



The University of Sydney & Intersect

Co-developing eResearch infrastructure:

Technology-enhanced research practices, attitudes and requirements

Summary Report

Research team: L. Markauskaite, J. Richardson, L. Hellmers,
M.A Kennan, J. Brattan, G. Pearson, A. Aditomo, F. Stevens

Report prepared by Lina Markauskaite, Anindito Aditomo and
Leonie Hellmers

2010-01-22 V1.1

Summary

This technical report presents descriptive results from the first stage of the survey that investigates existing technology-enhanced research practices, researchers' readiness to adopt eResearch, their needs and major barriers. This survey was conducted as a part of a larger project which aims to establish researchers' requirements for eResearch infrastructures and support, build researchers' awareness about eResearch potential, and engage with them to co-develop eResearch services.

The survey was conducted at various intervals between 4 May–2 July 2009 at four NSW universities: The University of Sydney, The University of Newcastle, The University of New England and The University of New South Wales.

It covered three main eResearch areas: a) data management, retention and sharing; b) technology enhanced research methods, tools and services; and c) research collaboration and dissemination.

The questions focused on several aspects, including researchers' present ICT use for research practices; their attitudes to, and awareness about, eResearch; barriers for broader use of eResearch; researchers needs for support and willingness to be involved in co-development of eResearch infrastructure. This technical report presents descriptive results and participants answers to all survey questions.

Acknowledgement

This study reported in this report was initiated and part-funded by University of Sydney ICT (Tools and Frameworks for Research Collaboration project) and Intersect Australia Ltd. The project is based at the University of Sydney, Faculty of Education and Social Work, Research Centre for Computer Supported Learning and Cognition (CoCo). We would like to thank DVCs/PVCs for Research of all four participating universities for their cooperation, and all participants of this survey for their time. The survey is running at four other NSW universities currently. Further universities interested in doing this survey are also very welcome. Contact: Lina Markauskaite (University of Sydney, CoCo Research Centre, l.markauskaite@edfac.usyd.edu.au) or Leonie Hellmers (Intersect, leonie.hellmers@intersect.org.au)

Research team

- Lina Markauskaite, Jim Richardson, Frankie Stevens, Anindito Aditomo, University of Sydney
- Leonie Hellmers, Intersect
- Mary Anne Kennan, Charles Sturt University
- John Brattan, University of Newcastle
- Graham Pearson, University of New England

Table of Contents

Summary	2
Acknowledgement	2
Research team	2
Table of Contents.....	3
Study background and rationale.....	4
Approach and procedure	4
Summary of findings	5
1. Participants' background	5
2. Research tools and methods.....	6
3. Research collaboration and dissemination	7
4. Research data.....	7
5. Data sharing practices.....	8
6. Data management and preservation	8
7. eResearch support, training, and services	9
8. Attitudes towards digital technologies in research	10
9. Further information and willingness to be involved in elicitation of eResearch requirements ...	10

Study background and rationale

Over the last decade, individual universities and governments in various countries, including Australia, have made significant commitments and investments in developing advanced technological infrastructures to support research – in short, eResearch¹. These investments have been fuelled by an explicit expectation that eResearch infrastructures will be taken up by broader research communities and will enhance the nature, quality and efficiency of research. Nevertheless, users and developers of eResearch infrastructures and services face numerous challenges embracing and scaling up eResearch.² What motivates researchers to embrace ICT-enhanced research approaches? What prevents them from adopting eResearch? What kinds of support do they need? The answers to these questions are central for making informed decisions about how one could improve eResearch uptake and maximise the benefits of research infrastructures. As a first step, this study aimed to investigate the role of technologies in current research practices; what kinds of challenges researchers typically face; and what kinds of eResearch solutions and support could be most beneficial for them.

Approach and procedure

This study was conducted using an online survey, which covered three main eResearch areas:

- 1) Data management, retention and sharing;
- 2) Technology-enhanced research methods, tools and services;
- 3) Research collaboration and dissemination.

The questions focused on four aspects:

- 1) Present practices and barriers for eResearch;
- 2) Priorities and requirements for new infrastructures, services and support;
- 3) Attitudes and awareness about eResearch; and
- 4) Willingness to be involved in future elicitation of needs and specification of requirements.

