CASE STUDY

A Digital Curation Centre Case Study
October 2013

Improving Research Visibility – Getting Data on the Institutional Repository RADAR

Angus Whyte, Digital Curation Centre

Section of the How to guide that this supports
Institutional data repositories

Introduction

This case study highlights collaboration between Arts and Humanities researchers and the RADAR institutional repository at Oxford Brookes University. The University's Sonic Art Research Unit (SARU) aims to make data a more visible research output, and improve the chances of it being accessible to researchers over the long-term through RADAR. Using the repository as a 'back end' for digital audio files, posters, photographs, podcasts, presentations and documentation allows them to be both centrally stored and catalogued, and locally presented on the SARU website. The study illustrates the role that digital data repositories can contribute to Research Excellence Framework assessment of impact and environment. So far the major challenges for the SARU-RADAR collaboration are:

- encouraging colleagues to change their practices
- planning what to archive and make citable,
- describing the selected material
- developing the Unit's strategy for dealing with its outputs.

Background

Oxford Brookes University is a relatively small and new (post-1992) university, with several hundred research active staff. The institution has an ambition to improve its research profile and infrastructure, and a strategy focusing on interdisciplinary research. It was one of the earliest UK institutions to develop a policy and roadmap for research data management (RDM) in response to funding body expectations. A steering group led by Professor Alistair Fitt, the Pro-Vice Chancellor for Research, was supported by DCC through its programme of institutional engagement.

In 2012 the group begin scoping research faculty requirements and raising awareness of funding body policy obligations. Actions were mainly undertaken through close cooperation between the Research Business Development Office (RBDO) and the Library. This led the Director of Research Support Sarah Taylor and the RADAR repository manager Rowena Rouse to lead an intensive series of Data Asset Framework interviews with research groups across the University's four Faculties.

RADAR (Research Archive and Digital Asset Repository) is a multi-purpose online resource bank for research outputs and online teaching materials. Set up in 2009 with support from the JISC Repositories and Preservation programme, RADAR was born from the realisation that an open access platform was needed to manage a wider range of outputs, from research publications, theses and dissertations through to teaching materials. The EQUELLA platform was seen as an appropriate solution for managing the varying metadata and access control requirements.

RADAR is now conceived in the RDM roadmap as a repository for 'final' outputs including research data collections. For repository manager Rowena Rouse "this is about using RADAR to create digitised special collections and making sure that items have rich metadata associated with them". Her involvement in DAF interviews on researchers' requirements identified opportunities that offer tangible benefits to individuals and groups, and one of these was the Sonic Art Research Unit (SARU).

In parallel with the institutional activity, in the last few years SARU researchers have been using more digital materials in their research and also disseminating these online. Research in the unit is practice based "we've got people building things, making instruments, writing scores, realising other

people's scores, making podcasts, making radio shows, doing installations. A very broad spectrum of activity taking place" says Dr Felicity Ford, an Early Career Research Fellow.

Successes on the road to long-term access

Identifying opportunities for practice-based research

SARU Research Lead Professor Paul Whitty and post-doc Felicity Ford have worked with RADAR manager Rowena Rouse to build a collection of digital material representing the Unit's research and practice, due for launch in 2014. The outputs to be made accessible include digital audio files, posters, photographs, podcasts, presentations and other documentation, for example musical scores. These are resources for dialogue between humanities research fields including Composition, Sound Art, Sound Studies and Field Recording. This has led to interdisciplinary collaborations, and a focus on site-specific practices, field recordings, and soundscape studies.

For lead researcher Professor Paul Whitty the RADAR repository offers opportunities; firstly to consolidate high quality masters of the research unit's digital assets, and secondly to ensure these are contributing to the Units research impact by being highly accessible through multiple online channels;

> "…it's very easy to upload an MP3 to an Internet sound mapping site or audioboo or SoundCloud or whatever, and relatively inexpensive, but what you really want is the high quality audio file stored somewhere……so you know its not only in one location on a hard drive."

Building the collection

Towards the related aims of a more persistent record of research outputs, and enhanced visibility of these research outputs, RADAR will provide direct access to them in the same manner as it currently does for other materials, organising these as collections of related objects. EQUELLA's capabilities to manage metadata and associated workflows for editing and defining access levels are attractive to RADAR manager Rowena Rouse:

> "…we can customise, I can create collections to deal with different types of subject areas and things so its very flexible. If you have a schema you can apply that to a collection, you can create workflows, and then decide on what you want to go public"

The collection also provides a back-end role for SARU online resources, enabling links to other online sites and tools that sound researchers and their audiences may use to access and render the content.

