Submission
2016 National Research Infrastructure Roadmap
Capability Issues Paper

Name	**David Williams**
Title/role	**Executive Director**
Organisation	**CSIRO (Internal Submission #16/561)**

This submission is a response to the request for feedback in relation to the National research infrastructure Capability Issues Paper, July 2016.

Prelude Commentary - General Comments as a preamble to questions 2,3,4,8,9,10.

The arguments around creating a piece of National Research Infrastructure need to take a **holistic** or "whole of life" approach to the decisions that **build, govern, maintain, enhance, enable access, review and decommission** a particular class of Infrastructure. "Whole of life" covers initial and ongoing capital, operations and ensure the necessary human capability is available to ensure a facility is "science ready". CSIRO regards the "science ready" status as critical to ensuring external scientists can maximise the profitable science capability of an infrastructure, and that the infrastructure itself is maintained and developed as a world class capability. The holistic approach is needed because of the interlinking and co-dependency between the elements. Furthermore, no institution or governance process wants to support the capital for a facility to see it "stranded" due to lack of operations funding or the lack of relevant technical and engineering expertise.

However, in an evolving and ever more digital world, even defining what would be classed as a National Infrastructure is a challenge in itself. Traditionally, it has referred to large physical facilities operated by one institution that are available on a defined access basis to the whole research community, because of the practical limitation on the number of scientists who can use the Infrastructure in a given time period. Scale and complexity have been criteria to define a small number of National Infrastructures that provide an ongoing long term need for the research community. It was on this basis that the Publicly Funded Research Agencies (PFRAs) in their 2016 submission defined three classes of infrastructure namely: Landmark; National; and Institutional. (Annex 1). These definitions are being challenged and CSIRO has recognised the following type of facility that could be considered a National Infrastructure:

- **A: Traditional Large scale and cost Physical Infrastructure.** The PFRAs in their 2015 submission to the Clark review identified these as **Landmark** Facilities. The Research Vessel, Supercomputers, Australian Animal Health Laboratory (AAHL), Square Kilometre Array /Australian Square Kilometre Array Pathfinder, the physical bio collections, the OPAL reactor and the Synchrotron all fit this category.
- **B: International Facilities that Australia have the ability to access.** These could be major infrastructure e.g. Square Kilometre Array (SKA) or a collective network infrastructure that gives Australia access to a whole system in return for partial involvement, such as the Integrated Marine Observing System (IMOS) program of Buoy and Argo float deployment.

- **C: Infrastructure as a network of individual facilities.** All PFRAs and universities own bespoke equipment that could be made available through networking to the wider community, access to 3D laboratories is an example. Collectively through a formal network mechanism these could be seen as a National Infrastructure in the same way that data sets are.
- **D: Software models as an Infrastructure; Data and Networks as an infrastructure.** In a data led world, the concept of national infrastructure not requiring to necessarily be a physical structure will become increasingly important in many fields. Software models are emerging as an Infrastructure; an example being the Australian Community Climate and Earth System Simulator (ACCESS) climate model. Digital data sets are a growing "asset" that can be considered as part of the National Infrastructure. All science is now creating larger datasets and CSIRO itself now has some 40 Petabytes of data that are in an accessible form for researchers.

CSIRO believe that government decisions to fund, build and operate National Infrastructure (or participation in international equivalents) will continue to be made on a case-by-case basis, judged against an underlying need of both national science priorities and/or national benefit assessment. Each case will require its own argument but the key policy issue is to ensure that the initial acquisition of the infrastructure is not considered in isolation of ongoing operational requirements; it is essential that a "whole of life" or holistic view is taken to ensure a decision results in a "science ready infrastructure" with clear and transparent governance and access rules.

Meeting this challenge was part of the basic rationale for CSIRO choosing in 2015 to manage the National Facilities (that it has operated over a long period) as a single business area, in order to provide a holistic view of both capital and operating requirements. These facilities (AAHL, Marine National Facility, Pawsey Supercomputer, Radio Astronomy and Bio Collections) are all owned and managed by CSIRO on behalf of the nation and have defined access regimes that are open to the entire community. Each has its own characteristics, but all are maintained as "science ready", and are available to the wider community on defined access terms and free at the point of access.

Response to specific discussion points proposed in Issues Paper.

CSIRO specific input to questions 2 to 14 are all set in the context of a need for the holistic approach.

Question 2: Governance issues.

It is not clear whether the proposed governance criteria relate to the decision making on which facilities to support, or the oversight of the facility once in place. All are reasonable, but assuming they are applied to decision making there are key criteria that need to be included. These are:

i. What novel science and/or innovation will the Infrastructure proposed address? CSIRO assumes that infrastructure does not need to be tied to explicit government priority areas - but that conversely, the research infrastructure required for priority areas is identified and prioritised.
ii. What is the transition plan for new infrastructure? Will it propose a finite life? If not, what is the whole of life proposal and how is funding foreseen?
iii. A mechanism for ongoing assessment of whether an infrastructure is meeting its goals.

Question 3 and 4: International facility approach.

Where appropriate a National Infrastructure should support access to international capability. Many areas of science require global connectivity to solve complex challenges. Some infrastructure will be beyond the capability of Australia alone, whilst others will demand collaboration.

