia, allowing hitherto unprecedented accuracy in determining patterns of biodiversity and in auditing the national biodiversity asset. CABC recommends that there is great potential to develop the Australian Biodiversity Census further, by continuing and enhancing the ALA activity of reorganising existing biodiversity information around the names listed and annotating the names with more information (such as images, descriptions, and barcodes and other genetic markers). This will enable a wide variety of users to identify study organisms and to transform a sometimes meaningless name into a meaningful concept. The ALA has commenced but not completed these activities. We envision that, with suitable investment, within a decade every organism named in the Australian Biodiversity Census could have associated with it a rich set of information. Such an asset would be globally unique.

[4] *Australian National Species Lists* in ALA documentation

Many specimens in our collections remain unidentified (estimated at over 80% for insect collections). A recent study (Bebber et al 2010) has identified collections as a major frontier in species discovery: most new plant species being described were already within collections at the time of their description, with over 25% of new species based on collections that were more than 50 years old. This suggests that of the estimated 70,000 plant species still to be described, over half have already been collected and deposited in herbaria. The case for other groups such as insects is probably even more heavily weighted towards collections being the optimum place for species discovery.

A key challenge exists in realising the potential inherent in the Distributed National Biodiversity Collections Research Infrastructure. Databasing of named specimens and identification of unidentified material in the collections would enormously increase the data available through the ALA to the research community, both in terms of taxon coverage but also spatial and temporal coverage. A sensible way to achieve this would be to tie future developments of the ALA closely to collections and genomics. For example, imagery of exemplars of named specimens and their critical features and genetic barcodes could be used to annotate the national species list. This reference set could then be used as a basis for identification of unidentified material and the recognition of new species. In our view this is an important area for future investment, given the value that has already been released from the collections data and the critical role that collections play in NRPs.

Response to Environmentally Sustainable Australia Expert Working Group

What are your views on the key future research directions identified and are there other key areas that have not been included?

The EWG has identified seven priority goals and the necessary capabilities to achieve a more environmentally sustainable Australia. The Distributed National Biological Collections Research Infrastructure has a clear role to play in helping to deliver on these goals, particularly in repurposing collections for long time series analysis, further developing the underpinning taxonomic knowledge required to make informed decisions and, although not explicitly mentioned, developing phylogenies that can be used to inform and target strategies to achieve the goals (see above under *ESA*). We note, however, that while identifying the need to collect data that will enable better projections for climate change effects in the oceanic regions surrounding Australia, the Discussion Paper makes no mention of a similar need in the terrestrial environment. Knowledge of all Australia´s biodiversity, including the identity and names of organisms, their genetic diversity, their phylogenetic relationships, and their functional role in ecosystems are critical underpinnings to effective sustainable management of both marine and terrestrial ecosystems. Environmental change risks decoupling interactions (e.g. pollinator/plant relationships) and shifting environmental envelopes faster than species can respond, leading to negative outcomes for Australian biodiversity. There is a clear need to further develop tools that can allow species interactions to be explored within an environmental and climatic envelope.

An important new goal for the Distributed National Biological Collections Research Infrastructure is to preserve material in Australian collections that can be harvested for genomic DNA. Tissue preservation techniques will be taxon specific to some extent; Australia's vertebrate collections lead the way in tissue vouchering, storage and use. These temporally and taxonomically defined genomic resources will become an increasingly valuable asset as next generation sequencing technologies improve and become more cost effective. Tissue samples and the genomic data that derives from them can be used in DNA-based identification strategies such as DNA barcoding, but more importantly have the potential to inform a huge range of other fields in global change adaptation research (Johnson et al 2011).

What are your views on the research infrastructure Capability areas identified, including their relative priority and their ability to support the current and future research needs?

CABC agrees that long term and standardised observations of the environment (including organisms in the environment) are important. We regard that equally important is adequate vouchering of material for future-proofing of these observations. Vouchers of specimens in both the monitoring and research framework provide quality assurance and control that means results can be verified at any time. In addition, adequate vouchered material in collections, tied to agreed checklists of the names of organisms, allows observational samples to be adequately identified and named. In a useful synergy, such activity also enhances the Distributed National Biological Collections Research Infrastructure. Specimens can also be repurposed in the future, providing extra (often unpredictable) research value as highlighted above.

It is important to recognise that vouchering incurs costs, both in making and permanently storing and curating the vouchers and in making information about the vouchers available. Any investment in this area needs to recognise these costs and provide a mechanism for cost allocation.

