I've collected my data, so what do I do with it now?

Research data management

Overview of the Training Programme

DATUM for Health

www.northumbria.ac.uk/datum

DATUM for Health – Healthy research needs healthy data

Overview of the Training Programme
Aims and Scope; Design; Outline Content & Materials; Recommendations

Aims and Scope

The *DATUM for Health* training programme is aimed at postgraduate research (i.e. doctoral) students (PGR) in health studies, including those whose PhD has a health focus but who are not necessarily registered in a school/faculty of health/medicine (e.g. in psychology, social sciences). The programme covers both generic and discipline-specific issues, focussing on the management of qualitative, unstructured data, and is suitable for students at any stage of their PhD.

It aims to provide PGR students with the knowledge to manage their research data at *every stage in the data lifecycle*, from its creation to its final storage or destruction. Students learn how to use their data more effectively and efficiently, how to store and destroy it securely, and how to make it available to a wider audience to increase its use, value and impact.

Programme Design – Model, Rationale and Format

The programme was designed to dovetail with existing research skills / research management provision in the university, in particular relating to PGR milestone preparation (project approval and mid-point progression), records management, DPA / FoI / copyright, ethics and governance. It was also informed by input from PGR students (via a questionnaire survey) and other research stakeholders (via a focus group) in health, particularly in terms of content and format enabling the preferences and already known needs of the target audience to be identified as well as the potential unknown needs (as expressed by other experienced stakeholders). A face-to-face format was developed comprising a mix of presentations, real-life examples, worked exercises, group work, class discussion and directed learning[1]. Management of qualitative research data in the health field raises particular issues of research methodology, data analysis and ethics. It was therefore important to include time to debate these issues and enable participants to learn from each others' experiences.

The programme design drew upon the pedagogic, research data / information management, digital curation and preservation knowledge and expertise of the collaborative project team. The DCC's digital curation session was also used to 'train the trainers' (i.e. build expert capacity, and will be delivered in-house in future.

The andragogic model of learning and teaching underpins the programme's design[1]. This model's philosophy places the learner (in this case the researcher) at the centre rather than the tutor. Its principles are:

- ***The need to know*** — *adult learners need to know why they need to learn something before undertaking to learn it.*
- ***Learner self-concept*** — *adults need to be responsible for their own decisions and to be treated as capable of self-direction*
- ***Role of learners' experience*** — *adults have a variety of experiences of life which represent the richest resource for learning. These experiences are however imbued with bias and presupposition.*
- ***Readiness to learn*** — *adults are ready to learn those things they need to know in order to cope effectively with life situations.*

- ***Orientation to learning*** — *adults are motivated to learn to the extent that they perceive that it will help them perform tasks they confront in their life situations.*[i]

In practice this translated into an approach that:

- was based on real life needs, sharing (students' / tutors') knowledge and experiences
- incorporated a lot of discussion (i.e. not solely presentation, information transfer)
- included practical activities and directed learning (to support learning by doing, reflection and practical application)
- was developmental, progressive (i.e. sessions built on each other, the design was holistic).

The rationale for adopting the andragogic model was that it helps researchers to execute their responsibilities effectively and be equipped to continue to acquire more in-depth skills as appropriate. It was successful and can be used by other HEIs for research data management skills training programmes.

To make the programme attractive to those who may not recognise the importance of RDM, it was given an interesting title *'I've collected my data, so what do I do with it now? Research data management programme'.'* It covered all of the stages of the data lifecycle and comprised four sessions, building upon each other:

Session 1: Introduction to Research Data Management
Session 2: Digital Curation 101 Lite
Session 3: Problems and Practical Strategies and Solutions
Session 4: Digital Preservation Coalition Roadshow – *Data for Life: Digital Preservation for Health Sciences*

Students were encouraged to sign up for all four sessions at the beginning. The aim was to hold the sessions at 2-weekly intervals but this was not possible. Sessions 2-4 were held at weekly intervals and four weeks after Session 1. The participants did not find this a problem.