SKA is one example that is being addressed today, and the work around the precursor activities and the establishment of the Murchison Observatory in WA shows how a national activity can pave the way for an important International capability being in Australia. Equally the implementation of SKA will involve the building of complementary SKA facilities South Africa. Australian science will benefit through access to all the data flows in return for the Australian contribution.

The CSIRO operated radio telescope observatory (ATNF) is open to international use based on merit review and at no cost. In return, Australia has similar access capability to overseas observatories such as the ALMA in Chile, and Hubble.

The AAHL is an example of an Australian facility that is part of a global reference capability, and at the same time used by overseas governments and industry on a fee basis.

In the environment world, the Argo ocean float system is one that gives Australia access to data from all Argo floats worldwide in return for its support of a subset of floats.

International facilities require the same rigorous whole of life assessment but should prevail when the analysis supports one or more of the following:

i. The science will be world class and preferably where collaboration would fast track innovation and scientific breakthrough on matters of global significance and urgency relevant to Australia;
ii. Where Australia is not a suitable geographic site for the facility;
iii. Where international reciprocal access gives Australia to a greater “whole”;
iv. The cost benefit of international partnership or access is better than building a National facility or would be beyond the scope of Australia to fund.

Question 5, 6, 7: Workforce skills, training and institutional roles.

The response covers all three questions.

Absolutely yes, and infrastructure development and support needs to recognise and cover these costs. All Infrastructure should be “science ready” and this involves ensuring suitable specialised expertise is required to ensure effective operations and so maximise the specific science research access. It should be recognised that science infrastructure can either be an “off the shelf” piece of equipment, or in itself a combination of research and engineering project that is on the edge of what is possible.

All significant infrastructure require the ongoing involvement of a core of practicing scientists to devote a percentage of time to understanding the equipment and its capabilities. An example of this is AAHL where core scientists collaborate with external partners and facility users to maximise usage, and develop new ideas or pre-clinical animal models that require high levels of biocontainment and biosafety. In addition, Infrastructure facilities require the correct mix of ongoing specialist technical, software and engineering capabilities to run the infrastructure, and ensure it is maintained. High cost and valuable infrastructure will lie idle without resident technical specialist in the host institution.

The responsibility to ensure the nation has the necessary skills requires the collective efforts of universities, TAFE's and research institutions. Research institutions should support the development of researchers and technical/engineering specialists who are trained and knowledgeable in the use of advanced science instrumentation and facilities, and provide ongoing training to stimulate outstanding performance.

In return the Infrastructure itself can provide a platform to provide training and skill development programs and access options suitable for students at all levels from secondary level through to PhD. The greatest benefit will be from programs with on-the-job/hands-on training embracing the full range of design, engineering, project management and digital/computational skills.

Question 8, 9, 10: Access regimes. Defunding or decommissioning. Financing models.

These three issues are interlocked, with both access and decommissioning having an impact on the funding model as part of the decisions to have government support for a National (or International) infrastructure. The basic principle to be used is that the access regime for a particular facility should be open and transparent.

A decision to finance a facility (Nationally or International) requires decision criteria based on one or more of the following principles:

i. Scientific excellence determination;
ii. National or Industry benefit for Australia;
iii. Potential co-investment by external sources in either/both Capital and Operations costs;
iv. "Whole of Life" costing of Capital, Operations, people and decommissioning. **Annex 2** provides more detail on what these include.

CSIRO believes that there are a range of categories and criteria for access regimes and each Infrastructure will need to be treated individually.

A. **Categories of access for those Infrastructure with limited access capability:**

i. Scientific merit. For merit based peer review access, the principle should be that a facility is "free at the point of access" and the user should be a *bona fide* researcher. Incoming researchers will still be responsible for the provision of additional new or specialist equipment facilities e.g. buoys for marine work. All resulting data from "free at the point of access" research should be placed in the public domain and publications in the science literature. A sub category of early career scientist allocation may be part of this process.
ii. National Interest Access. For example, AAHL is partially funded by DAWR to undertake a directed work program, and a small element as part of an international reference laboratory under the WHO. A determination of the percentage of time nominally allocated to this use should be determined in constructing the financial model.
iii. Investor access or fee paying, for Capital and/or Operations. The finance model needs to recognise this where accepted and apportion a percentage of usage that reflects the

investment. The allocation should take into account whether any outcomes are made available publically.

Any one piece of infrastructure can support one or all 3 models. However, to maintain competitive neutrality a true piece of national research infrastructure would not be created if a business could support a capability. Where a piece of infrastructure combines more than one role an important criteria will be to determine the “balance” of use in each role.

B. **Infrastructure with no access mode except for data.** There is a class of facility where access is through the resultant data, rather than, to the infrastructure itself. The IMOS program of Buoy and Argos float deployment is an example of this. The funding principles do not change. In this case the important issue is to ensure the data collected are curated and stored in an openly accessible form.
C. **Infrastructure where access is not limited.** Digital data sets and digital models which are a growing “asset” that can be considered as part of the National Infrastructure. In theory there are no constraints (other than network capacity) restricting access, and the debate needed is whether any other policy except the government “open data” access policy, should be implemented for Commonwealth funded datasets. Open policy would need to be consistent with protection of data for National interest and personal privacy reasons.
D. **International Infrastructure**. In this case access will be determined as part of the international consortia and an Australian decision to support would need to take into account whether the access regime meets Australian requirements before agreeing to fund and as an additional principle to the 4 set out above (see response to Q3 and 4).

