

# Library vs publisher vs archive: managing the UK topographic record



Kimberly C. Kowal  
Curator of Digital Mapping  
The British Library  
kimberly.kowal@bl.uk

Jonathan Holmes  
NGD & Corporate Data Manager  
Ordnance Survey  
jonathan.holmes@ordnancesurvey.co.uk



# Non-Print Legal Deposit Regulations

- By law, a copy of every UK print publication must be given to the British Library by its publishers, and to five other major libraries that request it. This system is called legal deposit and has been a part of English law since 1662.
- As of 6 April 2013, legal deposit also covers material published digitally and online, so that the Legal Deposit Libraries can provide a national archive of the UK's non-print published material, such as websites, blogs, e-journals and CD-ROMs.



CAMBRIDGE UNIVERSITY LIBRARY



## UK Legal Deposit Libraries

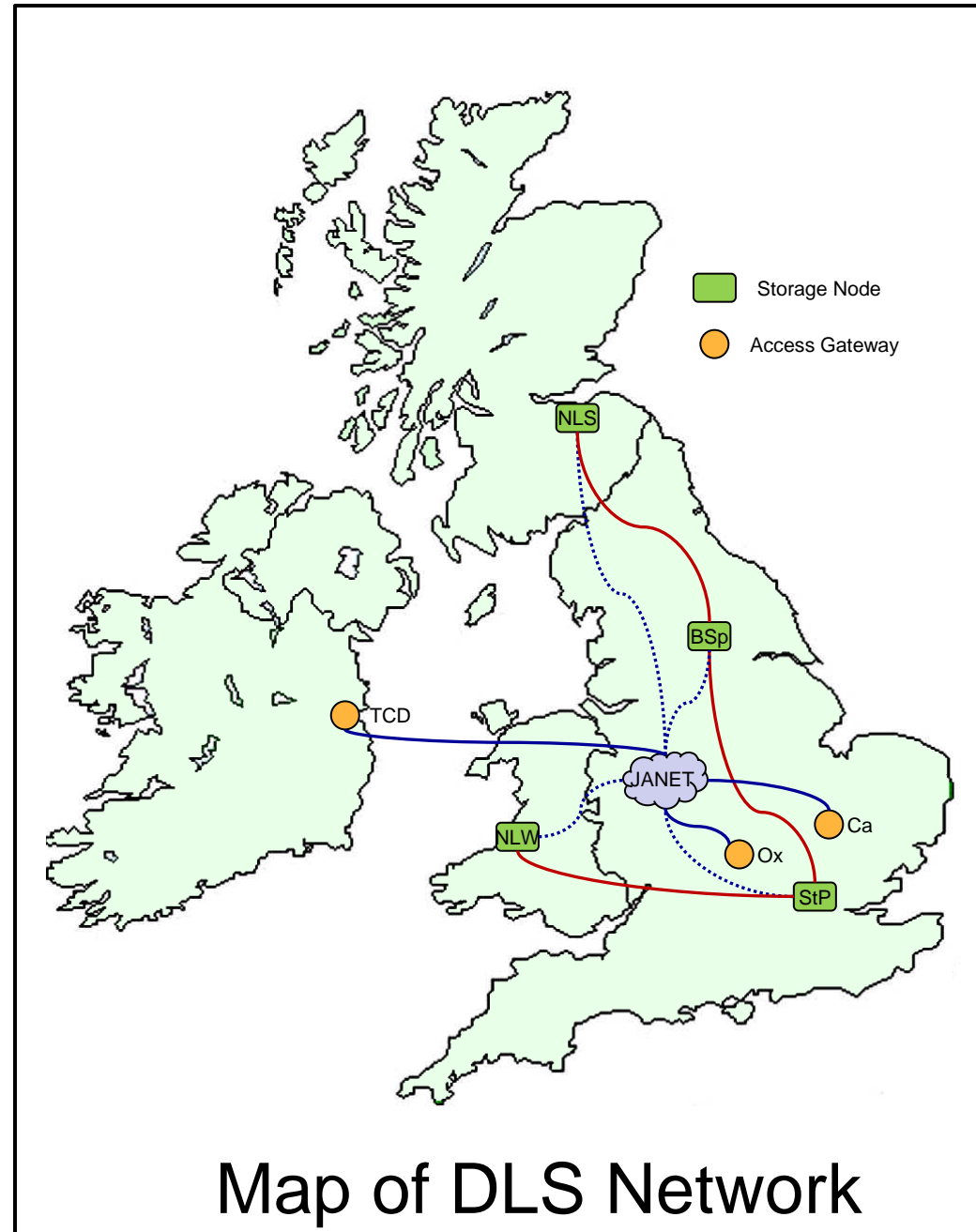
- The British Library
- Bodleian Library, Oxford
- Cambridge University Library
- National Library of Scotland
- National Library of Wales
- Trinity College, Dublin



