Digital Data and Archaeology: Management, Preservation and Publishing

Dr Katie Green
katie.green@york.ac.uk
Outline

- Digital Data and Archaeology
- Digital Data Challenges
- Case studies of bad practice
- Importance of data management and preservation
Archaeology

- Archaeology is **destructive**
- Comprehensive **records** of field work are **imperative**
Archaeological Data

- Published data is limited
- Majority of data recorded in grey literature
- Data is not easily searchable
- Assess to data is difficult
- Move from large data rich publications to **slim syntheses**
Digital Data

- Easier to create
- Easier to update
- Easier to access

Born Data - data created in digital format

Digitised Data - hardcopy converted to digital format

Image © www.digitalbevaring.dk
Software

- Lots of formats

Software used by archaeologists (identified more than once)
Software

- Lots of formats
- Become out of date rapidly

Software used by archaeologists (identified more than once)

Big Data Project

- 3D Studio Max
- ArcGIS
- AutoCAD
- BAE SOCETSET
- CODA
- ENVI / IDL
- ERDAS Imagine
- Golden Software Surfer
- Leica Cyclone
- MicroStation
- Pointools
- Polyworks
- RapidForm
- TerraScan
- Trimble Realworks
- Custom software
- MySQL
My lithics report is here, on floppy disc
Digital Data

- Easier to create
- Easier to update
- Easier to access

But Fragile!

- Digital data is encoded
- Digital data requires software and technology to present content
Why is Digital Data Fragile?
Why is Digital Data Fragile?

• **Deterioration** of the storage medium
  – Degrade – Bit rot!
  – Can be easily damaged
  – Can be easily overwritten
Case Study: NASA – again!

One giant blunder for mankind: how NASA lost moon pictures
Case Study: NASA – again!

One giant blunder for mankind: how NASA lost moon pictures

Copied over the Moon Landing tapes
Why is Digital Data Fragile?

• Deterioration of the storage medium
• Obsolescence of the software
Case Study: NASA

- NASA sent two Viking Landers to Mars in 1975
- Data recorded on magnetic tape
- Climate controlled environment

Photos: Courtesy NASA/JPL-Caltech
Case Study: NASA

- NASA sent two Viking Landers to Mars in 1975
- Data recorded on magnetic tape
- Climate controlled environment
- In the 1990s they could not decode the formats used
- Had to track down old printouts and retype everything

Photos: Courtesy NASA/JPL-Caltech
Software

• Lots of formats
• Become out of date rapidly

Big Data Project

Software used by archaeologists (identified more than once)
Why is Digital Data Fragile?

- **Deterioration** of the storage medium
- Obsolescence of the **software**
- **Obsolescence** of the storage medium
Preservation problems

Media Types

• Experience rapid change
Why is Digital Data Fragile?

- **Deterioration** of the storage medium
- Obsolescence of the **software**
- Obsolescence of the **storage medium**
- Obsolescence of the **hardware**
Technology

- Hardware experiences rapid change
Case Study: BBC Domesday Project

- 1986
- photographs, maps, etc
- 30cm laserdiscs
- BBC Microcomputers

http://www.bbc.co.uk/history/domesday/story
Case Study: BBC Domesday Project

- 1986
- photographs, maps, etc
- 30cm laserdiscs
- BBC Microcomputers
- In 2006 the laserdiscs were **obsolete** as was the hardware
- Rescue projects launched by The National Archives and Leeds University

http://www.bbc.co.uk/history/domesday/story
Why is Digital Data Fragile?

• Deterioration of the storage medium

• Obsolescence of the software

• Obsolescence of the storage medium

• Obsolescence of the hardware

• Failure to document data adequately
Case Study: Newham Museum Archaeological Service

Archive:
- approx. 150 excavations
- 6432 individual files
- 1500 excavation reports
- 700 database files
- 1200 geophysics files
- 200 separate projects
Case Study: Newham Museum Archaeological Service

Problems:

- No data structure
- No file naming standards
- No metadata
- No data documentation at all
Case Study: Silbury Hill

http://archaeologydataservice.ac.uk/blog/2013/08/jenny-ryders-day-of-archaeology-at-the-ads-a-silbury-hill-update/
Case Study: Newham Museum Archaeological Service

Lessons Learnt:

• Regular software migration needed
• Robust data documentation down to file level needed
• Need to plan for re-use
• Forward planning is much cheaper
• Data management strategy required

Image © www.digitalbevaring.dk
"Digital information lasts forever - or five years, whichever comes first."

(Jeff Rothenberg, RAND Corp., 1997)
Archaeological Data

2,000 years in the making
3 days to record
Backed up in 10 seconds
Lost forever?

How can we prevent this?
Protecting Digital Data

• Recognise data is as fragile as the archaeological record we excavate
• Stop archiving data as objects rather than computerised information
• Recognise the challenges of digital data
• Professionally archive digital material
• Create **Data Management Plans**

**GOOD DATA MANAGEMENT IS ESSENTIAL**