Defunding should be seen as part of a review process of whether the infrastructure is meeting its goals. This needs to be planned taking into account lifetime expectancy as part of the Governance process. A decision to cease the operation of a facility may require restitution of the site to pre facility condition, and with a holistic approach, a cost estimate should be made if this is required. The CSIRO view is that funding should not be provided initially, but that the costs should be part of a decision to decommission. For all the other infrastructure types a decision to defund is essentially a “merit of science” and will have little decommissioning cost.

In all cases of decommissioning, adequate notice should be given so that high priority programs (including PhD projects) can be completed, and users can transition to other science fields and the data can be transitioned to another facility.

CSIRO recognises that there are emerging **disruptive funding** models where the private sector build and operate a piece of infrastructure and the research community pay for specific access. The AARNET data network is a current example where a not for profit organisation has been created to provide a closed education and research network capability. This is self-funded by a set of PFRAs and universities. Data storage and compute is already encroaching in this space with cloud based systems and common data centres being established and utilised. A primary concern in this approach is the increase in ongoing operational costs this results in, at the expense of periodical capital needs. How this will impact on future Government schemes such as LIEF, EIF, future NCRIS, and one-off major Capital has yet to be assessed.

Question 11: Standard and accreditation requirements?

This needs to be addressed on a case by case basis. Examples of where either accreditation or standards need to be addressed are:

i. **Safety and HSE**. The AAHL infrastructure has to adhere to strict safety standards and protocols as does the Opal reactor, that are significantly higher than a normal infrastructure.
ii. **Standards**. Software developments need to meet industry standards if they are to translate to business opportunity.
iii. **Ethics**. There are ethics associated with many procedures that involve animals, or relate to personal data.

The Research Data Alliance (RDA), Open Geospatial Consortium (OGC), International Organisation for Standards (ISO), and World Wide Web Consortium (W3C) are each developing technologies and standards relevant to research infrastructure. Membership of these cross-disciplinary research infrastructure organisations would benefit from being co-ordinated nationally. There are additional domain-specific international bodies which ensure consistency with relevant standards.

Question 12, 13, 14: International or global best practice

There are examples of practices internationally that result in frameworks for decision making on infrastructure. In reality there is no absolute measure of "best practice".

Health and Medical Sciences

Question 15, 16, 17: Emerging directions and research infrastructure capabilities for Health and Medical Sciences right? International research infrastructure collaborations or other issues.

Historically some 70% of infectious diseases in humans have migrated from the animal population and there is no rationale to see this reduce in the future. Responding to a highly infective avian influenza essentially would require Australia to be self-sufficient in testing, development and manufacturing of agents. AAHL is seen as a critical infrastructure in a coordinated national response program.

AAHL also acts to link Australia into a number of significant 'invitation only' global disease and health security lab networks. An example is the BSL4 Zoonotic Laboratory network (BSL4ZNet) funded by the Canadian Department of Defence that links AAHL with similar skilled laboratories in the USA (Centers for Disease Control and Prevention, Department of Homeland Security, and Department of Agriculture), Canada (Public Health Canada and Canadian Food Inspection Agency), UK (Pirbright and Public Health England), and the German Federal Research Institute for Animal Health (FLI lab) in Riems. BSL4ZNet member countries can leverage the network through information exchange, access joint science funding and member labs, train and develop a specialised global BSL4 workforce, share otherwise confidential information and intelligence, and organise global support teams from across the network to respond to disease outbreaks and emergencies.

AAHL hosts a number of microbial, animal and plant collections to provide a supporting platform for performing essential biosecurity work. These include a calicivirus naive rabbit population, transgenic chickens and a ferret population. This enables biogenome projects to be performed, which in turn allow for the continuation of important work on biological rabbit control, avian influenza and high risk infective agent investigations.

The maintenance of AAHL as a secure research laboratory will provide Australia with a built in "insurance policy" to be able to respond in times of crisis and increasingly as a laboratory to do basic science on the issues. The core capabilities of the laboratory are its large scale PC3 and 4 laboratories, the professional work force and the in-house capability to undertake infective agent identification from laboratory analysis to supporting regional surveillance. The laboratory is also able to undertake amplification of agents through laboratory models, mitigation of infective agent activity through investigation of novel interventions and acceleration of responses, and develop products through pre-clinical expansion.

CSIRO sees benefit in leveraging allied relationship developments, and collaboration with major international projects. One such example is the 100 Thousand Genome Project being driven as a top level Ministerial activity as a standalone entity for the uplift in national healthcare services in the UK. This is an exemplar of dataset or object driven management, secure cloud architectures and applied information assurance processes.

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?

CSIRO supports the general directions in this section, especially the dependence of Australia's productivity and wellbeing on the environment, and the ecosystems services they provide; and the need to understand and manage soils, water and carbon in the landscape.

6.1.1 Integration. CSIRO agrees some areas will see a shift from data collection and management to integration, modelling and prediction. There is a need to place more emphasis on an integrated national observing system across the land – air – oceans domains that fuses in-situ and remotely sensed observations with a national modelling systems (such as ACCESS) and is linked to an integrated data management system(s) that supports data curation, discovery and analyses.