Other NCRIS areas such as BioPlatforms Australia need to be linked in to both the Distributed National Biological Collections Research Infrastructure and to soft infrastructure such as the ALA. BioPlatforms Australia has facilitated metagenomics research contributing to environmental research, and should be supported to enable further development into areas such as soil metagenomics, building a soil biodiversity database for Australia. However, as with all biodiversity studies, attributes such as DNA sequences need to be linked back to voucher specimens to have any future value. CABC believes this area that needs further consideration.

CABC strongly endorses the observation that the continuing decline in taxonomic capacity is an important constraint and major concern within the Research Infrastructure Capability Area. Species names and species concepts comprise framework knowledge that underpins research throughout the biological and environmental sciences. Examples given in the *Proceedings of the National Taxonomy Forum* held in 2007 are as relevant today as they were then, and indeed the decline in the taxonomic workforce has not been arrested. CABC would strongly endorse an approach that reverses the decline in taxonomic capacity.

CABC also notes the positive comments regarding better integration between observational data networks and modelling communities, but points out that both communities benefit substantially from interaction with the collections community.

What are your views on the existing funded facilities, including their ability to meet the current and future research needs?

The Distributed National Biological Collections Research Infrastructure has benefited in several ways from NCRIS investment in the ALA (and other capabilities), not only through the eResearch infrastructure it has created but also through successful engagement of a widely distributed research community, resulting in closer and more effective collaborations. The ALA has been an important first step in aggregating and mobilising data held within the Distributed National Biological Collections Research Infrastructure. The investment (NCRIS, EIF) in this innovative bioinformatics system only partially fulfils current and future research needs. In part, the ALA has successfully created systems that provide new and exciting ways to examine aggregated collections data; yet the amount of data available is limited. Many specimens are yet to be curated, digitised and made accessible to the ALA and the wider research community. We believe an important component of the next phase of investment will be targeted data capture that will enhance the systems created to date.

It is also clear from the list of funded facilities that synergies exist in several areas such as between TERN, ALA, IMOS and AusScope. This needs to be explored further and opportunities created for integration and interoperability. In particular links between ALA and various 'omics facilities would benefit from increased linkages particularly as in the long term the 'omics capabilities need to be linked back to verifiable species records or vouchers.

What are your views on the eResearch infrastructure identified, including their relative priority and ability to support the current and future eResearch infrastructure needs?

CABC agrees with the views of the EWG and the working groups of other Capability areas that managing the expansion in data available represents a major challenge. It is clear that substantially better data storage and computing capabilities to process rapidly escalating data sets, including those harvested or generated by the ALA and partner institutions, will be required in the next round of investment. Storage, backup and delivery of image data in particular to the ALA still present significant challenges for all institutions and this is an area of expansion.

What are your views on the cross disciplinary requirements identified including their relative priority and ability to support the current and future research needs?

CABC agrees that cross-discipline datasets, particularly those that allow biodiversity data to be put into the context of other data layers such as from earth science (modeling biodiversity data against geospatial layers) and social research (characterizing societal attitudes and responses to biodiversity) are of high importance. We note that development of within-country capacity in this area should not compromise access to outside-country data – biodiversity, particularly marine biodiversity, extends well beyond our political and territorial boundaries.

We also agree that the facilities provided through BioPlatforms Australia, particularly in soil metagenomics and ecogenomics, will greatly stimulate environmental research and should be a high priority. We reiterate the point made above that such research must be linked to authoritative, vouchered reference material (physical specimens, vouchered tissue samples and vouchered DNA sequences) in collections.

Are there other programs/issues/developments not listed that you consider could be a driver for future research infrastructure investments or may impact on such investments.

As noted above, CABC regards that taxonomic capacity – the ability to name, resolve, discover, identify and audit biodiversity, and hence to provide accurate, authoritative and comprehensive data on the nation's biota – is an underpinning framework science upon which virtually all other biodiversity and many environmental sciences depend. Maintaining taxonomic capacity is an infrastructural issue that potentially affects the success of investment in many disciplines contributing to an environmentally sustainable Australia. Existing and planned investment through NCRIS, such as into the ALA, TERN, OBIS and BPA, provides a unique opportunity, and a critical need, to substantially upscale taxonomic capacity in Australia, arresting and reversing its current decline. Considerable international and Australian effort (e.g. TRIN and the European EDIT[5] program) is investing in mechanisms and processes for eTaxonomy – the intersection of taxonomy with ICT, modern genetics and social research, to streamline and 'turbocharge' taxonomic productivity. We encourage the EWG to consider appropriate mechanisms for such investment, some of which are outlined elsewhere in this document.

