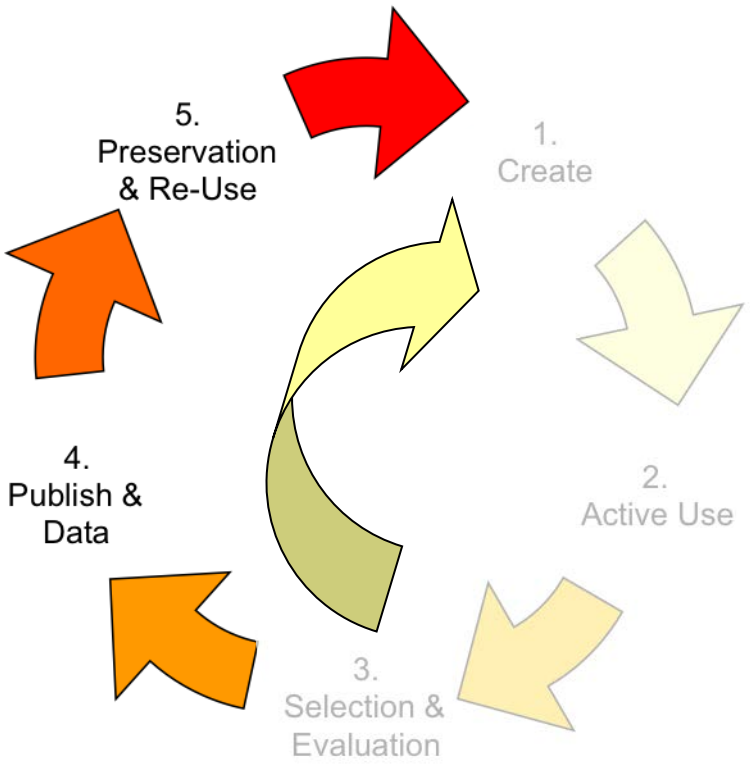
Archiving and Repositories

Holly Wright

ARIADNE is funded by the European Commission's Seventh Framework Programme

Data archiving in the digital lifecycle



Where data archiving fits

- At the start of the project - identifying the archive or repository to be used
- Before starting data capture – check archive requirements
- During the project – documentation + deposit
- On completion – deposit
- After the project – enabling access and reuse

Overview

Why Deposit?

1. Ensure Preservation
2. Provide Access
3. Professional Recognition
4. Follow Professional Standards
5. Meet Governmental Requirements
6. Meet Funding Agency Requirements

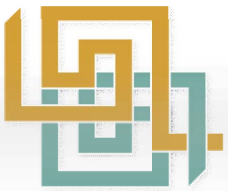
Data Deposit Requirements

AHRC

- Submission of a 'Technical plan'
  replaces the Technical Appendix as of 2012

- is essential:
  "where digital outputs or digital technologies are an essential part to the planned research outcomes".

  This should give a summary of those outputs, explain the technical methodology, technical support / experience, and address preservation, sustainability and use.



Archiving

What happens to data once its deposited?

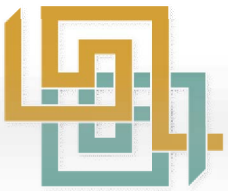
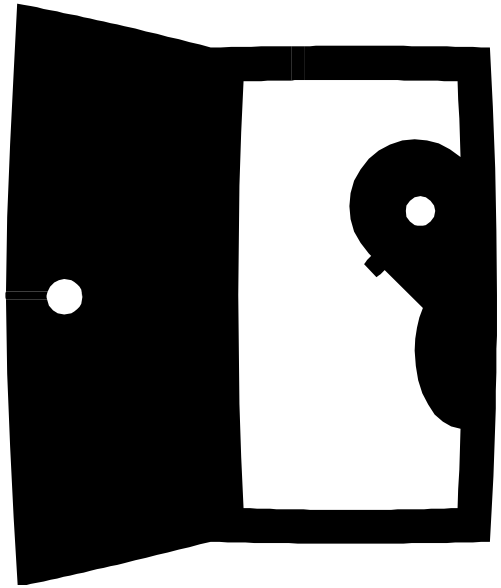


