Building an integrated research data management framework around Pure

UNIVERSITY OF CANBERRA

Research Intelligence

Pure

Your brilliance, organized for action

Make confident program decisions, streamline funding analysis and forge new partnerships with complete, correct, current and connected information around research activities.

ELSEVIER

Dr. Anesh Nair

Manager, Research Ethics, Integrity, and eResearch,
University of Canberra

"We wanted researchers to be able to manage project data through its lifecycle seamlessly. From planning and collection to storage and analysis, and publication to archival or disposal. This led to us embarking on an ambitious project to integrate three enterprise systems used at the University — Pure (Grants Management), ReDBox (research data management planning), and Infonetica (ethics). The end-to-end and streamlined research data management process enabled by the integration helps us adopt best practices for data management and creates a safe and secure data environment for all research stakeholders."

Executive Summary

The Australian Code for the Responsible Conduct of Research, which came into effect in 2019, lays down principles that underpin responsible research including how data is stored, accessed, and shared. But, more importantly, for the University of Canberra and other universities, the code places the primary onus of compliance on institutions and individual researchers. In response, the university undertook the research data management framework (RDMF) project in 2020.

The RDMF project envisages a seamless workflow throughout the research lifecycle, from grant award to research publication, by integrating the various technology platforms used at the university. In this case study, Dr. Anesh Nair, Manager, Research Ethics, Integrity, and eResearch at the University of Canberra, shares his vision of using technology to address the challenges of responsible research along with how they are using various Elsevier solutions and integration services to connect systems and platforms from multiple vendors to deliver a user-friendly experience for researchers and administrators.

University of Canberra: A young institution focused on research

The University of Canberra is a young and fast-emerging institution for academic and research excellence. Based in the heart of Australia's capital city, the university is home to nearly 10,000 students. They have taken rapid strides up the ranking charts in recent years and ranked 16th in the world and 2nd in Australia as per the 2021 edition of the Times Higher Education (THE) Young University Rankings.

Research at the university goes beyond academia to create solutions for real-world challenges through innovative research and applications. The university's efforts are focused on a number of high-impact areas that include climate change, life sciences, STEM, and governance. Another aspect of research at the University of Canberra is its commitment to an inclusive, collaborative, and global approach. The university regularly

engages with other academic institutions, government agencies, industry and businesses, and organizations from across the world.

Transitioning to the new era of responsible research

In recent years, we have witnessed an increased focus on responsible research at universities and educational institutions, particularly around issues of ethics, integrity, storage, and access to data. In line with this, Australia's National Health and Medical Research Council and the Australian Research Council developed the Australian Code for the Responsible Conduct of Research (2018).

The code, which came into force in 2019, promotes responsible and ethical research, lays down principles, recommends guidelines, and establishes a framework for conducting research. But more importantly, the code places the primary onus of adherence on institutions and individual researchers, marking a significant shift in how Australian universities and academic institutions manage their research and data.

Planning a research investigation is a complex process that includes multiple critical elements relating to data. This spans from collection methodology to validation techniques and secure storage to maintaining data for future reference, audits, and other uses.

Taking stock of research workflow, systems, and processes

In response, the University of Canberra conducted a need-gap analysis of tools and practices used by researchers. The study revealed the need for standardization and best practices for various phases, aspects and tasks of a research project. Some of the key findings include:

- **Lack of standardization:** Various departments and faculty members followed their own approach to research and data, which flowed down to their students and researchers.

Dr. Kyle Hemming

eResearch Analyst, Intersect

"In all my engagements with researchers and related academic staff during the rollout of this project, one thing that stood out was the relative lack of awareness about responsibilities and risks pertaining to research data. It can also be a laborious process, and for time-poor researchers, it is often seen as an additional administrative burden."

> "Having access to a structured, compliant, secure data framework offering several templated solutions defined by methodology can be very useful and reassuring for researchers. Moreover, storing the data with the university (as against individual researchers) ensures safety and compliance with regulatory norms. The new framework provides a helpful structure that was missing earlier and will go a long way in helping researchers and improving research integrity."

— Hugh Mason
PhD Student, Faculty of Business, Government & Law, University of Canberra

For instance, a particular department did well in capturing research output, and another department did not. And despite being part of the same institution, there was little sharing of best practices and knowledge.