¹ DEST/DCITA, *An Australian eResearch strategy and implementation framework: Final report of the eResearch coordinating committee*. 2006, Australian Government, DEST/DCITA. [URL](#)
NCRIS Committee, *Review of the National Collaborative Research Infrastructure Strategy's Roadmap*. 2008, Australian Government: DIISR. [URL](#)

² Ribes, D. & Finholt, T.A., Tensions across the scales: planning infrastructure for the long-term. In Proceedings of the 2007 international ACM conference on supporting group work. 2007, ACM: Sanibel Island, Florida, USA. [URL](#)
Edwards, P.N., Jackson, S.J., Bowker, G.C., & Knobel, C.P., Understanding infrastructure: Dynamics, tensions and design. Report of a workshop on "History & theory of infrastructure: Lessons for new scientific cyberinfrastructures", 2007, NSF Grant 0630263, Office of Cyberinfrastructure. [URL](#)

In total, participants were asked to respond to 40 questions, most of which required them to choose from a range of options, and allowed a short comment. Eight questions asked participants to provide open narrative answers. The survey instrument included the following sections³:

1. Background information
2. Research tools and methods
3. Research collaboration and dissemination
4. Research data
5. Data sharing practices
6. Data management and preservation
7. eResearch support, training, and services
8. Attitudes towards digital technologies in research
9. Further information and willingness to be involved in elicitation of eResearch requirements

The online survey was conducted in 4 May–2 July 2009 at four NSW universities:

- The University of Sydney
- The University of Newcastle
- The University of New England
- The University of New South Wales

Email invitations were initially distributed via Deputy Vice-Chancellors Research inviting all academic staff, research students and research support staff to participate. Further invitations and reminders were sent directly to some interest groups, centres and faculties within universities. None of the invitations targeted other institutions, however researchers from other institutions were allowed to access and complete the survey. After the survey closed, the data was downloaded, anonymised and imported in SPSS. Multiple-choice answers were analysed using descriptive statistical analysis. Total and valid percentages are reported in this report. Percentages reported in the summary are based on valid answers. Anonymous answers were sorted in the alphabetical order and presented in this technical report in original format. Summaries of the answers to these questions are based on preliminary interpretative reading of responses.⁴

Summary of findings

1. Participants' background

1. Sample size. In total 658 participants took the survey and 537 (82%) of them completed it to the end. Percentages reported in this summary are based on valid answers, ie those participants who completed most of the survey.

³ Research instrument was developed in close collaboration between the University of Sydney ICT unit (and members of the Tools and Frameworks for Research Collaboration project) and Intersect staff.

⁴ More detail analysis of qualitative responses is planned in the next stage of the analysis and will be reported separately.

2. Affiliations. Survey respondents were from the following institutions: 29% from UNSW, 26% from the University of Sydney, 17% from the University of New England; 16% from the University of Newcastle, and 12% from other institutions or did not indicate their affiliation.
3. Disciplines (RFCD codes). About 24% of respondents indicated that their major disciplinary area is medical and health sciences, and between 10%-12% indicated one or several of the following disciplinary areas: social science, humanities, arts; biological sciences; information computing and communication sciences; and education. All other disciplines (RFCD codes) were indicated less often.
4. Disciplinary clusters. Overall, about 24% of respondents indicated RFCD codes that belonged to two or more disciplinary clusters. Of the rest respondents about 22% of participants represented social, behavioural and economic sciences; 20% medicine, biology sciences and biotechnology; 11% engineering and environmental sciences; 7% physics, chemistry and geosciences; 7%- mathematic and information and communication sciences; 5% humanities and creative arts; and 3% other disciplines.
5. Primary role. Respondents quite evenly represented different levels of research experience: 23% were graduate students; 24% early career researchers; 21% middle career researchers and 22% senior long experience researchers. The remaining 8% of respondents were other university staff (mainly general staff, librarians, research assistants, IT and research support) and 2% of respondents held emeritus, honorary, visiting or adjunct appointments.
6. Awareness about eResearch. About 57% of respondents answered that they had not heard the term eResearch before the survey. This indicated that respondents have different levels of awareness about eResearch.

2. Research tools and methods

7. Software use for data handling and analysis. Only 13-17% indicated that they don't use spreadsheets or databases for data handling and analysis, while about 83-87% of respondents indicated that they use these two softwares often or occasionally. In contrast 78-84% of respondents indicated that they do not use the following four types of software: data mining; GIS or other special software; digital voice recognition and transcription; special visual or audio analysis software.
8. Digital research methods: needs for support. Asked to list digital research methods and tools they don't use, but would use if supported, respondents indicated a wide range of research tools. Frequently mentioned tools included software for statistical, mathematical and financial analysis (21%), modelling and simulation (21%), visualisation and visual data analysis (19%), digital voice recognition and transcription (18%), data mining (16%), and qualitative textual and linguistic data analysis (16%)⁵.

