POLICY BRIEF #1

THE FUTURE OF AUSTRALIAN EUROPEAN E-RESEARCH COOPERATION

March 2019

EPIC

Europe's ICT innovation partnership with Australia, New Zealand and Singapore

www.epicproject.eu

info@epicproject.eu

@EPIC_ProjectEU

Recommendations

Australia and the EU should jointly:

* Develop openness in research based on their existing efforts. This requires the collaboration of government, industry and academia.
* Develop collaborative funding models in the field of e-Research to leverage existing resources.
* Develop virtual labs that go beyond open data.
* Investigate options for developing marketplaces for open data.
* Investigate opportunities to enhance e-Research in the humanities, the arts and social sciences.
* Develop robust data management plans to facilitate good practices by researchers in their regions, and also worldwide.
* Evaluate the economic benefits of e-Research.

SUMMARY

Science and research are increasingly digitally-driven and digitally-dependent. The acceleration of wide-spread sensorisation, coupled with large databases of historic data, promises new opportunities for researchers to utilize a wealth of data. A new generation of young researchers is eager not only to exploit these resources, but also to openly share their data, tools, and solutions with peers and the public. Australia and the EU have already collaborated in key e-Research domains such as biological science, astronomy, ocean and climate research. The European Commission currently finances the Mesopp project[1], for example, which delivers good scientific results on ocean e-infrastructures. Continued mutually beneficial exploitation of e-Research opportunities will require new approaches which go beyond open data and new sources of funding, and more nuanced assessments of the benefits of open and digital research.

INTRODUCTION

The digitisation of science has arrived

An expanding set of digital tools and improvements to globally accessible digital infrastructure are changing industry practices as much as citizens' lives. Science has been particularly quick to adopt digital methods and formats in its work practices. This includes data, as well as discourse, tools, and methods; even funding and infrastructure may be accessed digitally.

This digitisation of science is a major driver behind the internationalisation of research, leading to a workforce shortage of individuals with appropriate skills. Such internationalisation is not a luxury. Rather, it has become a necessity, in particular for industry as innovation and knowledge for innovation are sourced globally. However, it is often plagued by significant practical challenges such as travel costs, time zone differences, working styles etc. A recent round of panel sessions on EU-Australian research cooperation identified key trends for future internationalisation of e-Research with a focus on long-distance collaboration.

Targeted e-Research developments in combination with good e-Research processes and infrastructure will facilitate easier and more productive cross-discipline and therefore cross-silo collaboration.

Support for e-Research is seen as a significant investment in Europe. Europe's Digital Single Market[2] policies include specific actions for an improved e-Research environment.

DRIVERS OF E-RESEARCH

The motivation behind e-Research activities cannot be limited to "more open" processes, or "data re-use". Especially in Europe and Australia, three key drivers for agencies to fund e-Research are to (i) improve research quality, especially the reproducibility of

[1] http://www.mesopp.eu

[2] https://ec.europa.eu/digital-single-market

research, (ii) enhance the use of more effective research results beyond academia, i.e. research translation, and (iii) increase the productivity of all aspects of research activity.

The effectiveness of these motivations is dependent on the extent to which they are acknowledged when data infrastructure is first established. Simple provision of research data in a repository is not enough – thought must be given to the potential usage of the data.

Applications in climate research are at the forefront of international cooperation in e-Research. Similarly, satellite imaging has proven an important driver of Europe's cooperation with Australia, in particular since the availability of Copernicus[3] data for Australia.[4]

Today, much effort goes into adapting existing data for a specific need, often resulting in individualised and non-standard solutions. A possible way forward, in particular for EU-AU cooperation, is the creation of virtual labs. Such labs provide not only access to data but also tools that facilitate virtual experimentation, and are characterised by clear rules, ease-of-use, and data curation.

Some jurisdictions and research funders have moved to supporting openness in research. For example, the EU's new Open Science Cloud[5] provides an opportunity to improve international engagement in initiatives that support collaborative and open innovation. However, making data findable, accessible, interoperable, and reusable (FAIR) across country borders remains a non-trivial challenge which needs to be tackled in close cooperation across jurisdictions. The future lies in strengthened collaborations between academia, industry, and government agencies. At first sight, e-Research and open data/open science approaches often focus on quantitative aspects of "more", "better", "faster", but there are other reasons for supporting e-Research:

* ***Making the best of available data***

Several fields of science have been transformed from being non-data driven to highly data-driven. Biology, for instance, is an example of a scientific field that is experiencing an explosion of data. The advent of massively parallel sequencing techniques, and other high-throughput genomic and proteomic approaches, has dramatically increased the amount of data available to biologists. However, many biologists are still developing the skills to make the best use of this available data and rely on collaborative interaction with bioinformaticians to analyse and interpret these large data resources.