Research outputs to be stored on RADAR span the variety of normally ephemeral artefacts from performance-based research mentioned above. The front-end includes SARU's website (www.sonicartresearch.co.uk) and project-related off-shoots, hosted on third-party social media sites such as blogging platform Wordpress and audio sharing site Soundcloud. This should provide multiple routes to discovery, as illustrated below, with RADAR providing the persistent archive for objects that are linked to from the SARU website and social media, as well as facilitating discovery by others unconnected to the disciplines involved.

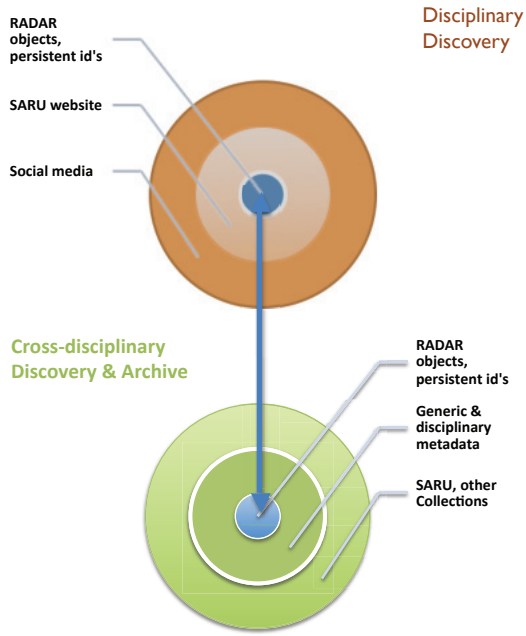

Figure1. Routes to Discovery and Long-term Access

Post-doc practitioner Felicity Ford has provided much of the impetus and effort towards establishing SARU's archival collection on RADAR. Ford was partly inspired to explore the opportunities to make this openly available by the example of UbuWeb, a US resource hosting scores, texts, video and sound clips from of avant-garde artists; "…what I think is amazing is being able to see properly organised assets that are 40 years' old". Ford, whose research and practice explores the domestic soundscape and field recordings, has also been motivated by frustration in trying to cite material only to find the links to it have broken: "…reading journal articles about projects in my area, when I go to look at the website that they're citing from 2006, the website has completely disappeared. And this is a huge problem. So I'm really keen to preserve some of the long term projects."

[1]Sonic Art Research Unity home page. Available at: http://www.sonicartresearch.co.uk/home/

Making research more visible

Alongside the goal of better long-term accessibility, the researchers are ever watchful of the need for short-term promotion. Identifying their projects, podcasts, radio shows, academic presentations, PowerPoint presentations etc on RADAR helps promote the visibility of the department.

Symposia and festivals provide a similar focal point for sharing as conferences and workshop series in other disciplines. They include the *Sound Diaries* project with its annual Symposium, and *audiograft* an annual festival of Contemporary Experimental Music and Sound Art curated by SARU. Digital resources from these events are accessible from the SARU and audiograft.com sites, with embedded links to files published elsewhere, for example in Consumer Waste (consumerwaste.org.uk) – a site described as "a low-impact imprint for the publication of contemporary experimental music", and partly run by fellow SARU practitioner Stephen Cornford.

Managing copyright and other intellectual property issues needs dealt with on a case-by-case basis. Restrictions on reuse are not felt to be particularly contentious; open access is both accepted and necessary for promotion, but reuse less so. As practising artists the researchers typically do not want to waive all rights by default so while Creative Commons licensing is used it is frequently with 'no derivatives' terms; "…it's trying to strike a balance between making the resources available, but not making them completely plagiarisable" according to Ford.

Challenges

Using the collection to demonstrate impact

Practice-based researchers can face data management challenges in demonstrating the impact of their work, for example through the Research Excellence Framework (REF). Practice in digital arts and humanities inevitably extends beyond the conventional types of scholarly output, and impact case studies may require these outputs to be traced over 10 years or more. This is a further incentive to organise and preserve evidence as Paul Whitty relates: -

> "It's important for us in terms of the REF, there's an impact case study we're writing around some of the Sound Diaries projects, and so this is going to be really useful for us to know where everything is, really, I mean, just as simple as that. And also I think in terms of writing the REF environment statement and talking about how we care for our data… it's becoming more and more important amongst schools of arts as they realise that what they're doing in their processes is actually collecting data, because I think a few years ago we wouldn't have thought of it in those terms, and of course everyone had a hard drive full of research images and drawings and sounds and we wouldn't necessarily have thought that it was something that should be preserved. But I think now, more and more, we're realising that because… research organisations like the AHRC are so interested in funding innovations in methodology – I think it's really important that we preserve the data relating to those methodologies.