CSIRO welcomes the acknowledgement of the critical role for Australia's national capability in the weather and climate prediction model system (ACCESS) in the Environment and Natural Resource Management capability domain. ACCESS is a key element of CSIRO's new Climate Science Centre, and its long-term strategic plan. Australia's weather and climate research community is now poised to transform this investment into a nationally significant platform for weather, climate and environmental research.

ACCESS, and its model components – especially the land surface scheme (CABLE) – play a critical role in driving integration across timescales (weather through climate) and across the elements of the Earth's climate system (atmosphere, oceans, land, cryosphere etc). Improved model performance and predictions will result from fusing models such as ACCESS with high quality observations across these domains, which are provided by national collaborative research infrastructure facilities such as

TERN and IMOS, and those provided by agencies such as CSIRO (GASLAB and ICELAB) and the Bureau of Meteorology (e.g. the Cape Grim Baseline Air Pollution Station).

In the marine and coastal domains, we need to continue to support the development of socio-ecological decision support tools, such as Atlantis. However these models are being seriously constrained by inadequate data streams and poor process understanding. Ongoing and enhanced observing, and process studies are equally crucial in these areas.

6.2.1 Atmospheric observations. Continued and enhanced atmospheric observations will drive better understanding of processes and models.

Observations of atmospheric composition (aerosols, greenhouse gases), radiation, and water use and carbon uptake by the biosphere (land, ocean) are important in our research to understand processes and improve weather and climate predictions. In combination with remotely sensed data they enable through models assessments of water availability, carbon budgets and greenhouse gas emissions; and the data needed to inform "State of the Environment" reporting.

The OzFlux facility (a part of NCRIS-funded TERN) provides radiation, water use (evapotranspiration) and carbon flux (net uptake and loss) in Australian ecosystems is critical, and is well positioned to undertake additional key measurements in a broader suite of ecosystems and landscapes (especially agricultural, urban, and coastal). Continued access to the international Weather monitoring network of observations is also important alongside the Cape Grim Baseline Air Pollution Station.

6.2.2 Marine environment. Key challenges for marine industries and the environment are around how to balance human activities and environmental stewardship, and how to better manage multiple and often competing sectors. The science challenges are enormous to deliver on the industry and community needs of the science, we need to:

i. Develop coupled socio-economic and biophysical decision support systems that explicitly deal with trade-offs and uncertainty;
ii. Develop new understanding of, and tools to deal with, cumulative impacts within and across sectors (going beyond simple linear overlays and additive assumptions to tackle the nonlinear and interactive nature of these impacts);
iii. Identify new cost effective and smart technologies and observational approaches to support existing and emergent industries, and regulators;
iv. Improve national marine data systems that bring together all data sets (government, industry and universities) not just rapidly and interactively but in such a way as to allow for information discovery and knowledge creation; and
v. Establish marine baselines and monitoring programs so that change can be assessed and impacts attributed (e.g. climate, impacts of oil spills etc.).

CSIRO recognises the critical and enduring role, such as that provided by the National Biological Research Collections, the Marine National Facility (MNF) and the Integrated Marine Observing System (IMOS) and understands the benefits for science and society that could result from:

i. Maximising the utilisation of the MNF to support integrated marine observations (IMOS program) alongside an improved coastal vessel capability with greater national coordination;

ii. Improved IMOS ecological capability through better in-situ autonomous and sensorised platforms (e.g. AUVs, enhancing gliders and Argo floats) to include biological sampling and further coastal and shelf observing capability; and
iii. Extending existing marine data systems to include all data sources, including collaboration across public and private domains.

6.2.3 Terrestrial systems. The statement “groundwater constitutes more than 95 per cent of Australia’s fresh water” is misleading as much of it is too saline for drinking or agricultural purposes, particularly in the internally draining arid and semi-arid regions where evaporation considerably exceeds rainfall. Australian groundwater currently makes up around a third of our total water consumption, although this varies from location to location.

CSIRO notes that there would be great benefit from a better integrated terrestrial – atmosphere observing research infrastructure to support:

i. Sound management of natural resources including water, carbon and nutrient resources for environmental and production benefits;
ii. Monitor, assess, and predict ecosystem responses to climate change and variability;
iii. Support disaster management and early warning systems needed to meet Australia's priorities in national security; and
iv. Ensure that Earth system models used to underpin Australia's policies and commitments to international treaties adequately represent Australian terrestrial ecosystem processes.

6.2.4 Solid Earth. Our understanding of the solid Earth is fundamental to many areas of societal well-being and the Issue Paper does not adequately cover the priorities required. In particular, mineral exploration and discovery, energy resources, sustainable basin bound resource management should be a focus. Data is particularly relevant to our understanding of the Australian continent where our knowledge base is razor thin and mostly limited to near surface information.

Infrastructure developments we see benefit from are:

i. Auscope building the tools and data infrastructure to extend our knowledge base at depth through a ‘geoscience square kilometre array’; and
ii. The UNCOVER national initiative. This will be a decade long research program funded from a variety of sources that will have significant research infrastructure requirements from drill hole access, to monitoring and imaging sensors, to data collation and delivery and specialised analytics – discovery.
iii. To complement UNCOVER, Australia needs to develop an integrated distributed network of geophysical and remote sensor deployments and geochemical sampling and analysis, that will form a geological telescope looking inward not outward and supporting research from applied mineral discovery, to new energy systems (unconventional gas, waste storage and geothermal) as well as the underlying fundamental research that facilitates our understanding of how the Earth works and how it supports life and society.