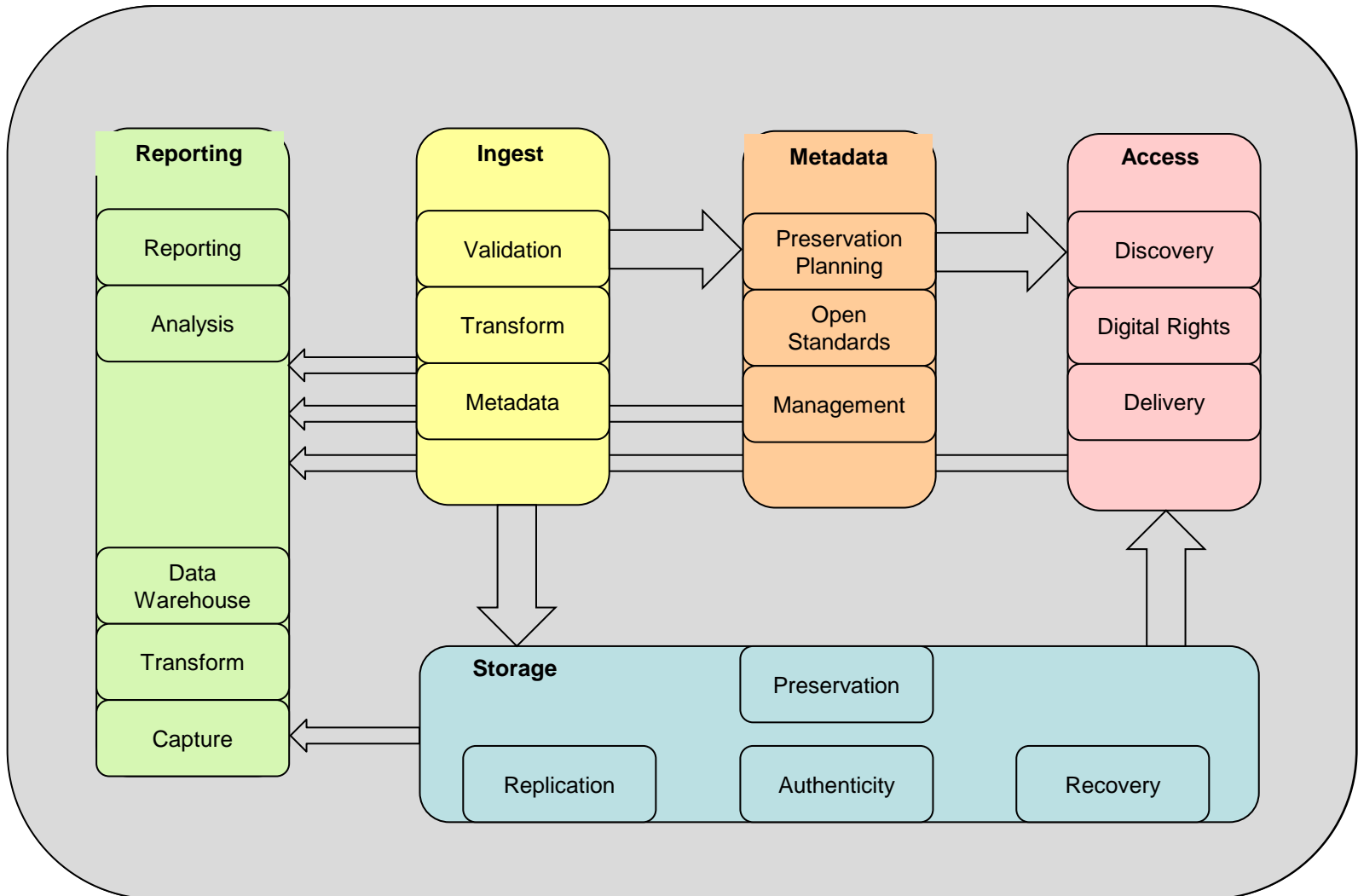
Bodleian Library  
UNIVERSITY OF OXFORD

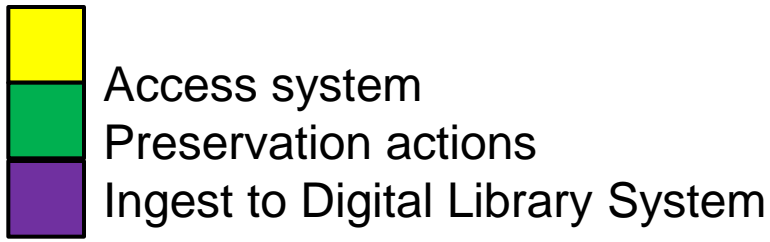
# Digital Library System Mission

- Provide a secure place to store digital material created and acquired by the BL, forever
- Acquire Digital Legal Deposit Material on behalf of the other Legal Deposit Libraries and share it with them
- Ensure that all material is permanently findable and accessible
- Provide the widest possible access whilst respecting the wishes of rights holders
- Make Digital 'Business as Usual' at the BL