Image © www.digitalbevaring.dk

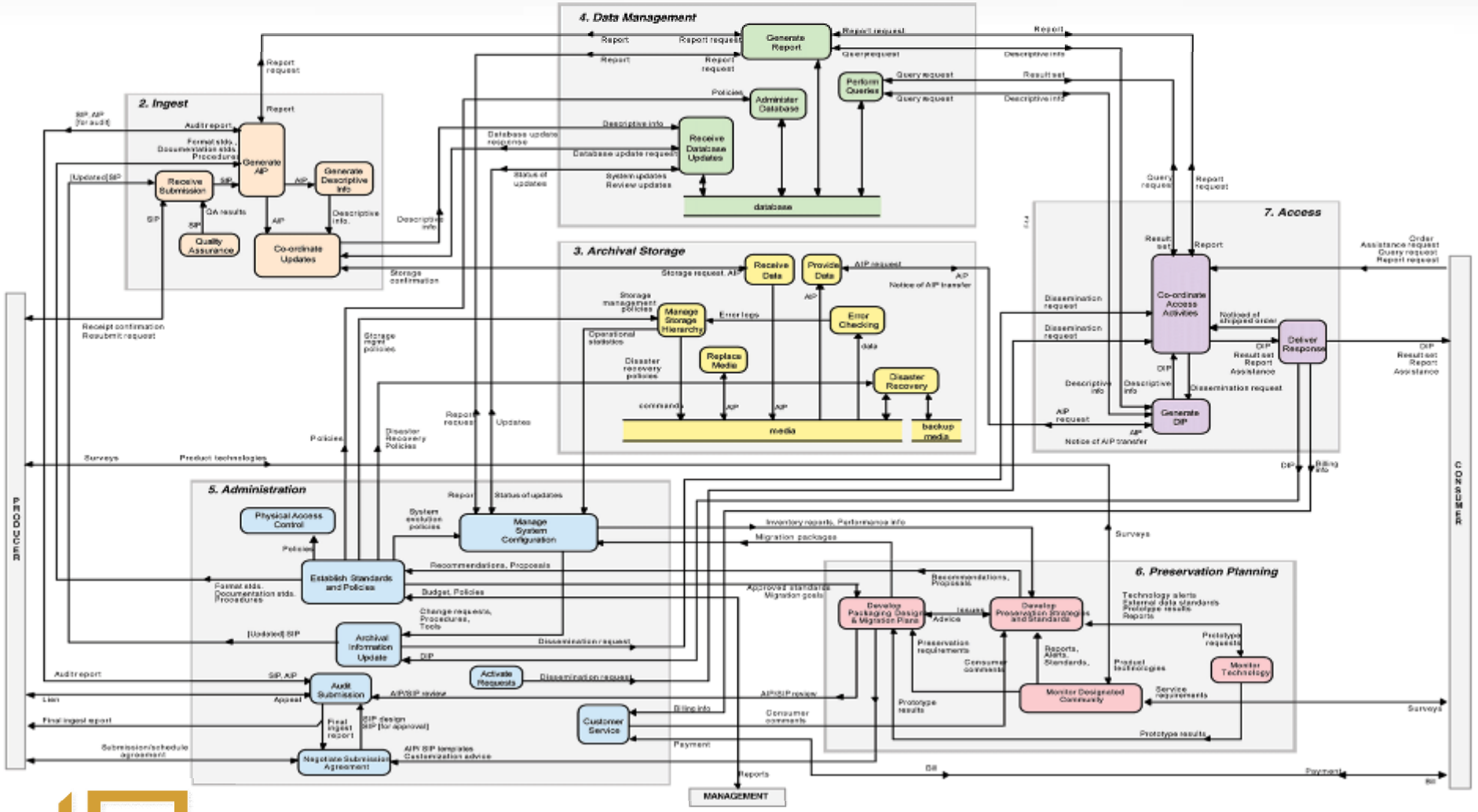
Archiving

Behind the Scenes at the ADS

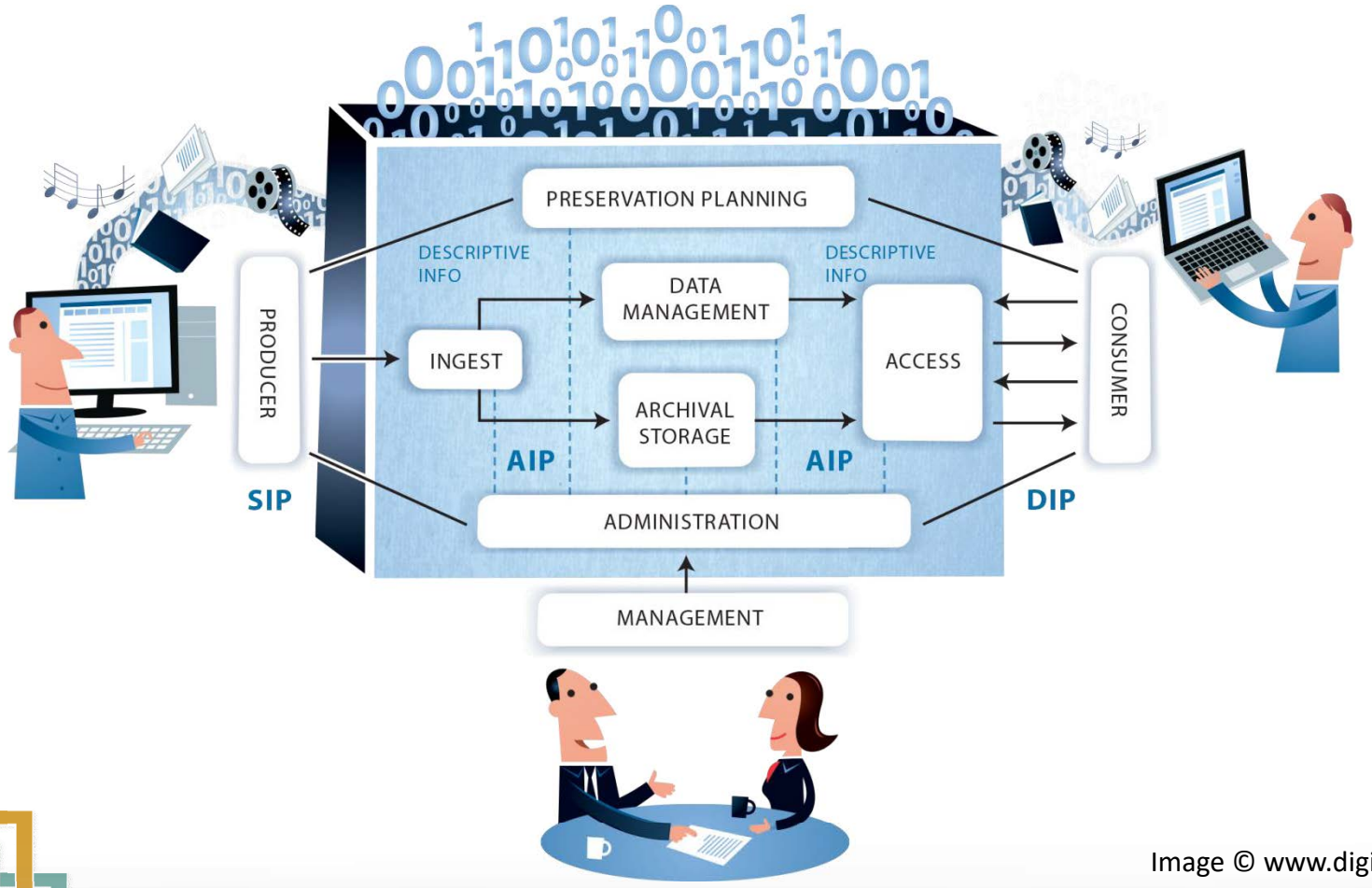
- Follow the **Open Archival Information System** (OAIS) reference model
  - International ISO standard 14721
- Ensure the multiple and regular **backups** and the **renewal** of storage media
  - Virtual Server
  - Tape backup at University of York & Hull
  - Deep Store
  - Five year rotation strategy
- Use data **migration** strategies



Archiving


4. Data Management
Generate Report
Perform Queries
Administer Database
Receive Database Updates
database
2. Ingest
Receive Submission
Generate AIP
Generate Descriptive Info
Quality Assurance
Co-ordinate Updates
3. Archival Storage
Receive Data
Provide Data
Manage Storage Hierarchy
Error Checking
Replace Media
Disaster Recovery
media
backup media
7. Access
Co-ordinate Access Activities
Deliver Response
Generate DIP
5. Administration
Physical Access Control
Manage System Configuration
Establish Standards and Policies
Archival Information Update
Activate Requests
Audit Submission
Negotiate Submission Agreement
Customer Service
6. Preservation Planning
Develop Packaging Designs & Migration Plans
Develop Preservation Strategies and Standards
Monitor Designated Community
Monitor Technology
MANAGEMENT