The interactive format and length of sessions (2.5 hours in accordance with other PGR training sessions with the exception of the final full day session) meant that the ideal group size was 20-25 people.

Outline Content and Materials

Session 1: Introduction to Research Data Management (2.5 hrs)
Learning outcomes:
- an understanding of research data management needs and issues
- the ability to start planning for managing own research data using a DMP

Directed learning tasks:
- start developing a personal research data management plan using template, in preparation for Session 2
- watch the 3-minute video of Louise Corti, UKDA, Essex University on 'How can researchers ensure that they'll be able to share, archive or re-use sensitive data?'

Session 2: Digital Curation 101 Lite (2.5 hrs)
Learning outcomes:
- appreciation of how data curation can support and safeguard research
- understanding of the data curation lifecycle
- identification of the processes and activities involved in good practice for research data management
- awareness of the free services and tools available to support data curation

Directed learning tasks
- think of the problems you've experienced with managing your research data, to share in Session 3
- think of any good tips or systems you've used for managing your research data, to share in Session 3

Session 3: Problems and Practical Strategies and Solutions (2.5 hrs)
Learning outcomes:
- awareness of research data management problems
- knowledge of practical solutions to research data management problems
- ability to develop a research data management strategy for the remainder of their PhD and in their future research career

Directed learning tasks:
- Reflect on your strategies and procedures for research data management, e.g.
  - Do you need to take any further actions to ensure your research data is appropriately anonymised and capable of being shared in the future, if appropriate?
  - Do you need to revise your file / document naming system, and / or the file structure for storing your research on your university drive?
  - Are the ways in which you physically store your data sufficiently safe and secure?

Session 4: Digital Preservation Coalition Roadshow – Data for Life: Digital Preservation for Health Sciences (all day)
Learning outcomes:
- awareness of digital preservation issues
- knowledge about digital preservation approaches
- awareness of the wider research data management agenda beyond PhD study and of initiatives in other organisations
- networking

Directed learning tasks:
- Make an entry in research diary / portfolio about research data management strategy and actions going forward
- Make an entry in training needs analysis document about research data management
- Meet with supervisory team to report on research data management learning and plans going forward

Note: this session was a public event which attracted ~50 participants

The programme drew upon relevant existing materials, tailored them appropriately and developed additional discipline-specific materials where needed. A house style (brand) was developed for the materials with the mantra ''*healthy research means healthy data'* and used to create coherence. Materials were in different formats to support different cognitive styles (e.g. PowerPoint presentations, notes of the discussion, templates, a customised search engine, videos from the Incremental project, DCC / DPC / UKDA leaflets / documents). They (or links to them) were stored on the virtual learning environment (VLE) for future access. Due to the nature and constraints of their research PGR students cannot always attend training sessions therefore sessions were audio recorded for potential use at different times and in different places.

Recommendations

Based on the experience of delivering the programme and its evaluation, the materials have been refined. Session 4 was not planned as a sustainable, repeatable session. Participants felt three sessions were sufficient and therefore the introduction to digital preservation (presented by Dr William Kilbride) has been incorporated into Session 2 on digital curation, and the directed learning tasks after session 4 have been incorporated into the tasks after Session 3. Session 2 has been rename to - 'Data Curation Lifecycle'.

The programme therefore comprises three sessions, and the training materials comprise:

- this overview of the training programme with model and rationale for its design
- PowerPoint slides for each session with separate accompanying notes for tutors to run the session listing all handouts/support materials used
- handouts / sample answers for learning activities that were produced by the DATUM project team. Other handouts (freely available from other organisations on the Web) are accessible through the URLs given.
- resources list

It is also recommended that Session 1 (or equivalent) be compulsory training for all new PGR students.

---

[i] Quoted from: http://www.learningandteaching.info/learning/knowlesa.htm referring to Malcolm Knowles. Knowles MS (1990) *The adult learner: a neglected species.* 4th ed. Houston: Gulf Publishing.