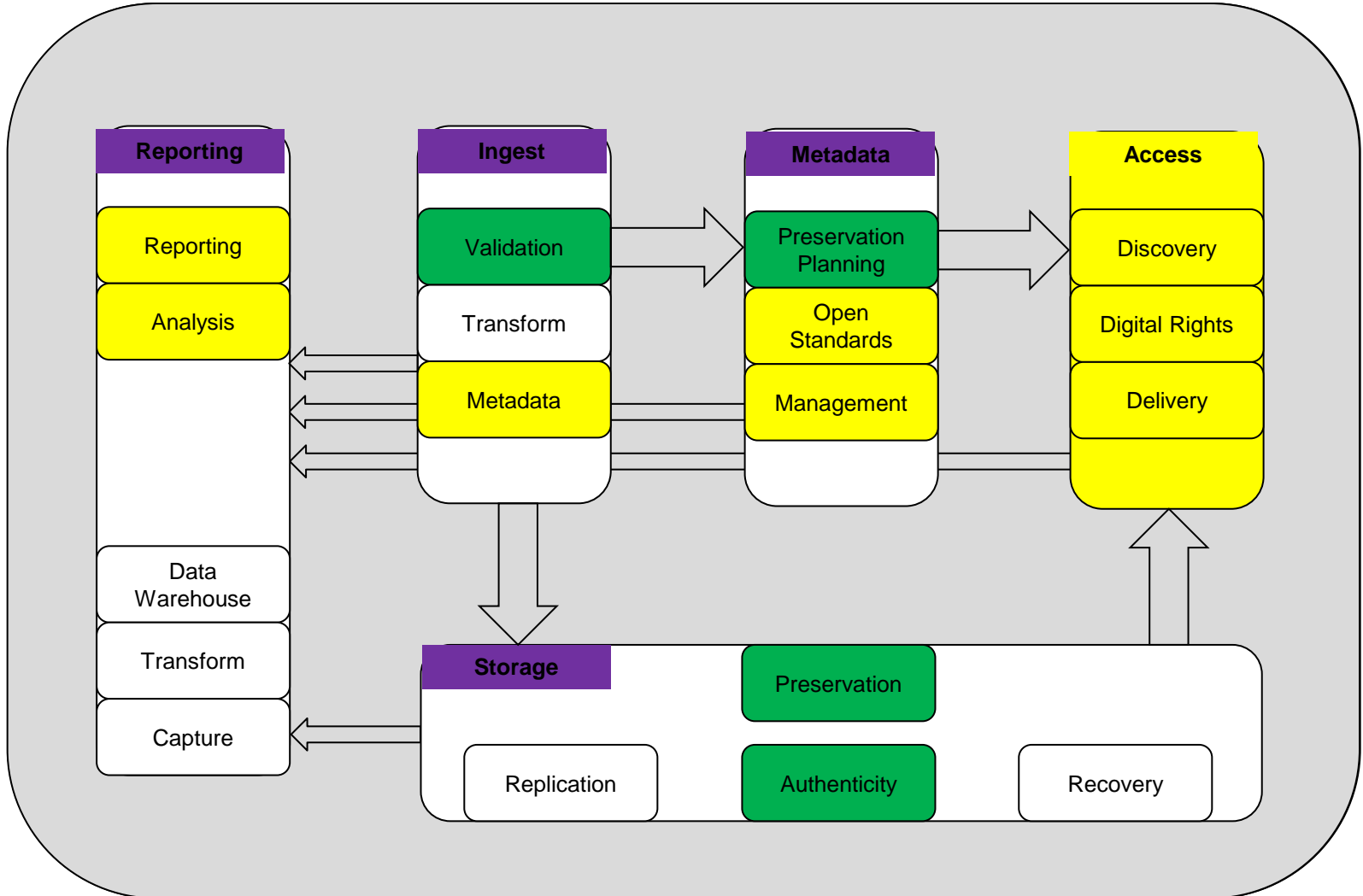


# DLS Software Components





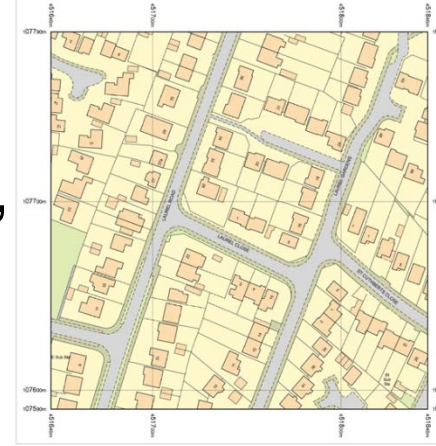
# Components-spread



# Ordnance Survey (OS) digital data

Commercially valuable, large scale mapping of GB,  
source scales 1:1,250; 1:2,500; 1:10,000

- 1998-2007 Land-Line (ntf)
- Land-Form Profile (dtm)
- 2006- onwards OS MasterMap Topography Layer<sup>®</sup> (gml)
- 2006- onwards OS MasterMap<sup>®</sup> Integrated Transport Network<sup>™</sup> (ITN) Layer



“In the spirit of long-term archiving”

Obligations and Restrictions:

Required secure, user rights management framework  
to control application and data access

Access limited to readers inside the deposit libraries

Printing – A4 extracts for non-commercial users

Digital copying for the Library (*long term preservation*)

No downloading

Usage and security reporting

# Geospatial datasets: Access

**Software 1 - standalone application with data stored in each LDL, 1999-2006**

**Software 2 - networked customised web mapping application, hosted and managed, 2006-2013**

Cooperative arrangement with formal LDL approval

•European Union procurement - 'restricted procedure' - Jan - May 2006

**Software 3 - networked customised web mapping application, hosted and managed 2014-2019**

Cooperative arrangement with formal LDL approval

•European Union procurement - 'restricted procedure' - Jan - Aug 2012

The system was to provide a generic and cost-effective infrastructure designed to:

- Be flexible and scalable, accepting additional data of various types and from numerous sources
- Take advantage of open source technologies, which are cheaper, more effective, and widely used
- Hold metadata about map content in a more open, accessible form, allowing for easy transfer into other repository systems