- **Need to improve data storage:** The study found that research data storage practices varied making it difficult for the university to trace the provenance of data and track intellectual property (IP) created. Additionally, grant and funding providers have mandated structured and approved research data management plan (RDMP), so the issue of data storage needs to be addressed.

- **Paper-based processes were in use:** Many processes involved during the research lifecycle were paper-based, thus making the task laborious and ineffective. For example, researchers needed to fill out an MS Word document to submit their RDMP, which made the process cumbersome for researchers, faculty members, and administrators.

- **Multiple tools operating in silos:** The university used numerous digital tools and platforms across the research lifecycle. These are from different vendors, and despite being connected in the overall workflow, the platforms operate in silos making it complex and repetitive for users. From the university's perspective, the absence of a single source of truth prevented them from making the most of the research conducted across the institution.

Prof. Charles Lemckert

Honorary Appointment, Faculty of Arts & Design, University of Canberra

> "Enabling users to develop a data management plan at the initiation of any project ensures that the data is appropriately located and stored as per regulatory standards. The deployment of Pure and the integration with various related systems seem well suited for the purpose."

Designing an integrated research data management framework

An outcome of the research workflow analysis is the research data management framework that the University of Canberra commenced in 2020. The project, done in partnership with Integration Delivery Services team from Elsevier, envisages a seamless workflow for users and stakeholders involved with the integration between the major platforms they currently use:

- **Pure** is a world leader in Research Information Management Systems (Current Research Information Systems) that brings research information into a single, versatile and secure platform.
- **Data Monitor** helps universities and institutions track and analyze missing datasets stored across thousands of research data repositories.
- **Digital Commons Data** is a research data repository that helps researchers, administrators, and data curators to store, manage, curate, share and preserve data.
- **ReDBox** by Queensland Cyber Infrastructure Foundation (QCIF) helps universities collect information and metadata about research projects across the institution.
- **Infonetica** is software that automates and streamlines processes related to research ethics throughout the project lifecycle.

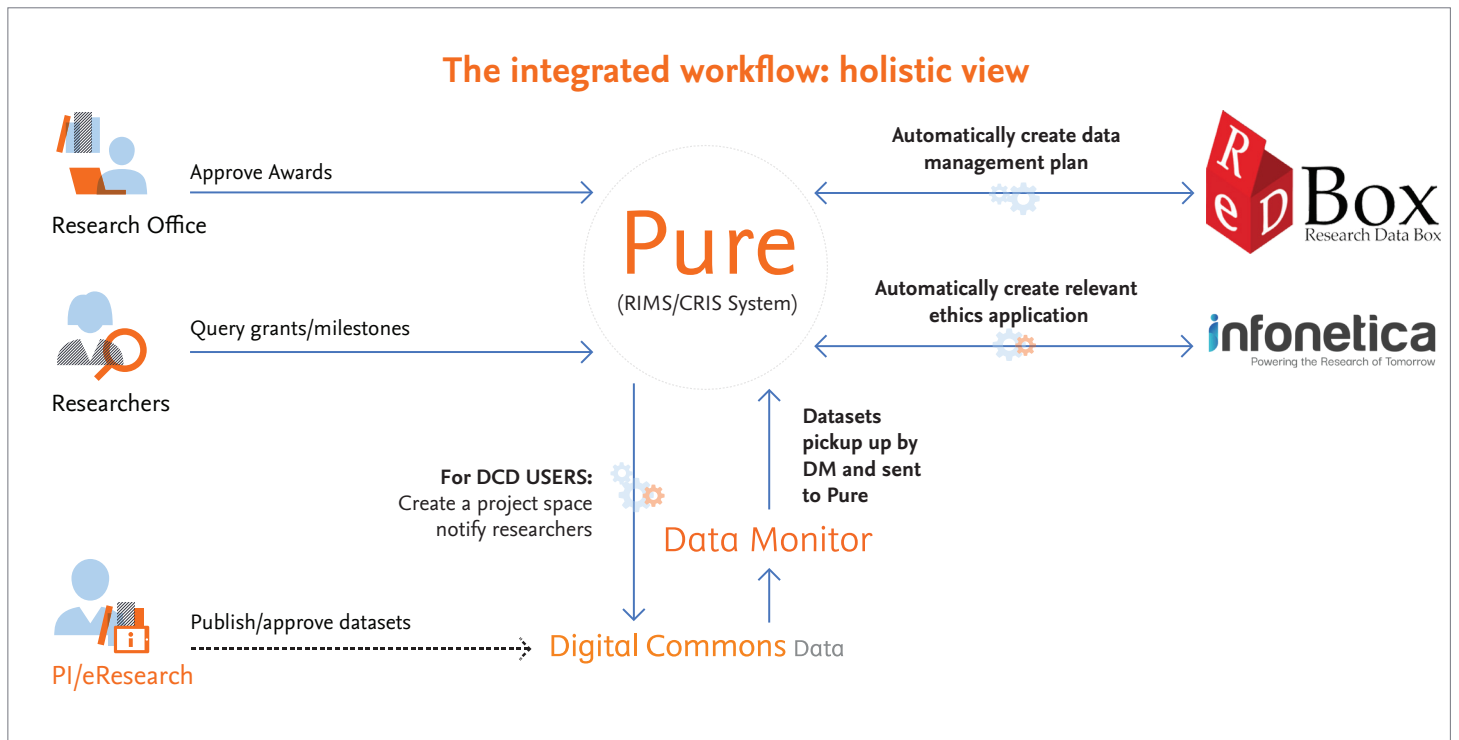## The integrated workflow architecture with Pure at the core