* ***Helping researchers benefit***

For researchers, citation rates increase dramatically when they publish their computer code and data alongside their papers. As increased citations lend impact to a researcher's work, this provides a personal motivation to embrace e-Research when appropriate tools and resources are made available.

* ***Improving research quality***

Another reason to invest in e-Research is to drive up research quality. Using international data repositories has become an important driver of research quality, for example through competitions. Also, it helps to establish factual standards which can support the development of common approaches.

* ***Exploiting the translational power of e-Research***

Governments have an incentive to increase the use of data beyond the original research environment, i.e. translation into practice within the public and private sectors. Assistance with the transformation of research is needed to achieve this.

The Australian Government has made, and continues to commit to, significant investments in e-Research.[6] The National Collaborative Research Infrastructure Strategy (NCRIS) provides a national network of world-class research infrastructure facilities that support high-quality research with the aim of fostering greater innovation in the Australian research sector and the economy more broadly. Similarly, Europe's new Open Science Cloud aims to create a virtual environment with services for research data that are open, seamless, and free at the point of use. It seeks to combine Europe's data infrastructure with high-performance computing resources and targets 1.7 million EU researchers and 70 million science and technology professionals in data-driven science. In addition, various investments by Europe's H2020 programme targeted Open Science and the provision of open research data.[7]

Current challenges

There is increasing demand from researchers to include international research data in their own projects with varying degrees of integration. This presents numerous practical, but also organisational, challenges which are currently being addressed through largely individual and non-standard solutions. Unfortunately, collaborative international funding models of significant size and reasonable simplicity are difficult to find. Although there are many initiatives and working groups, funding for cross-border and joint open e-Research schemes is scarce. Even where collaboration and open access are focused, commercial use of

[3] http://www.copernicus.gov.au/

[4] The Australian federal government's data explains Sentinel features in more detail: https://blog.data.gov.au/news-media/blog/near-real-time-satellite-images-show-us-real-australia

[5] https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud

[6] See Australia's Open Data platform: https://data.gov.au/data/ and the corresponding funding opportunities for Australian organisations: https://www.pmc.gov.au/public-data/open-data/platforms-open-data

[7] https://ec.europa.eu/programmes/horizon2020/en/h2020-section/research-infrastructures-including-e-infrastructures

data may come with added administrative burdens and costs.

Human aspects play an important role in the further development of e-Research. This ranges from research in humanities, arts, and social sciences as disciplines with specific challenges for e-Research, to the human aspects in international research cooperation.

Integrated platforms in the domain of humanities, arts and social sciences provide a broad range of challenges. Research in these fields sometimes results in disparate and non-standardised large data sets such as those that exist in disciplines like psychology, sociology, political science, and humanities (e.g. archaeology, linguistics, history). To fully tap into the potential of open e-Research resources for innovation, much more cooperation in these fields will be needed.

Human factors also include the necessity that collaborators from around the globe understand the motivations, objectives and environments of their colleagues. For this, there is no substitute to conducting meetings and work in person. Despite substantial progress in collaborative tools, there is little consensus today on which of these tools to use – from data sharing to word processing to video conferences – especially when working across sectors and continents. Therefore, and for the time being, meeting personally can and should remain a key element, enhanced by the use of digital tools, to ensure efficient international long-distance collaboration.

Digitisation of research is opening science not just for scientists, but also for a global public. This enables more widespread access to scientific publications, and also goes much further in making available data, software, even infrastructure together with instructional videos and tutorials. This is a transformative step for scientific knowledge from being a “club good” to a truly public good.

POLICY IMPLICATIONS

This increased digitisation is changing the relationship between science and society and, consequently, altering science policy. The boundaries of what is to be considered a national research policy are much less clear today than just a decade ago. Policy makers with an interest in sustaining national competitiveness are faced with large multinational players – in academia and industry – who opportunistically source research and innovation globally. E-Research and the digitisation of science therefore need to be integral in the context of domestic policies that address international collaboration in research and beyond.

The intrinsic complexities of international interoperability require an inclusive approach in which all relevant actors from research/academia, government and industry are involved. The specific challenges of not-for-profit versus commercial interests require an open-minded approach. Often today, we are lacking nuanced data that clearly demonstrates the mutual benefits of open and collaborative approaches in e-Research. This is particularly true for cross-border cooperation.

