

Methodology notes for Institutional Data Management Blueprint (IDMB) Questionnaire Dataset

Data Collection:

First Iteration – 8th-31st of March

- Research Staff and Students in Archaeology, Electronics and Computer Science, Engineering Sciences
- Data is collected through an iSurvey (in-house system) online survey.

Second Iteration – May 2011

- Research Staff and Students across a wider range of faculties
- Changes to Q1.2 (School) and Q1.4 (Area of Research) to allow the survey to be used with all faculties and research areas across the university. To avoid an unmanageable list of research areas respondents see a list of areas for their faculty and an 'Other – please specify' option.
 - Within iSurvey School & Research Area were 1.2-1.2i, Research Role 1.3, Do you currently hold/store any research data 1.4
 - For the purposes of SPSS variable names the same numbering as the original survey was used for both iterations.
- 'Sharing not required' was added as an option to Q2.18b (Concerns about allowing others to access your data once the project is finished).

Initial processing of data

- Data is downloaded from iSurvey. Where respondents had not finished the survey their data was deleted if no questions in Sections 2 & 3 (the bulk of the survey) had been answered.
- For the second iteration all results from the 'Area of Research' questions are combined into a single column.
- The column headings (question text) were replaced with the variable name (Q followed by the question number, where there are multiple columns for a single question (multiple response) additional text was added).
- Imported into SPSS, data labels and multiple response sets added by SPSS syntax files.

Combining data from both iterations

- Data was combined in excel – where an individual had responded to both, the results from the second iteration was kept.
- A column was added indicating which iteration each response was from.
- Data in the 'Area of Research' question in the first iteration was improved for ECS and Engineering Sciences (checking their research group where needed).

Data analysis

- Carried out in SPSS, using syntax files to allow the same analysis to be repeated easily.
- Results exported to excel, and percentages recalculated where needed (to exclude the proportion who didn't answer the question).

Dataset modification for sharing

- The following fields were deleted (including all qualitative data to ensure anonymity) – 34 out of 134 variables.
 - Name (Q1.1)
 - IP Address
 - Other details - School (Q1.2A)
 - Other details - Research Role (Q1.3A)
 - Other - Area of Research (Q1.4A)
 - Funding body details (Q2.1A)
 - Other - ownership of research data (Q2.1B)
 - Other details - data characteristics (Q2.2A)
 - Other details - data types (Q2.3A)
 - Compatibility issues (Q2.4A)
 - Experiences of reusing data from previous projects (2.5A)
 - Experiences of using external data (2.6A)
 - National data centre or data archive details (Q2.7A)
 - Other Details - person responsible of managing data (Q2.7B)
 - External commercial web data storage facility details (Q2.8A)
 - Other provided file server e.g. by School unit details (Q2.8B)
 - Other details - location to store research data (Q2.8C)
 - Length of retention for compliance reasons (Q2.11A)
 - Other details (method to overcome storage issues) (Q2.12B)
 - External commercial web data storage facility details (Q2.14A)
 - Other details - location to backup research data (Q2.14B)
 - Details of data service - UKDA etc. (Q2.15A)
 - Other details - keeping track of where data is stored (Q2.16A)
 - Details of access to data given to others during lifetime of a project (Q2.17A)
 - Other details (issues in sharing data during lifetime of a project (Q2.17C)
 - Details of access to data given to others after a project has ended (Q2.18A)
 - Other details - issues in sharing data after a project has ended (Q2.18C)
 - Detail of data management plans requested by funder(s) (Q2.19A)
 - Details of School policies (Q2.20A)
 - What is the biggest problem for you with regard to managing and storing your data (Q2.22)

- How can the University including your School iSolutions and the Library make data management and storage easier for you (Q2.23)
 - Whether they were willing to be interviewed (Q2.24)
 - Email Address (Q2.24A)
 - Do you have any further comments you would like to make (Q3.1)
- Dataset converted to (SPSS) Portable format – this truncated many of the variable labels (see below) and removed the multiple response sets.

Abbreviated Name	Extended Name
DateFini	DateFinished
DateStar	DateStarted
Iteratio	Iteration
Timetake	Timetaken
Q2.12Aex	Q2.12Aexthd
Q2.12Aot	Q2.12Aoth
Q2.12Aus	Q2.12Ausb
Q2.14com	Q2.14comp
Q2.14e_1	Q2.14extstor
Q2.14ext	Q2.14exthd
Q2.14isb	Q2.14isbs
Q2.14isf	Q2.14isfs
Q2.14myd	Q2.14mydoc
Q2.14sch	Q2.14schlfs
Q2.14tap	Q2.14tape
Q2.16elo	Q2.16elog
Q2.16loc	Q2.16locdb
Q2.16pap	Q2.16paplog
Q2.16rem	Q2.16remdb
Q2.16spr	Q2.16sprdsht
Q2.17Bco	Q2.17Bconf
Q2.17Bdo	Q2.17Bdoc
Q2.17Bfo	Q2.17Bform
Q2.17Bli	Q2.17Blicen
Q2.17Bot	Q2.17Both
Q2.17Bsh	Q2.17Bsharenr
Q2.18Bco	Q2.18Bconf
Q2.18Bdo	Q2.18Bdoc
Q2.18Bfo	Q2.18Bform
Q2.18Bli	Q2.18Blicen
Q2.18Bot	Q2.18Both
Q2.18Bsh	Q2.18Bsharenr
Q2.2obse	Q2.2observ
Q2.3audi	Q2.3audiotape
Q2.3data	Q2.3database
Q2.3di_1	Q2.3digitalvideo

Q2.3digi	Q2.3digitalaudio
Q2.3exce	Q2.3excel
Q2.3fiel	Q2.3field
Q2.3imag	Q2.3image
Q2.3powe	Q2.3powerpoint
Q2.3slid	Q2.3slides
Q2.3vide	Q2.3video
Q2.6resa	Q2.6resas
Q2.7natd	Q2.7natdatac
Q2.7nobo	Q2.7nobody
Q2.7resg	Q2.7resgp
Q2.8exth	Q2.8exthd
Q2.8exts	Q2.8extstor
Q2.8mydo	Q2.8mydoc
Q2.8othf	Q2.8othfs

Michael Whitton Aug 2011 mw2@soton.ac.uk

The Institutional Data Management Blueprint project is funded by JISC under the Managing Research Data Programme.

Acknowledgements: JISC programme manager, Simon Hodson; University of Southampton project contributors

<http://www.southamptondata.org/about-us.html>