As part of the plan, the integration began with Pure, which the university was already using for grants management. And since

Pure offers deep integration capabilities, it could also serve as a core to connect other systems and platforms to automate various processes for researchers throughout the research project.

The following architectural view of the integrated workflow outlines specific points of connection between platforms mapped to critical events throughout a research project lifecycle.

- **Grant award to RDMP:** A research project commences when the Research Office awards a grant and the participants are notified in Pure. At the same time, an RDMP workflow is created in ReDBox with pre-filled metadata extracted from Pure, including project title, description, funding information, researcher and project participant details. This allows researchers to get on with the RDMP without spending time duplicating information that already exists in the system. Once the RDMP is approved, a project ID is created in Pure.
- **Grant award and ethics:** The awarding of a grant in Pure triggers the creation of a workflow in Infonetica if ethical

---

The integrated workflow: holistic view

Research Office — Approve Awards → **Pure** (RIMS/CRIS System)

Researchers — Query grants/milestones → **Pure**

Pure ↔ Automatically create data management plan ↔ **ReDBox** Research Data Box

Pure ↔ Automatically create relevant ethics application ↔ **Infonetica** Powering the Research of Tomorrow

For DCD USERS: Create a project space notify researchers

Datasets pickup up by DM and sent to Pure

Data Monitor

PI/eResearch — Publish/approve datasets ⇢ **Digital Commons** Data

to ensure compliance with data storage and related best practices.

- **Draft to dataset curation:** As project members finish their research and share a draft for publication, a curation workflow is initiated in Digital Commons Data. In this stage, moderators look at the data for adherence to FAIR Data Principles (findable, accessible, interoperable, and reusable) and other sensitivities involved in publishing the research. The platform also facilitates comments, revisions, and iterations until final approval is received.

- **Curation to publication:** The final approval for the dataset triggers the creation of a DOI (Digital Object Identifier) and publication in the institutional data repository according to the visibility criteria and restrictions. At the same time, the research gets added to the Data Monitor index, making it available to the research community around the world.

- **Publication to showcase and reporting:** The final step of the process takes place where it began, in Pure. Following the publication of a dataset, Pure imports it from Data Monitor and automatically links it to related profiles and publications. Additionally, with Pure as the central repository for all research and related information, universities and institutions can use it for mandatory reporting and to engage with stakeholders, funders, potential partners and prospective researchers to showcase the university's research prowess.

Impact of the project on the research function

- **Enabling researchers:** This project has resulted in the creation of an end-to-end and streamlined research data management process for all researchers at the University of Canberra. It has enabled the university to meet its requirements for providing users with the means to create high-quality data management plans for their research projects.

considerations are involved in the project. This includes the participation of humans or animals in research, working on genetics, DNA, or diseases and other fields that handle sensitive data and require regulatory oversight or restricted access. The project is greenlit in Pure once mandated authorities sign off on ethical approvals.

