Developing an audit methodology suited to research data assets

Sarah Jones, Seamus Ross, Raivo Ruusalepp

Data Audit Framework Launch
British Academy, London
1st October 2008
The problem

How can organisations realise the value of their research data assets when it is unclear:

• what data is held;

• where it is located;

• and how it is being managed?
A recommendation

“JISC should develop a Data Audit Framework to enable all universities and colleges to carry out an audit of departmental data collections, awareness, policies and practice for data curation and preservation”

Expected benefits

Broadly speaking, auditing data brings three main benefits:

• prioritisation of resources (efficiency savings)
• ability to manage risks associated with data loss and irretrievability;
• realising the value of data through improved access and reuse
Development questions

- What is the scope of the Framework?
- Is it applied at department or institutional level
- What is a ‘data asset’?
- Who should do the audit?

- What information should be collected?
- How should this information be collected?
Methodology

Workflow and tasks
Stage 1. Planning the audit

START: Decision to conduct a Data Asset Audit

Organisation

1.1 Appoint an auditor

1.2 Establish a business case to approach the organisation

1.3 Conduct initial research to plan the audit

1.4 Set up the audit

1.4.1 Ask for documentation about the organisation

1.4.2 Fill in Audit Form 1

Audited organisation

Collector and send the relevant documentation

Auditor

1-2 days

Elapsed Auditor time

1 week
Stage 2. Identifying and classifying data assets

**Organisation**

- 2.1 Analyse documentation
  - 2.1.2 Start filling in Audit Form 2
- Can organisation staff be interviewed extensively?
  - Yes: 2.3 Prepare interview questions
  - No: 2.2 Prepare a survey questionnaire

**Auditor**

- 2.2.1 Identify staff who should answer the questionnaire and send the survey
- 2.3.1 Identify staff who should be interviewed and agree interview times
- 2.3.2 Complete Audit Form 2, if possible start filling Audit Form 3

**Audit Form 2**

- Inventory of data assets
- For audit

**Audit Form 3**

- Data asset assessment
- Audit Form 2, Inventory of data assets

**Approve data asset classification and assessment**

**Set up a meeting with organisation's management to assess data assets**

**Finalise classification and assessment of data assets**

**4-6 days**

**2 weeks**
Stage 3. Assessing the management of data assets

Organisation

3.1 Prepare interview questions
3.2 Complete Audit Form 3 for each data asset
Answer interview questions

Auditor

3.1.1 Identify staff who should be interviewed and agree interview times
Audit Form 3: Data asset assessment

Elapsed Auditor time

3-4 days

2 weeks
Stage 4: Reporting results and making recommendations

Organisation:
- Act on the audit findings to improve the environment of data assets
- END: Audit completed

Auditor:
- 4.1 Assess data assets and produce audit report
- 4.2 Amend audit report if necessary
- Discuss the audit findings and report
- Approve the audit findings and report

Elapsed time:
- Auditor time: 2-3 days
- 1 week
Inventory of data assets

- Name of the data asset
- Description
- Asset manager (responsibility)
- Reference (location)
- Classification
- Comments (general and on classification)
Detailed assessment

- ID
- Data creator(s)
- Title
- Description
- Subject
- Date
- Purpose
- Source
- Updating frequency
- Type
- Format
- Rights and restrictions
- Usage frequency
- Relation
- Back-up and archiving policy
- Management to date
Testing the method

- Glasgow University Archaeological Research Division (GUARD)
- Innovative Design and Manufacturing Research Centre (IdMRC) at University of Bath
- School of GeoSciences at University of Edinburgh
Lessons learned

- Ensure appropriate timing
- Plan well in advance
- Adopt the most appropriate method
- Scope the work carefully
- Collect additional information early on
Redevelopment

- Identifying and classifying now one stage
- Differentiation made between man hours and elapsed time
- Guidance added on scoping / granularity
- Element names changed
Broader significance?

the loss was entirely avoidable

officials ignored repeated warnings about serious flaws in data security

woefully inadequate processes for handling data

criminally careless government departments

a “muddle through" culture

horrendous carelessness with highly sensitive information

slack practices