|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Meeting title:** | | DataPool Steering Group Meeting | | |
| **Date:** | | 12th November 2012 | **Time:** | 12:30-15:30 |
| **Location:** | | John Roberts Room, Staff Social Centre (B38), Highfield Campus, University of Southampton | | |
| **Present:** | | Dr M. Brown; Dr L Carr; Prof S Cox; Dr G. Earl; Prof J Frey; Prof P. Nelson (Pro VC Research); M. Ployaert; W. White (Chair); .Dr O Parchment (on behalf Dr P. Hancock)  L. Corti (Associate Director, UK Data Archive, University of Essex); Dr H. Snaith; NERC  Project Managers: D. Byatt; Dr S. Hitchcock | | |
| **Apologies:** | | Prof A. Wheeler (Provost and DVC); G. Pryor (Associate Director, Digital Curation Centre, Edinburgh); S. Rumsey (Digital Collections Development Manager, Bodleian Libraries, University of Oxford); Dr P. Hancock; | | |
|  | |  | | |
|  | **Agenda item** | | | |
|  | **Minutes**  The previous minutes were accepted | | | |
|  | **Matters Arising and Update Report**   1. **Criteria to identify significant data**   This is a work in progress and is being informed by some of the data interviews that have been taking place. J Frey directed the groups attention to the concept of “intelligent access” in the context of the recent Royal Society report *“Science as an open enterprise*” (<http://royalsociety.org/policy/projects/science-public-enterprise/report/>). This suggests it is a case of “not all, not immediate” with the key question being “Does the data support the endeavour?”   1. **Additional Guidance topics**   The existing topics already covered were listed and a request for advice on further topics was made. There was a comment on the heading “Restricting Access to Research Data” as this was thought to be quite negative. This was accepted by the Group, but the section had been written to address specific concerns identified within the University during the Policy development process. An earlier section did address the issues around sharing  **Action:** Project to consider alternative heading | | | |
|  | **Policy engagement and integrated service support**  There was a useful discussion about the balance of individual expertise/networks and service structure in order to ensure a robust service.  NERC have moved to a more formal requirement for a plan, albeit a short one page document, with the aim of identifying cost.  Further development should involve contracts staff especially as they may have very useful contacts within the organisation. | | | |
|  | **Training model**  Increasingly there is a more co-ordinated approach to training and support for post-doctoral researchers in the institution. Although some ‘special’ training takes place there is a move to harness the knowledge and expertise in providing such support. Research data management training will make use of this.  The Steering Group emphasised the need to raise awareness of the role of good data management from the start. Post graduate training will be co-delivered with a PhD and Library or other appropriate staff.  We will need to make links to draw in and address the needs of the early career researcher especially with grant writing events.  The opportunities for making use of data during the life of the research were discussed in the context of training. One possibility was the use of data in the upgrade process and this had been considered. Issues related to this include privacy, how the data might be presented, how it would be used. If this were to develop it would impact on the training and emphasise the need to be thinking about the presentation of the data earlier. Some felt that arguments should stand in their own right with the data following later. This might include the publishing of the data alongside the thesis and there were other issues associated with this – should it be separate records/different embargoes.  Progress was still on-going on the development of training for support staff. | | | |
|  | **Multidisciplinary engagement**  The paper on the Imaging case studies was discussed. This would look at the data type being used, who was using it, what was best practice. Work had already resulted in the availability of some equipment being made known.  The ethics process involved had helped enhance the processes and approach and had demonstrated a need for a stronger link between data management and ethics processes.  There was a discussion on re-use of data that had been cleared for one purpose being requested for another. Some people are not really thinking about ethics approval in light of potential re-use. The importance of thinking about ethics approval in light of potential re-use of data and data sharing was noted.  Related to this was increasing awareness about other groups who might be interested or able to use the data created. This meant that recording the assumptions that were made in the creation of the data was important. Lessons might be learnt from some of the work being carried out by the National Centre for Research Methods and it would be worth looking into this.  There was also work on the visualisation of social science data. | | | |
|  | **Data Catalogue and Storage**  The development in SharePoint and ePrints were outlined. ePrints Soton would become the data catalogue for the University.  There were two very different examples of storage offers – Edinburgh and Bristol. There seems to be no coherent view in HE on the best way forward. We are currently building a case for service development and investment priorities. Work is under way with Associate Deans Research to obtain robust requirements information  Who does what is funder specific with some councils offering full or partial storage for data. It was not clear that NERC stored all data at the moment and that was an issue.  The issue of significance was discussed. Some data is almost always unique e.g. environmental data, some data might take 20+ years to become relevant, e.g. longitudinal studies data.  Need to consider the curation required and if there are different levels. Some data just needs to be catalogued; some needs curation and additional metadata. All key to the costs associated with data management.  Storage needs to be able to link data and code and syntax to make the data re-usable. | | | |
|  | **Linked initiatives for the next phase**  There was a review of the Web Observatory, a data preservation and archive facility.  Another linked development was LabTrove, an open software electronic laboratory notebook. This was a highly collaborative piece of software and researcher centric. It would provide the option of a URI in the lab book and dataset. It was hoped that it would be available as an open source and as a commercial product.  DOI minting is being investigated using DataCite, initially on eCrystals but would be rolled out for use on ePrints. Work would be required on what should or should not be allocated a DOI. | | | |
|  | **AOB** | | | |
|  | Date of Next Meeting 12th March 2013 | | | |