Question 19, 20: International research infrastructure collaborations or anything else to be included or considered in the 2016 Roadmap for the Environment and Natural Resource Management capability area?

CSIRO sees benefit in considering the following possibilities:

A national facility to future proof Australian innovation in environmental forensics. CSIRO, ANSTO and university partners have identified that there would be a benefit in creating a distributed national facility in high precision environmental measurement in order to future proof Australian innovation in environmental forensics. Benefits would accrue from leveraging:

i. On-stream leading technologies (e.g. stable and radioactive noble gas environmental tracers);
ii. A diversify of applications (e.g. addressing needs of national significance including emerging nuclear fuel cycle, existing legacy site risk assessment, and rehabilitation from industrial use and mining); and
iii. The training of a workforce-ready technical specialists so that the infrastructure can be fully exploited and maximum value for the nation realised.

A national facility for design and fabrication of ground-to-space remote sensing systems and DataCube processing. Australia relies 100% on data from foreign-controlled satellites with no control or sensors customised for Australian landscapes. In return for the data, Australia has provided Cal/Val sites across Australia, but these are underfunded. Australia has also made the core Datacube software open source, and is now striving to fund the development of a system capable of holding all data received from a new generation of satellites.

In the shorter term, ensuring Australia maintains its ability to develop Datacube and maintain Cal/Val sites across Australia will ensure we continue to access the satellite data streams. The issue remains that the ability of Australia to benefit from satellite data at no cost, is dependent on the on-going goodwill of those nations that build and operate the satellites. There is the opportunity for Australia to "have skin in the game" through Australian designed and built Earth observation capabilities and the associated ground calibration and high-end data processing system. Possible options range from a coordinated national program, which supports a small but important set of data into the global network, to joining the European Space Agency with the Australian contribution being in Earth observation.

Synthetic Biology (SynBio) is an emerging interdisciplinary field and disruptive technology about to revolutionise medical science to provide greater impact. It is the convergence of diverse domains of biotechnology, evolutionary biology, molecular biology, systems biology, physics, chemistry, computer engineering and informatics, electrical engineering, and genetic engineering, among others. SynBio involves the in silico design, fabrication, and construction of new biological parts, devices, systems, and machines, as well as the re-design of existing, natural biological systems, for useful purposes.

There are currently no DNA synthesis/laser sorting facilities in Australia, which is a critical emerging foundational technology underpinning advances in synthetic biology, and Australia would benefit from in-country access to such capabilities.

Water security – quantity and quality. Research infrastructure critical to water security provides (i) quantitative and predictive, at-scale capacity to monitor and detect change in the quantity, fate, transformation, interconnection and distribution of surface and groundwater resources and their constituents (including contaminants) and (ii) low cost pilot-scale water treatment, and resource and energy recovery capacity to recycle and reuse low grade waters, to bank water for future use.

Critical Urban Zone Observatories. Integrated systems understanding our living zone are crucial to managing natural and engineered environs, from the Great Barrier Reef or to city scapes. Linkages and measurements with advanced sensing systems at relevant scales drive understanding and application, and require demonstration site scales where infrastructure is in place for decades. A concern is that there is an assumption that urban environment research infrastructure already exists, thereby excluding it from the Roadmap. However the continued existence of Australia's building research capability cannot be assumed.

The transition to an urbanised society is acknowledged in CSIRO's original megatrends. Building and construction represents 8% of GDP, 16% of businesses, and employs 9% of the nation's workforce. Australia's 18 largest cities generate 80% of its GDP. The size and design of housing are major factors in determining energy and water consumption. Research into urban infrastructure can impact economic, employment, safety and sustainability aspects for Australia. Its exclusion from future research infrastructure planning, whether due to assumed capability or oversight, should be redressed.

Nuclear fuel cycle. Commonwealth and State-based interests in the nuclear fuel cycle are precipitating thinking and response at a global scale. Ambition for developing facilities in Australia to manage a significant fraction of the globe's high-level nuclear waste (used fuel)[1] are ambitious in scale (estimated combined capital and operating costs are A$145.3 billion and total revenue A$257 billion over the 120-year life under the baseline scenario)[2]. There is also a significant demand to find and characterise new water resources in remote areas of Australia to promote economic development through mining and mosaic-style agriculture.

New solutions in environmental measurement will be demanded by new industries to measure more directly and precisely. Emerging technologies, tuned to Australian conditions are already in development (ANSTO's tritium facility with improved detection limits, CSIRO's Noble Gas Facility, University of Adelaide's Atomic Trap Trace Analysis Facility).

Advanced Physics, Chemistry, Mathematics and Materials

Q 21, 22, 23: Emerging directions and research infrastructure capabilities for Advanced Physics, Chemistry, Mathematics and Materials? International collaborations or anything else that needs to be included.

CSIRO is in general supportive of the description of current capability and emerging needs. On **astronomy,** the SKA should be added to the potential new infrastructure. CSIRO also notes that Astronomy Australia Limited is a funding entity not a piece of infrastructure and should be removed, and that the Australian National Telescope Facility should more correctly be referred to as the Australia Telescope National Facility.