⁵ The percentage indicates the proportion of the respondents who wrote at least one answer to this question. This applies to all open ended questions: 8, 14, 15, 19, 23, 30, 32, 33.

9. High performance computers. About 13% of respondents indicated that they use HPC facilities for their research. APAC – was the most often mentioned HPC facility.

3. Research collaboration and dissemination

10. Extent of research collaboration. Less than 23% answered that nearly all their research is individual, while the remaining majority of respondents were involved in collaborative research. About 44% of respondents indicated that nearly all their research is collaborative and about 33% indicated that about half of their research is collaborative.

11. Loci of research collaboration. Most research collaboration occurs within research groups (70%) and with other universities outside Australia (52%) and within Australia (50%). The extent of collaboration outside the research group, but within the same faculty or other faculties within the same university is smaller, 35% and 31% respectively. Only about 30% of collaboration is beyond academia and about 22% with non-university research agencies.

12. Collaboration media. The most common media for research collaboration used by more than 91% of respondents were email, face-to-face meetings and telephone (Only 1-2% of respondents indicated that they don't use email and face-to-face meetings for collaboration). In contrast, about 94% of respondents indicated that they don't use virtual research environments and about 88% indicated that they don't use special project/task management tools. Less than 27% of respondents indicated that they occasionally use collaborative document writing and social networking tools.

13. Dissemination of research findings. Conventional publishing (e.g. journals, books) was the main form of research dissemination - used often or occasionally by more than 96% of respondents. About 67% of respondents indicated that they often or occasionally publish in online proceedings; 60% said that they publish in e-journals; and 46% publish in open access e-journals. 36%-45% of respondents indicated that they use institutional, project or personal websites, blogs or wikis or institutional repositories to disseminate research findings. Less than 20% of respondents indicated that they use digital disciplinary repositories or publish in open access, commercial or scientific societies' ebooks.

14. Dissemination and collaboration: needs for support. Asked to list main areas of research collaboration and dissemination that would most benefit from ICT support, respondents among many other areas often identified synchronous communication (e.g., videoconferencing), internet related tools and needs (e.g., websites), a range of needs related to data handling (e.g., research repositories), e-publishing and project management.

4. Research data

15. Nature of research data. Asked to describe the nature of digital and non-digital research data, among many other data types, respondents most often listed survey data (35%), interview transcripts and speech data (25%), laboratory measurements (24%) and a range of visual data.

16. Data collection methods. More than 87% of respondents often or occasionally collect/create data themselves or as a part of a team. More than 71% of respondents obtain data directly from other respondents (often or occasionally); about 57% of respondents obtain data from the third party organisations or data archives and repositories, and 24% from commercial online sources. Most respondents use data obtained from somebody else only occasionally, and most often collect or create data themselves.

17. Data digitisation. Respondents have a significant proportion of their research data and materials in digital format: more than 50% of respondents indicated that about 81-100% of their data and material is in digital format and further 23% of respondents indicated that 61%-80% of their data and material is in digital format. 9% of respondents indicated that 0-20% of their data is digital.

18. Volume of digital data. Most respondents store quite small amounts of digital data: about 47% of respondents indicated that they currently store less than 10GB and further 36% indicated that they store less than 1TB of digital research data. Nevertheless, some respondents indicated that they store quite large amounts of data: 12% between 1 and 10TB and 3% more than 10TB.

19. Data storage needs. About 33% of respondents indicated that they need more storage for research data. Asked to indicate how much storage they need in total, 21% of respondents indicated less than 1TB, 37% - between 1TB and 10TB and 29% - 10TB and more.

5. Data sharing practices

20. Access to data. About 48% of respondents do not allow researchers from outside their research project or team to access their research data, while 43% allow some access and only 9% allow access to all their research data.

21. Data sharing methods. Of those, who allow access to their data, 73% indicated that they typically provide access via privately negotiated access. About 37% of respondents also publish data online; 29% submit data when they publish in ejournals; 24% deposit to repositories and only 13% provide access via third party.