The need to plan for impact begins with the earliest stages of file organisation according to Felicity Ford:

> "Working on the RADAR archive has changed my strategy for how I manage my own files and how I think about …when I create a resource, where is this going to go? And, how am I going to give it the maximum life that it can have? If you can't demonstrate your impact, then you lose in REF as far as I understand…and it's also [that] time is finite, so how much time have I got to spend putting stuff on Vimeo or putting stuff on SoundCloud when I don't actually know how useful that's going to be to me long term as a researcher… whereas if I put it on RADAR then it is useful to me as a researcher, so, yeah, it's changed those things".

Lessons learned and long-term ambitions

Building support and developing strategy

A key point for all involved has been the need to build support around the link between managing data assets in the short term and their longer-term visibility. According to Ford "trying to get the whole department engaged is not the easiest thing". One barrier she sees is that arts practitioner-researchers do not typically deal with their source materials as highly structured data. Another is the focus on short-term promotion: "everyone in our department suffers from the same thing, which is that we're all…public working artists trying to get our work out there, so the issue of publicising what we do is constantly an issue for funding and for getting people to come to our events".

The RADAR-SARU strategy is to pursue the link between long-term visibility and short-term savings on time managing resources. As Ford puts it; "…once it becomes clear that you only have to upload things once and then you can always link to it and it will always link to all your other stuff, then … That's the gold standard that we're aiming for with it, with the SARU archive."

The Audiograft festival will be used to publicise the archive. In its pre-launch phase the priorities are to make data management across RADAR and the third-party platforms involved (WordPress, SoundCloud etc) less labour intensive. Currently this work is handled by Felicity Ford and Rowena Rouse and "…it's not sort of the thing that you can just sit down and do half an hour on…you have to sit down and spend a day on it and have no interruptions."

Paul Whitty sees the key point as "we also need to empower other members of SARU, and say to them, you can do this yourself". To address this, the group are taking steps to:

- Train other SARU practitioners in the archival processes, and raise awareness of the need to select what to keep in RADAR.
- Reduce effort by identifying workflows that use spreadsheets to manage data transfer between RADAR/EQUELLA and the WordPress platform.

> "I think when you link storage and dissemination, that'll be the moment people will suddenly realise that this is a really awesome way of working"
>
> - Paul Whitty

Selecting resources and structuring information

SARU's data is of highly varied types. Defining metadata and populating the repository with this at a suitable level remains a challenge that has required restructuring the collection on several occasions. The main challenge has been the diversity of items to be described, and the consequent need to structure, catalogue and link the resources to make them usable in future.

Selecting data resources for archiving should lead to them being citable by other researchers and practitioners. In the short term the main issue around citation is identifying which outputs should be openly accessible and citable as a record of research and practice, and which preserved as outputs that practitioners see as valuable but do not want distributed in easily copyable formats.

Selection is about prioritising what to archive, for example through an annual review of data assets that should have been archived in the previous year and plan for the year ahead. Considerations here include the ephemeral nature of the social media used, and determining which aspects of their context will significantly affect how future users will interpret the archived resource.

Archiving audio and visual material can involve extra efforts to maintain digital masters as well as copies for lower bandwidth dissemination. These copies often need to be in different file formats, with long-term preservation implications. The Internet Archive has proved a useful tool for creating copies of files in a variety of dissemination formats.

The group have yet to encounter practical issues about recommending how their resources should be cited e.g. when multiple versions of a resource exist, or there is a need to identify a specific item in the resource, such as a short sequence in a long recording. Tracking how the RADAR collection is used is very important to SARU, and, as well as citations, the extent to which items are accessed on social media will be keenly monitored on sites like ImpactStory.

Further information

SARU (Sonic Arts Research Unit) Available at: http://arts.brookes.ac.uk/research/arp/saru.html

RADAR (Research Archive and Digital Asset Repository) available at: http://www.brookes.ac.uk/go/radar

Preserving Moving Pictures and Sound. Richard Wright, DPC Technology Watch Report 12-01. March 2012

Centre for Digital Music (C4DM) Sustainable Software for Audio and Music Research. Available at: http://soundsoftware.ac.uk

ImpactStory. Available at: ImpactStory.org

EQUELLA further information at http://www.equella.com/

A pilot report on deploying EQUELLA as a data repository was carried out by the Jisc ADMIRe project at University of Nottingham see Berry, M and Parsons, T. EQUELLA data repository pilot. Available at: http://admire.jiscinvolve.org/wp/files/2013/05/ADMIRe-EQUELLA-Research-Data-Repository-Pilot.pdf

EQUELLA is a registered trademark of Pearson
SoundCloud is a registered trademark of SoundCloud Ltd
WordPress is a registered trademark of the WordPress Foundation. Vimeo is a registered trademark of Vimeo LLC

Follow DCC on Twitter: @digitalcuration and #ukdcc