---

[1] Baseline scenario assumes 50% of accessible market or 138 000 tHM (tonnes of heavy metal) and 390 000 m^3 of intermediate level waste

[2] Source: Nuclear Fuel Cycle Royal Commission Final Report 2016

CSIRO also foresees transformation in certain sectors of manufacturing through **Additive printing and nano-particle production in kilogram levels**. Currently CSIRO operates “Lab 22” that offers a 3D printing capability and this should be added to the list of existing capabilities. This facility underpins the ability to develop new capability for Australian industry/business by providing: design to powder production; additive manufacturing; modelling of processing and paints; durability assessment; and accreditation.

There is scope for Australia to identify some relevant niche space science missions to planets within our solar system through the development of niche areas such as Cubesats instruments. NASA in its future space science missions is planning to utilise Cubesats and deploy them from a parent ship to make specific observations that are then transmitted back to Earth through the parent ship. This novel new approach to planetary science offers the opportunity for a low cost entry into space science instrumentation for Australia, building on our strong relationship with NASA through the Canberra Deep Space Communication Complex operated by CSIRO.

Understanding Cultures and Communities

Q 24, 25, 26: Emerging directions and research infrastructure capabilities for understanding Cultures and Communities? International collaborations or anything else.

Flexible, communicating digital ecosystems should pay attention to how and why data is obtained and used. This requires new infrastructure capabilities that can help coordinate and communicate the scientific, policy, or practical purposes of citizen science and other forms of data, and to set limits on what is considered reliable or unreliable information.

Citizen science is recognised as key to enhancing surveillance systems, and scientific observations of a range of environmental and medical conditions, and to assist individuals and communities to make better decisions. Advances in technology and the rise of citizens collecting data, has been supported by brokering technologies that enable large volumes of data to be processed and translated to help advance science, inform public policy, and influence thoughts and actions. Blocking, authorising, filtering and translating individual contributions is the key to maintaining trust in citizen science data. Criteria by which trust in this data is evaluated can be different and often needs to be reconciled. It is not enough to create digital technologies that create scientifically accurate data because local communities, patients and land managers often generate, legitimate and trust information that accords with local knowledge and understanding.

National Security

Q 27, 28, 29: Emerging directions and research infrastructure for National Security? International research infrastructure collaborations or anything else.

The definition of National Security would benefit from a broader definition to include potential dual use capability of defence infrastructure, for example, environmental intelligence infrastructure such as ocean prediction has also revolutionised maritime and civilian safety and retrieval operations.

9.2.2 Current capabilities and emerging capability needs – Cyber Security. CSIRO agrees with the observations in this section. The industry led growth centre has the potential to drive a capability development in national cyber skills and commercial growth, which will flow on to provide a broader uplift in national posture. The conventional "closed wall" vertical mechanisms of managing national cyber security, will need to broaden horizontally beyond the defence constructs, and will require industry and government collaboration to yield benefit, uplift in posture, knowledge and real-time applied resilience across both technology based delivery, and human governance/response capability. The security approach to the National Research Infrastructure should align with the National Cyber Strategy. https://cybersecuritystrategy.dpmc.gov.au/assets/pdfs/dpmc-cyber-strategy.pdf

9.3.2 Desirable new capabilities – Cyber security. CSIRO agrees with the assessment that Australia has disparate and subscale research capabilities in this field. CSIRO sees benefit in a federated approach to people, process and technology, driven from the Cyber Growth Centre, not just for innovation and product based growth, but also for real time operational application.

Other desirable new capabilities – Biosecurity and Water. CSIRO has identified its capability under the Biosecurity Section (9.1.1) in the response to questions under the Health Section in questions 15-17. This is a clear overlap of issues and serves to strengthen the case for the maintenance and expansion of the role of AAHL. Similarly the Section on Water Security (9.3.3) is also addressed in our response to questions 18-20.

Underpinning Research Infrastructure

Q 30, 31, 32. Emerging directions and research infrastructure capabilities for Underpinning Research Infrastructure? International research infrastructure collaborations or anything else.

The identified directions and capabilities are broadly correct. These facilities are at the core of Australia's push as an innovative nation, and underpin virtually all important current science research.

High Performance Computing (HPC). Maintaining peak supercomputing facilities, such as the NCI and Pawsey Supercomputing Centre, are critical and should be seen as complementary to internal computing facilities for primary compute power that the well-founded internal facilities of a research focused institute should maintain as part of its basic infrastructure.

However, the definition of Peak (Tier 1) HPC is traditionally defined as a compute capability that is in the top 200 globally. Given the pace of development, this means that Australia needs an ongoing mechanism to ensure it maintains a peak compute capability for the nation. CSIRO can see great benefit in supporting two HPC computers (currently NCI and Pawsey).

The objective for Australia is that at any one time, one of the two computers should aim to be in the global top 20 computers. Whilst ambitious, this is one of the criteria that is considered when trying to attract world class researchers to remain or come to Australia. To achieve this requires the maintenance of the two national supercomputers that are refreshed in capital terms on a time-scale that maintains one as being in the top 20 globally at any moment in time.