- Allow controlled access to accommodate legal deposit regulations
- Reflect developments in web-mapping technologies and improvements to the map viewing interface

# Geospatial datasets: Preservation



## Scalable Preservation Environment

- SCAPE: EU project to develop tools and services to for large-scale digital repositories
- Research Datasets testbed: Identification, validation and checksumming of a complex corpus
  - A tool to create/check fixity of files
  - A tool to identify the files in the corpus
  - A tool to validate files adhere to the format specifications, where appropriate
- Integration of SCAPE toolset with DLS



# Geospatial datasets: Digital Library System

## Metadata + Ingest

### Creating a new Content Stream for Geospatial Data

- The Library uses the Metadata Encoding and Transmission Standard (METS) schema to convey the metadata required to manage the digital objects held within the Digital Library System (DLS).
- **DLS Strategic Ingest:** Metadata Extension Repository is meant to cover all content streams
- Defining a Content stream:
  - mapping onto the SI METS profile.
  - creation of data model
  - developing software to create METS file and ingest

# Geospatial datasets: Digital Library System

## Ingest + Metadata

### Creating a new Content Stream for Geospatial Data

- What makes it different from other library materials?
- Usage unfamiliar
- Diverse formats, complex relationships
- Unfamiliar descriptors
  - Coordinate reference systems
  - Cartographic representation
  - Topology (data structures)
  - Project files, data packaging

# How do the data producers handle managing data in a repository

- The Ordnance Survey
- The National Archives

Over to Jonathan...