- **Grant award to project record creation:** Following RDMP and ethical approvals, a project record is created in Pure, and once again, all the metadata gets transcribed into the project record, which then becomes a single point of reference for all information and data related to the project.

- **Project record to workspace and active dataset:** The creation of a project record in Pure triggers the formation of a dedicated workspace in Digital Commons Data. At the same time, authorized participants are notified, invited to collaborate, and given access to data based on their roles. Participants can also refer to their approved RDMP within the same workspace

- **Enhancing data governance:** In addition to the functioning of the enterprise system, the pre-launch roadshow for faculty members enabled the university to communicate its expectations and the responsibilities of researchers. Relevant training modules and related pages on the new eResearch portal allowed the university to reiterate the messaging of the need for robust data governance.

- **Infusing effectiveness and efficiency in the process:** Integrating the systems to deliver a seamless experience has helped reduce administrative burden for researchers. It has helped make the research data management process more accessible and has whetted the appetite for similar initiatives in the future, including the integration of Pure and Infonetica.

Learnings from the project for the research function

Projects where disparate systems need to be integrated come with a fair share of challenges, ranging from engaging with users and internal stakeholders to design a suitable framework to dealing with vendors and solution providers to realize the vision into a platform that delivers value for researchers. Having led this multi-faceted integration at the University of Canberra, Dr. Anesh Nair shares learnings from the project:

- **Keep it simple and focus on the big picture:** When you are connecting multiple systems, complicating matters can make it seem like a daunting task. Keeping it simple and underlining the essential connections that need to be made across systems makes the project manageable. Once those connections are established, we can tweak and add features later. This also helps the project take off without weighing it down with minute details right at the start.

- **Working for the project, not the client:** We often come across vendors who are content to tick off the boxes and get done with a project. However, in a project with multiple moving parts, it is essential to have a partner and a collaborator willing to think out of the box to deliver a project as envisioned.

- **Working with multiple vendors has a learning curve:** This is perhaps the most underrated aspect of project management in today's globalized world and work environment. Companies from different parts of the world have different cultures and working styles, and differences in time zones and language also add to the complexity of a project. Thus, all involved parties must invest time and energy in understanding each other and establish a mutually understood working relationship right at the beginning.

- **Some shortfalls may be inevitable:** While planning an integration of this nature and scale, we always strive to perfect the system before going live. However, some shortfalls may be inevitable. For example, we realized during testing that all metadata did not flow through as expected. This was mainly due to gaps in data at the source, which was another university enterprise system. And the data shortcomings got replicated through Pure and the rest of the integrated systems. To address this, the team worked towards developing prompt workarounds and manual interventions, which ensured that the final integration outputs were as desired when the system went live.

- **Tailoring the message to maximize uptake:** After going live with the integration, it was important that we got the message out to researchers, as well. So, as we rolled out a roadshow, we discovered that it was crucial for us to tailor the messaging for various faculties and users. For instance, some preferred an all-inclusive talk about research data management followed by a demonstration on how to use ReDBox and information on the integration itself. Others preferred these sessions to be divided into bite-sized chunks. By working closely with each individual faculty, we were able to customize our outreach which resulted in a high turnout for our outreach sessions and a deeper engagement with users.

Pure

The world's leading research information management system that enables evidence-based decisions, simplifies research administration and optimizes impact, reporting, and compliance.

Data Monitor

Data Monitor puts institutions in the driving seat of their Research Data Management journey by providing visibility on their institution's entire research data output.

Digital Commons Data

Digital Commons Data is a secure and trusted generalist repository to publish datasets following the FAIR principles of findability, accessibility, interoperability and reusability.

For more information scan the QR code or visit:

elsevier.com/solutions/pure

Elsevier offices

ASIA AND AUSTRALIA
Tel: + 65 6349 0222

JAPAN
Tel: + 81 3 5561 5034

KOREA AND TAIWAN
Tel: +82 2 6714 3000

EUROPE, MIDDLE EAST AND AFRICA
Tel: +31 20 485 3767

NORTH AMERICA, CENTRAL AMERICA AND CANADA
Tel: +1 888 615 4500

SOUTH AMERICA
Tel: +55 21 3970 9300