#1 Joint open data and e-Research initiatives

The future lies in strengthened collaborations between academia, industry and government agencies. International cooperation provides a particular challenge and has become a necessity. This includes work on standards for data, services and infrastructure to ensure access, interoperability, and ease-of-use along with the development of supported services. Policy coordination (bilateral, multilateral and global) around certain data collections (for example, national surveys) would improve the reusability of such data and coordinating efforts around rewards for the reuse of software may lead to increased efficiencies beyond the reuse of data.

#2 Establish a collaborative international EU-AU funding model

There are clear benefits of open research data and e-Research, and most agencies and research policy makers have astutely become aware of this. However, it remains challenging for Australian and European researchers to initiate e-Research activities due to the lack of fit-for-purpose funding models – with only few exceptions. Easily applicable funding models for realising EU-AU e-Research projects should become a primary target of future EU-AU cooperation.

#3 Value beyond research

In order to establish longer-term collaborative funding models, it will be necessary to provide evidence of the benefits of e-Research – both scientifically and economically. For example, it may be possible to undertake research around the GDP benefits of wholly open data, as opposed to commercial or free for non-profit arrangements.

#4 Joint virtual labs

A trusted data repository is not simply a place to store data. Making data open does not necessarily make it reusable. A possible way forward, in particular for EU-AU cooperation, is the creation of virtual labs. Such labs would provide not only access to data, but also tools that support virtual experimentation and are characterised by clear rules, ease-of-use, and data curation. Virtual labs eliminate the users’ infrastructure costs of using/visualising the data. Such labs may involve payment for dedicated services, but this in turn opens the possibility of adding quality services on top of the data.

#5 Test new ideas of common interest

Robust workflows are needed to involve users in the design of e-Research platforms, so they are fit for use. Data needs to be evidenced and quality-assured to create trust and a solid basis for further exploitation. This is the first step for research data to become a commodity in trust-building initiatives: processes, policies, paradigms, and ways of working. A joint initiative to test the application of robust data management plans for research endeavours has the potential to assist researchers in their data management activities, but also in improving the often poor quality of current data management plans.

#6 Create marketplaces for data

Using open data can be very expensive, for example because of the lack of quality, metadata, interfaces, interoperability, etc. A data marketplace can provide an incentive to make data available, acquired and utilised. Use of data from such marketplaces would necessarily attribute public recognition to this use, and to the original provider. It could work like a stock exchange, which is simply a registry of input and output transactions. Although some data marketplaces already exist, they are by no means common. Most of the necessary work towards such marketplaces is around policies and procedures relating to shared areas of interest, from the medical domain to climate data.

The EU and Australia could provide trials of dedicated data markets. It takes time and money to clean and package the data in order to render it fit for use. In many circumstances, it may be more efficient to package the data and charge for this data-as-a-service. In addition, publishers need to find their role. They still rarely apply quality assurance for data that is underlying publications. Europe and Australia are well placed to experiment with such marketplaces at an international level.

#7 Identify application areas of shared international importance

Australia and the EU already have e-Research ties in a broad range of application domains from astronomy to biology, genetics, climate and medical research, and music. Further collaboration in recognised areas of international cooperation is recommended. For example, the United Nations' Sustainable Development Goals (SDGs) provide an outstanding frame of reference for the development of e-Research collaborations. The SDGs can be used as a guideline for initiating thought around policy and legal frameworks and needs e.g. socially diverse data.

Further References & Links

A. Treloar, Design and Implementation of the Australian National Data Service. Presented at the 4th Int. Digital Curation Conference, December 2008. Available at https://www.researchgate.net/publication/238741036

European Open Cloud Initiative – Building a competitive data and knowledge economy in Europe. COM(2018) 178 final. http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=15266

HPC declaration on HPC. http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=43815 (accessed 8/Nov/18).

NCRIS, The Australian Research Data Infrastructure Strategy. Available at https://www.education.gov.au/ncris

Proposal for a Council Regulation on establishing the European High-Performance Computing Joint Undertaking. COM(2018) 8 final. http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=49292

How to cite

When quoting information from this report, please cite the following: „EPIC consortium (2019).The Future of Australian-European E-Research Cooperation. EPIC Project policy brief #1."

This policy brief was prepared as part of the EU project EPIC. The views expressed in this brief are solely those of the authors and not of the European Commission or its services.

Europe's ICT innovation partnership with Australia, New Zealand and Singapore

www.epicproject.eu

info@epicproject.eu

@EPIC_ProjectEU

This project has received funding from the European Union's Horizon 2020 research and innovation programme (ICT) under grant agreement No 687794.