CSIRO supports a co-ordination of the capital investment cycles such that each computer is renewed on a 4 year cycle, and refreshed on a 2 year cycle. This can be achieved without any change in current governance. The current position in this respect is dire. The NCI has just received what is effectively emergency support to ensure its continuation beyond its current 4 year life. The Pawsey Board has agreed that the Pawsey computer requires funding for a capital replacement in 2017/18 to remain viable.

There may be benefit in moving to a single unitary capability for Peak (Tier 1) HPC. However, this would require a recasting of the operational funding model for facilities. This is because the current operational funding is different for each computer. And furthermore, each computer has a different mix of peer review and directed access, as a result of the funding model for operations. CSIRO is happy to provide logistical support for a new national governance approach.

Below Tier 1, CSIRO maintains its own Tier 2 capability which is a powerful compute environment, as do many Australian research institutions. Given this, CSIRO does not see an argument for continued Commonwealth investment and believes all research institutions should be able to maintain a relevant scale Tier 2 compute capability as part of their basic infrastructure.

To maintain competiveness with other nation states, Australian HPC will need to be an international activity that actively seeks out partnerships and collaborations within our region. CSIRO is being actively sought out, and is already in conversation with Singapore and South Korea. International HPC leaders such as the USA, European Union, and Japan also have a desire to foster greater linkages with Australia.

Increasingly as commercial cloud based compute capability develops, the ability to use cloud rather than "own" compute hardware will develop. Commercial cloud access does have significant benefit in terms of fixed costs and capital needs, in that costs are usage related and hence variable.

At this time there are potential barriers to cloud. One is that cloud compute is not at a level that can currently replace HPC, and hence for the foreseeable period cloud cannot replace HPC. A second is cost. Cloud compute costs can spiral very quickly in a "research" mode. Good management systems are needed to ensure this is not the case. A third is security, which needs to be considered for all activity in the cloud. Given these points, CSIRO believe cloud compute is an area to be considered for Tier 2 compute and not Tier 1 at this time.

CSIRO also sees the potential of quantum computing as a capability to make a significant step forward in HPC capability. Although not yet scalable, the development of this capability has great potential benefit.

AARNET. The AARNET internet network is constantly upgrading and providing additional bandwidth and is striving to support all education and research establishments. As a self-funded body by the institutions using the network, this can only be assured by relevant financial contributions, and CSIRO supports this approach.

Access and Authentication. This is a barrier to full and open cooperation between institutional partners without each partner having to recertify external people entering their systems. Setting aside normal security and commercial issues, this is critical in the light of the new export laws which mean institutions and individuals have legal obligations to ensure access to certain facilities and systems are controlled. For physical facilities this is relatively easy to control. For both compute and data access it is a more challenging issue (and could be considered under national security).

CSIRO recommends that the issue is addressed in the cyber security world to see if a common procedure for identification can be found. However for CSIRO, it is also driving the entire approach and control to our internal compute and data files, with a layered approach of security being considered.

Digitisation. CSIRO notes that the National Biological Research Collections are missing from the Table on page 26 and should be included as critical existing infrastructure. In the new world of "big data" the digitisation of the Collections will greatly increase their value to science and innovation, releasing the potential of Australia's biological and genetic diversity. Emerging genomics technologies present significant opportunities for research using the Collections as a resource with the potential to deliver impact in bio-based industries such as agriculture, fisheries, aquaculture, food, health, pharmaceuticals, manufacturing, and biosecurity. Some examples are:

i. Bio-prospecting for better crops by looking at wild varieties genomes e.g. soybean and cotton, and silk chemistry for materials engineering.
ii. Bio-security by enabling rapid species identification tools, and looking at close relatives of known "problem" species to assess the dangers.
iii. Bio-remediation by for example, identifying those sub species that are drought adapted genotypes for ecological restoration.
iv. Biodiversity policy and decision-making as in the marine protected areas planning e.g. Fishmap, PNG Kakoda biodiversity assessment.

CSIRO sees great benefit in a new data infrastructure through the digitisation of national biological reference data held in the National Biological Research Collections to position Australia at the very forefront of biodiversity eResearch infrastructure, utilising the Atlas of Living Australia as the baseline for open access by users.

Improving access and use of the massive number of physical artefacts and specimens held by CSIRO and others, representing Australia's unique biodiversity and physical attributes is essential. Enabling richer access to these collections through digitisation will significantly enhance the ability to use these resources for important new discoveries (potentially providing breakthroughs in sectors such as health, medicine, agriculture, resource industries, environmental remediation etc). CSIRO responded to the question of digitising its bio collections in Q18 and deals with the management of all digital data in Q 33 - 35.

Data for Research and Discoverability

Q 33, 34, 35 Emerging directions and research infrastructure capabilities for Data for Research and Discoverability? International research infrastructure collaborations or anything else that needs to be included or considered.

CSIRO broadly supports the recommendations of the 2015 Australian Research Data Infrastructure Strategy. In essence, CSIRO supports a policy that all publicly funded digital data sets should be available to researchers. There would be benefit in making research data published by statutory authorities (e.g. Bureau of Meteorology, GA, ABS, CSIRO) visible alongside research data. CSIRO however recognises that there will be certain limitations on data for issues such as national interest, personal confidentiality, and in joint funding with business.