22. Reasons for restrictions. The most often indicated reasons for restrictions on accessing data were privacy and confidentiality issues (57%) and competitive research advantage (44%). About 18-24% of respondents indicated other reasons, such as: commercialisation potential, ethical issues, technical difficulties and lack of incentive; and about 15% of respondents indicated licensing issues and lack of usefulness of data for others. Only 7% of respondents indicated that there is no reason for such restrictions.

6. Data management and preservation

23. Data management and preservation issues. About 43% of respondents indicated that they face data management and preservation issues and further 20% were not sure. Among many other

issues, concerns associated with data storage, backup, archiving, preservation, security, confidentiality and other ethical issues were mentioned quite often.

24. Data management plan. About 54% of respondents indicated that they or their research team do not have an explicit data management plan and further 15% were not sure or didn't know.

25. Data storage during the project. More than 81% of respondents indicated that they store their data during their project on internal hard drives. More than half (53-55%) also indicated that they store their data on USB drives and external hard drives and about 40% indicated that they store data on a local area network and CDs/DVDs. Only 17-18% of respondents use departmental or central IT/university computing centres for data storage. Very few respondents (2-3%) indicated that they store their data during their projects in disciplinary, state/national data centres /repositories or commercial storage facilities.

26. Data storage after the project. About 79% of respondents indicated that they themselves typically store and care for their data after the end of a project. Less than 6% of respondents said that this is carried out by university or faculty services and less than 1% said that their data after the project is stored in national, state or disciplinary repositories. 4% indicated that nobody stores or take care after their data after the end of a project.

7. eResearch support, training, and services

27. ICT support for research. More than half (56%) of respondents indicated that ICT-related support for their research is typically provided by IT staff in their faculty or school and 40% indicated that this is done by central university IT. About 14-23% of respondents indicated that they typically receive IT support from colleagues, research personnel, IT staff in their research group, PhD or graduate students. About 11% respondents said that nobody provides them with such support.

28. Need for IT support. About 29% of respondents said that present ICT support minimally or not at all matches their research needs; 31% answered that it matches needs only moderately and 36% indicated that it matches need quite well or very well. About 5% of respondents indicated that they don't need IT support.

29. eResearch related bodies. Asked about their familiarity and use of services provided by five different Australian eResearch bodies, more than 93% indicated that they had never heard about Australian Access Federation (AAF). More than 81% also had never heard about Australian National Data Service (ANDS), Australian Research Collaboration Service (ARCS), and National Computational Service (NCI). 74-75% respondents had never heard about Intersect and National Collaborative infrastructure Strategy (NCRIS). Only 3-4% of respondents indicated that they use the services of Intersect, NCI and NCRIS.

30. Areas for ICT support. Asked to list specific areas of ICT technical or human support that would enhance their research capacities, the respondents indicated a broad range of services. Often mentioned areas were discipline-specific expert support (e.g, expertise in data analysis), general

expert support (e.g., IT personnel), data management and access to specific and general resources (e.g., research software).

8. Attitudes towards digital technologies in research

31. Importance of eResearch. More than 70% of the respondents indicated that eResearch is very important or important for the future progress of their research field with further 21% indicating that it is only moderately important. Only 7% of respondents said that eResearch is of little importance or unimportant and 2% that ICT is not relevant to their research. Asked to explain their answers, respondents often mentioned the importance of computational tools, digital data, access to ejournals and other online resources and collaboration for their research fields.

32. Future applications of ICT. Asked to describe the most important applications of ICT in their research fields, respondents often mentioned virtual research communication and collaboration, access to and collection of data (28%); various areas related to data storage and management, and various, often computation intensive, software applications.

33. Future challenges. Asked to describe the most important challenges for the application of ICT in their research field, the respondents often indicated a range of issues related to training and ICT expertise (26%), various data management (19%) and physical data storage (15%) related issues, adequate funding (16%), data analysis related issues (15%) and adequate ICT support (14%).

9. Further information and willingness to be involved in elicitation of eResearch requirements

34-35. Willingness to participate in eResearch focus groups and discussions. More than 30% expressed their willingness to participate in follow up discussions aimed at developing eResearch support and services in NSW.

36-38. Information about the survey and eResearch. About 59% of respondents indicated that they would like to receive a brief report of the survey, 52% answered that they would like to receive announcements about other eResearch activities from Intersect and 60% indicated that they would like to receive such announcements from their own university.

40. Final comments. About 10% of respondents wrote some final comments. Among positive and negative comments, a number of respondents expressed concerns about their low familiarity with eResearch terminology and ICT.

Statistical results are available in the Full Technical Report