# Background



- Ordnance Survey
  - Founded in 1791
  - Responsibility for mapping Great Britain (England, Scotland, Wales)
  - Currently a Government Trading Fund
- The National Archives (TNA)
  - Formerly the Public Records Office
  - Founded in 1838 to reform the keeping of government and court records



# Background

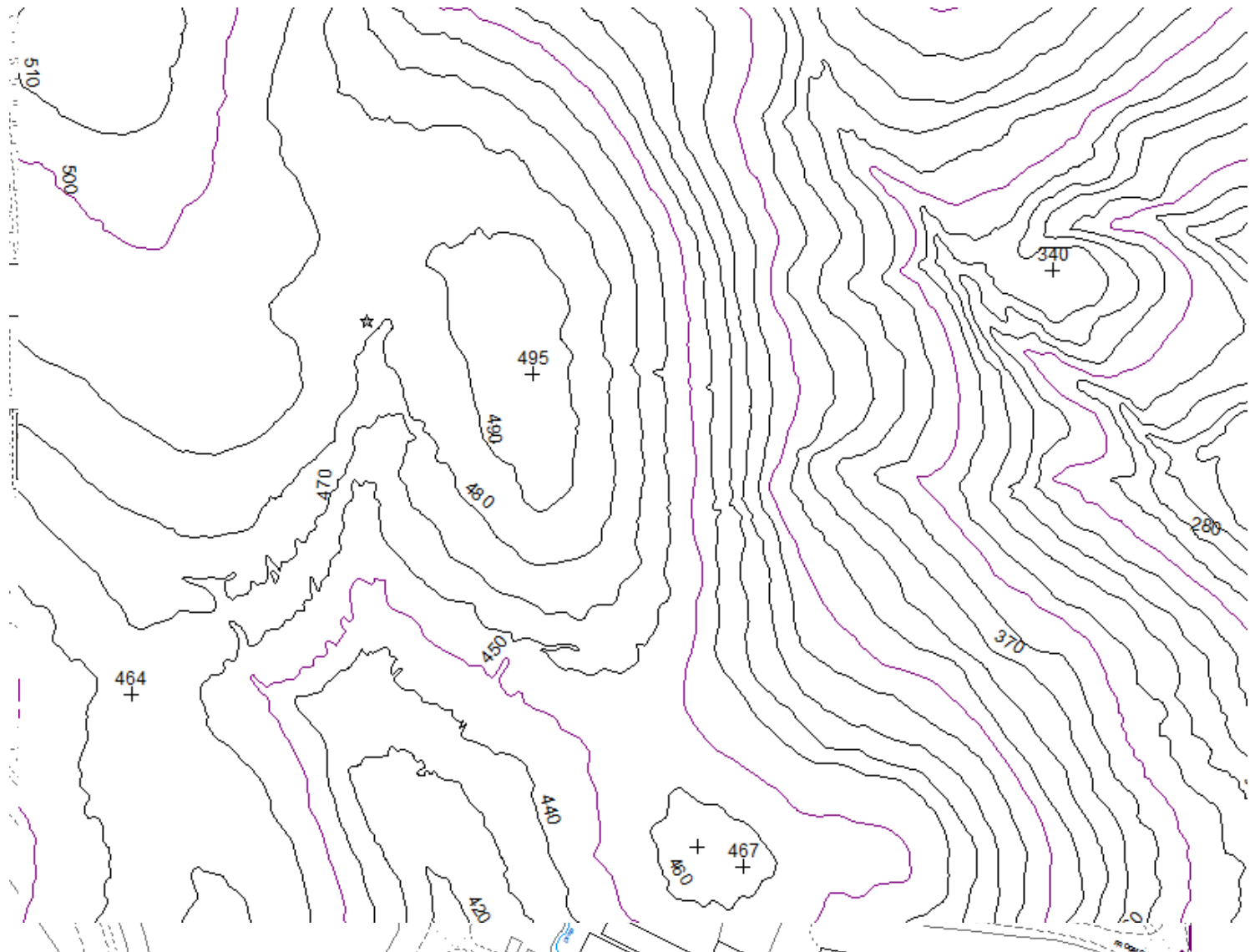


- In 1999 Ordnance Survey and The National Archives signed an agreement for Ordnance Survey to become the first Electronic Place of Deposit (POD)
- Stores Landline, Landform Profile and OS MasterMap

