Death of a repository

UKSG eNews 414

02 Feb 2018

Andrew Gray, Goldsmiths, University of London

Goldsmiths, University of London has been involved in the long (endless?) march to 'open' since 2007 when we established our institutional repository Goldsmiths Research Online (GRO) and not long after, our nascent data repository, Goldsmiths Data Online (GDO). We were key partners in the JISC projects Kultur, which aimed to establish a model institutional repository for visual and creative arts, and Kaptur which was developed to help define the nature of research data within the visual and creative arts. In addition the Defiant Objects project looked at how non-traditional outputs could be described and accommodated within repositories with attendant metadata. Goldsmiths has continually investigated and been committed to understanding the needs of our researchers and to how research outputs in arts, humanities and social sciences need to be described, captured, preserved and shared in an online research environment. As of October 2017 we decided to merge our data repository with our research repository and thus close Goldsmiths Data Online.

Increasing costs and decreasing budgets prompted us to look closely at how GDO was being used, whether it represented value for money and whether it achieved its purpose to capture and share research data. Goldsmiths is an arts, humanities and social sciences institution but also has successful disciplines outside this portfolio including psychology and computing. A key area of the institution's research is 'practice as research' or just 'practice research' (the term we use at Goldsmiths) where often creative practice/activity forms an integral part of the methodology of the research. This strand of research, one that we have investigated in all our collaborative JISC-funded projects, is an area of continued debate, but also the one area where we were successful in obtaining research data for our data repository.

On the date that we transferred our research data to GRO all five records of research data were practice research, four from our Art Department (all from one AHRC-funded project) and one from Computing. Each of the records of datasets we held within our data repository comprised of a range of formats and file types including audio, video, images (born digital and digitised) and text. Reviewing these datasets on GRO we can see several issues that highlight the complexity of translating practice into expected standards of data. There is minimal contextual information, many files are in non-archival formats and the understanding of copyright is shaky.

The small number of datasets is mirrored by our partner institution from the Kaptur project, University of the Arts, London (the only other partner institution to establish a separate data repository) having six records, and Glasgow School of Art having two records. A look at the dates on which the records were created in all repositories shows that they are nearly all in the years 2014-2015, two years after the end of the Kaptur project. None of this reflects the amount of work that has gone into understanding research data in an arts context or that is continuing, as evidenced by the recent 'Does my data look good in this?' workshop in 2017. What it does reflect is that

research data for the arts still has a long way to go before it reaches a position where it is shared as part of a researcher's practice. It also reflects the difficult relationship that the arts have with research and the questions "When does art become research?" and "How does a research institution value art when it's not research?", which are questions that also affect and reflect the outputs in our research repositories and ones that I'm unable to answer here.

While it might be unfair, it is useful to look at how our research repository GRO fares. As of October 2017 we had 15,725 items and 1,971,752 downloads. For the five items in our data repository we had 237 downloads. The difference is stark but illustrates how research repositories in comparison to data repositories have been supported over the last few years. A combination of JISC support, research funder policy and government policy (HEFCE/Finch report) helped establish institutional repositories and an open access landscape. A similar landscape of funding and support for research data has not been matched, although within STEM disciplines the uptake and need for research data management and repositories has been more clearly articulated and established by bigger institutions and research funders.

Support and understanding of practice research/arts research in an online environment with an open access ethos requires continual discussion and attention. Issues and concerns identified in the Kultur and Kaptur projects over how the work is represented and described, and about ownership, copyright and funding need to be addressed with each new researcher. Although REF and open access does not intend to marginalise such research, our limited resources mean only so much of our attention can be devoted to having the conversations that ensure this research is properly represented and accessible. This becomes more evident with research data, and while 'Initial decisions on the Research Excellence Framework 2021' (REF 2021) encourage the effective sharing and management of research data that goes "above and beyond the REF open access policy requirements", we can see from our own experience that without resource and/or 'requirement' this does not happen.

Much of the research produced within arts is not funded and where it is funded there is often no recognition by the funder of the existence of research data or the value of it being shared. This is still an area of confusion to art practitioners. Our case study/contribution to the UCL LEARN project describes how practice research data can be managed productively but also highlights that "there is insufficient funding and understanding on the part of the artists and institutions involved as to how or even why it is worth making this data available" ('RDM in the Performing Arts'). Such research often floats under the radar of our/any systems to manage research (research office/institutional repositories) and escapes the structures and requirements of an institution and therefore operates on its own desires and needs.

If one of the purposes of sharing research data is for it to be used, you would want it in an area that is used. Moving our research data to GRO we believe will be of benefit in terms of engagement with the data since it is already a recognised site for research with a healthy monthly engagement. We are currently in the process of implementing the mining of doi's for datasets (alongside theses and exhibitions), will be populating Primo Central and are already aggregated by Core.

If research data within arts and social science institutions and those with a lean towards arts and practice research in general is to become visible, useful and valued then we need to address the issue of resource. At present current research imperatives do not allow the development of a data management service that responds to diverse research practices. Until this is done then we will continue to underestimate and overcomplicate research data within the arts.