References

ABRS, 2006. *ABRS Survey of Australian Taxonomic Capacity 2003/2006 Part b resources of Australian Natural Science Collections*. (Accessed 2011_05_01; http://www.environment.gov.au/biodiversity/abrs/publications/other/taxonomy-survey-2003/index.html)

Bebber D.P., Carine M.A., Wood, J.R.I, Wortley, A.H., Harris, D.J., Prance, G.T., Davidse, G., Paige, J., Pennington, T.D., Robson, N.K.B, Scotland, R.W., 2010. Herbaria are a major frontier for species discovery. *Proceedings of the National Academy of Sciences* 107, 22169–22171.

European Roadmap for Research Infrastructure: Roadmap 2008. European Strategy Forum of Research Infrastructures (ESFRI), ISBN 978-92-79-10117-5 pp. 1-100. European Communities 2008.

Hartley, C.J., Newcomb, R.D., Russell, R.J., Yong, C.G., Stevens, J.R., Yeates, D.K., La Salle, J., Oakeshott, J.G., 2006. Amplification of DNA from preserved specimens shows blowflies were preadapted for the rapid evolution of pesticide resistance. *Proceedings of the National Academy of Sciences* 103, 8757–8762.

Johnson, K.G., Brooks, S.J., Fenberg, P.B., Glover, A.G., James, K.E., Lister, A.M., Michel, E., Spencer, M., Todd, J.A., Valsami-Jones, E., Young, J.R., Stewart, J.G., 2011. Climate change and biosphere response: unlocking the collections vault. *Bioscience* 61, 147-153.

Pitcher, C.R., Doherty, P., Arnold, P., Hooper, J., Gribble, N., Bartlett, C., Browne, M., Campbell, N., Cannard, T., Cappo, M., Carini, G., Chalmers, S., Cheers, S., Chetwynd, D., Colefax, A., Coles, R.,Cook, S., Davie, P., De'ath, G., Devereux, D., Done, B., Donovan, T., Ehrke, B., Ellis, N., Ericson, G., Fellegara, I., Forcey, K., Furey, M., Gledhill, D., Good, N., Gordon, S., Haywood, M., Hendriks, P., Jacobsen, I., Johnson, J., Jones, M., Kinninmoth, S., Kistle, S., Last, P., Leite, A., Marks, S., McLeod, I., Oczkowicz, S., Robinson, M., Rose, C., Seabright, D., Sheils, J., Sherlock, M., Skelton, P., Smith, D., Smith, G., Speare, P., Stowar, M., Strickland, C., Van der Geest, C., Venables, W., Walsh, C., Wassenberg, T., Welna, A., Yearsley, G,. 2007. *Seabed Biodiversity on the Continental Shelf of theGreat Barrier Reef World Heritage Area. AIMS/CSIRO/QM/QDPIF CRC Reef Research Task Final Report*. http://www.reef.crc.org.au/resprogram/programC/seabed/final-report.htm

Porch, N., Jordan, G.J., Price, D.M., Barnes, R.W., Macphail, M.K., Pemberton, M., 2009. Last interglacial climates of south-eastern Australia: plant and beetle-based reconstructions from Yarra Creek, King Island, Tasmania. *Quaternary Science Reviews* 28, 3197–3210.

Proceedings of the National Taxonomy Forum 4–5 October 2007 Australian Museum. ISBN 978 0 642 56858 8. Published FAST/ABRS copyright FASTS.

Roberts, C., Ingham, S., 2008. Using ancient DNA analysis in palaeopathology: a critical analysis of published papers, with recommendations for future work. *International Journal of Osteoarchaeology* 18, 600–613.

[5] See http://www.e-taxonomy.eu/

Scientific Collections: Mission critical infrastructure for Federal Science Agencies. A report of the Interagency Working Group on Scientific Collections (IWGSC). ISBN 978-0-9819500-0-6.

Woodward, F.I., 1987. Stomatal numbers are sensitive to increases in CO2 from preindustrial levels. *Nature* 327, 617–618.

Zink, A.R., Reischel U., Wolf, H., Nerlich, A.G., 2002. Molecular analysis of ancient microbial infections. *FEMS Microbiology Letters* 213, 141–147.

Contact:

Professor David Cantrill

Chair, CABC

David.Cantrill@rbg.vic.gov.au

Tel: 03 92522301

Fax 03 92522413

Mob. 0438 543 271