CSIRO recommends the following be considered:

A. Key principles for nationally automated digital databases, are:
    i. Research data should be open by default;
    ii. Access to research data should be provided as close as possible to the source or 'point of truth';
    iii. Any embargo periods must be reasonable in duration and take account of the 'shelf life';
    iv. Data interfaces (APIs) should adopt community standards, protocols, and conventions;
    v. New standards or conventions should build on existing standards if possible, and must be fully and openly documented;
    vi. Technical *standards* are part of the research infrastructure, and their development should be funded; and
    vii. Well-governed common scientific vocabularies must be published in a way that makes them easily usable as linked data on the web. Development and maintenance of common vocabulary content must be through a transparent process. (ANDS hosts a technical platform, but has not addressed the associated governance issue.)
B. The 'big public databases' (Bureau of Meteorology, CSIRO, GA, and ABS) need corresponding 'big public interfaces'. Adoption of standard interfaces, and guidance to the user community on accessing these, are required.
C. Institutional and technical connections between government data providers and research infrastructure operators must be improved. Key blockages and costs at the interface between State/Territory governance. Current research infrastructure approaches tend to take on the cost of data aggregation and integration, and minimise costs for data providers. The costs of infrastructure operations can be minimised by addressing data interoperability challenges at the source, by working with data providers to establish common data exchange models and data access mechanisms.

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.

CSIRO has two comments on the capability areas:

The first is that Agriculture and Food should have greater representation in the Issues Paper and be considered at capability level. This is based on the importance of agriculture to Australia and the industries it support (2/3 of the continent is used in one form or another for agricultural production). CSIRO sees opportunities to consolidate and invest further into digital agriculture technology based infrastructure to improve efficiencies in farming systems and delivering productivity gains. The integration of new emerging technologies (e.g. full wave-form LiDAR, full range hyperspectral cameras) and the repurposing or translation of "phenomics" tools to farmers would require ongoing investments to enable rapid technology transfer to the industry. Opportunities exist for delivering "e-tool platforms" to farmers for real-time adaptive response to specific cropping situations. Similarly, a new capability for non-invasively measuring animal productivity as a function of animal movement and pasture consumption would deliver significant benefit to the industry and the Research community. Such infrastructure does not exist despite the fact that the technology are now available to enable it.

A second issue is to note the National Innovation and Science audit is being undertaken and may well have outcomes that are relevant to this review.

Annex 1 Categories of NRI

Title	Characteristics	Examples
Landmark	Multi-country, multi-user, multi-institutional partners, multidisciplinary or mission focussed Unique, very large-scale, complex facilities which are international in capacity and used to support research undertaken to address international and national strategies and priorities Funding arrangements developed on a case by case basis Budgeted within host's core budget Capital cost: $100m and above	ASKAP (a global example of the Square Kilometre Array) Australian Animal Health Laboratory Australian Synchrotron OPAL research reactor RV *Investigator* Super Computer Infrastructure Atlas of Living Australia collections*
National Facilities	Multi-user, multi-institutional partners, multidisciplinary or mission focussed Unique, large-scale, complex facilities which are used to support research undertaken to address national and regional strategies and priorities Hosted by an organisation(s) on behalf of the research community which accesses and utilises the facility Budgeted separately from hosts core budget Priority established by broader research community Capital cost: $10M - $100m	AARNet Integrated Marine Observing System National Deuteration Facility Nanofabrication Facilities National Imaging Facility National Sea Simulator Other NCRIS facilities
Major Institutional	Predominantly for internal institutional usage Research focus set by institution, with a strategic purpose ranging from specific research sectors or locations to broader national strategies and priorities Budgeted as part of institutional budget Institutional governance Capital cost: $10 million - $50 million	Coastal Research Vessels General Laboratories Industry funded facilities Specialled industry sector specific laboratories

Annex 2: Whole of Life Funding

Investment in new research infrastructure requires the establishment a whole of life costing to include: of funding principles that are assessed at the point of investment decision. In

(i) ***Capital funding***: This should take into account all sources of funding including co-investment from the entity that will own and/or operate the asset as well as from third parties (government and non-government). It should cover ongoing maintenance capital, refurbishments and the replacement of major components.

(ii) ***Operating funding provided to operate and maintain the asset*****,** should be at a level to allow the asset to be operated optimally: This would include expenditure for the day to day operations and maintenance of the asset, including necessary scientific, technical and engineering staff. This should take into account all sources of funding including contributions from the entity that will own and/or operate the asset and from third parties (government and non-government), including user charging arrangements.

(iii) ***Make good and decommissioning***: While accounting standards may require a liability be recognised for make good and/or decommissioning expenses, funding should not be provided initially. It should be noted that a future financial commitment will be created and that funding will be considered as part of the independent review if the review recommends the asset be decommissioned.

(iv) ***Depreciation / Replacement Funding***: CSIRO recognises this as a problem today, and sees benefit in an approach that would recognise assets as owned by the Commonwealth Government. Thus the operating institution would need an operating loss approval to cover the depreciation expense and this should be reflected in the NPP.

(v) ***Review***: The timing of an independent reviews of the infrastructure should be formally agreed. The outcomes of the review should be used to inform decisions on the continued role of the infrastructure, the continuation of funding or the provision of decommissioning and/or make good funding.