# Landline



# Landform Profile



# OS MasterMap



Slide





# Why Ordnance Survey became a PoD



- `<osgb:descriptiveGroup>Network Or Polygon Closing Geometry</osgb:descriptiveGroup>`
- `<osgb:descriptiveTerm>Polygon Closing Link</osgb:descriptiveTerm>`
- `<osgb:nonBoundingLine>true</osgb:nonBoundingLine>`
- `<osgb:physicalLevel>50</osgb:physicalLevel>`
- `<osgb:physicalPresence>Closing</osgb:physicalPresence>`
- `<osgb:polyline>`
- `<gml:LineString srsName='osgb:BNG'>`
- `<gml:coordinates>403999.020,411237.110 404000.000,411237.260 </gml:coordinates>`
- `</gml:LineString>`
- `</osgb:polyline>`
- `</osgb:TopographicLine>`
- `</osgb:topographicMember>`
- `<osgb:topographicMember>`
- `<osgb:TopographicLine fid='osgb1000002102987241'>`
- `<osgb:featureCode>10113</osgb:featureCode>`
- `<osgb:version>1</osgb:version>`
- `<osgb:versionDate>2011-05-03</osgb:versionDate>`
- `<osgb:accuracyOfPosition>1.0m</osgb:accuracyOfPosition>`
- `<osgb:changeHistory>`
- `<osgb:changeDate>2011-05-03</osgb:changeDate>`
- `<osgb:reasonForChange>New</osgb:reasonForChange>`
- `</osgb:changeHistory>`
- `<osgb:descriptiveGroup>Network Or Polygon Closing Geometry</osgb:descriptiveGroup>`
- `<osgb:descriptiveTerm>Polygon Closing Link</osgb:descriptiveTerm>`
- `<osgb:nonBoundingLine>true</osgb:nonBoundingLine>`
- `<osgb:physicalLevel>50</osgb:physicalLevel>`

# Place of Deposit System



- Written in-house using Visual Basic and Unix
- Simple system
- Ingests a snapshot of the live product
- Two backups are made of the data
  - One stored locally and one remotely
- File structure is also simple
  - Product
    - Year
      - 100km grid square

# Place of Deposit System



- Data specification and product specifications are stored in the system
- Metadata – very limited
- There is no viewing tool. Data has to be extracted and loaded into a GIS

# Ordnance Survey – Place of Deposit



**Ordnance Survey  
Place of Deposit**

Ingest Data

Create CRC Values

Check CRC Values

Exit

V2.1 Mar 2012

A screenshot of a software menu titled 'Ordnance Survey Place of Deposit'. The menu is displayed on a light blue background and contains four cyan buttons with black text: 'Ingest Data', 'Create CRC Values', 'Check CRC Values', and 'Exit'. The version number 'V2.1 Mar 2012' is located in the bottom right corner of the menu area.

# System use



- Data is used internally
- Used externally in response to Freedom of Information requests
- By 2019 – Need to start to make data archived for TNA available for viewing publicly
- Under the terms of the Act very limited copying for the purposes of:
  - Private study;
  - Non-commercial research, criticism, review and news reporting, (provided that any copy is accompanied by the following acknowledgement ©Crown Copyright. Reproduced by permission of Ordnance Survey®;or for
  - Parliamentary or judicial proceedings.
- How we do this is yet to be determined

# Challenges for Ordnance Survey



- Only some of the products are stored in the PoD
- Many of the other products are strewn across drives, networks and servers or in cupboards.
- Much of this data has a short / medium term use internally but it is not stored securely.
- Storage and maintenance in the medium term requires a more sophisticated system than we currently have.
- What do we do with the data longer term?
- Should the archiving processes of Ordnance Survey and the British Library be linked with some sort of accession plan?

# Conclusion



- OS, TNA and the LDLs are three separate bodies with different but similar obligations. There are clear synergies between them and there are clearly opportunities for the three organisations to work together.
- Should the same data be held in more than one place for more than one purpose? This seems inefficient.
- Can this data just be held once for the nation?