Open Archival Information System (OAIS) reference model



Open Archival Information System (OAIS) reference model

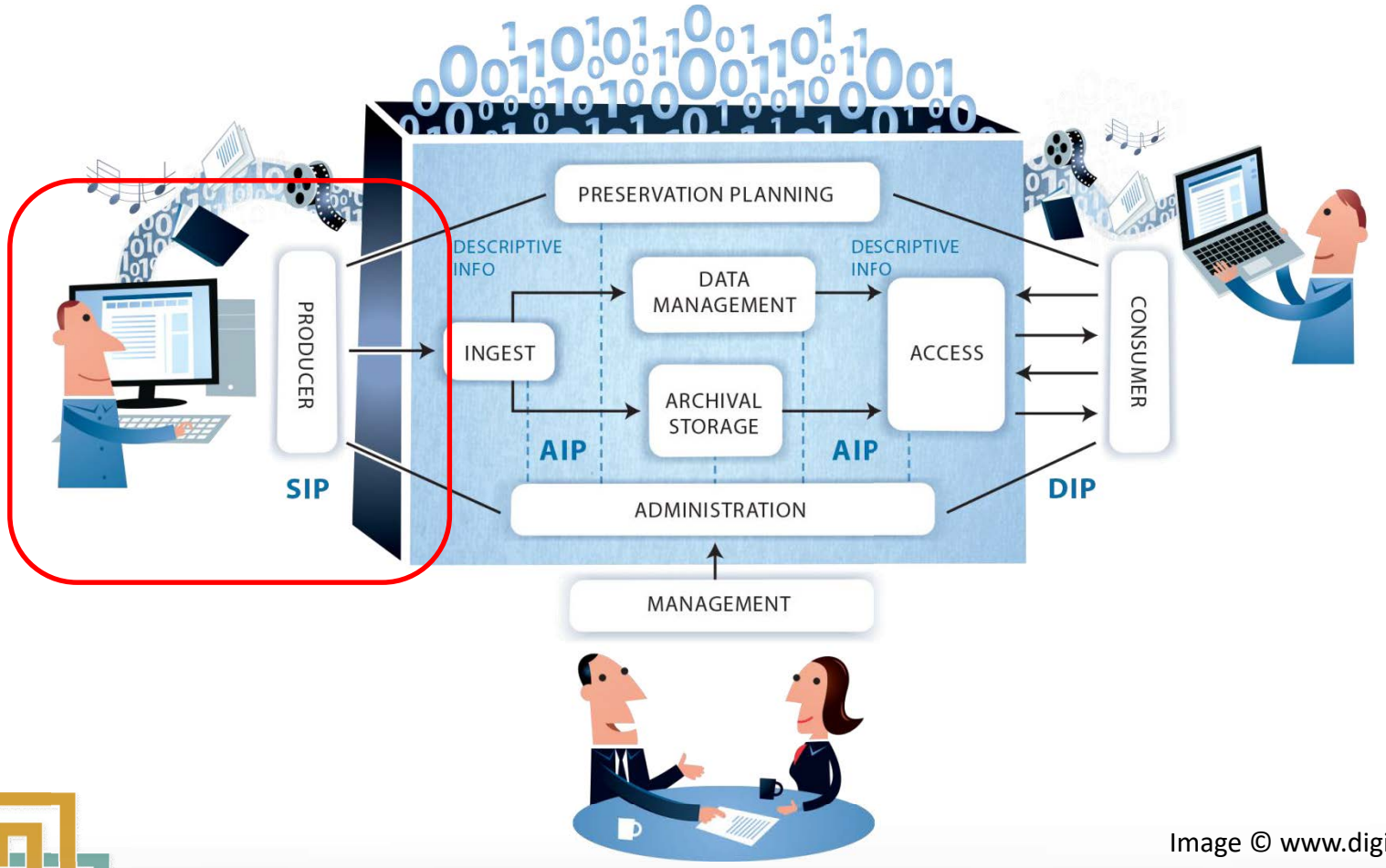
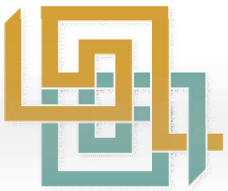


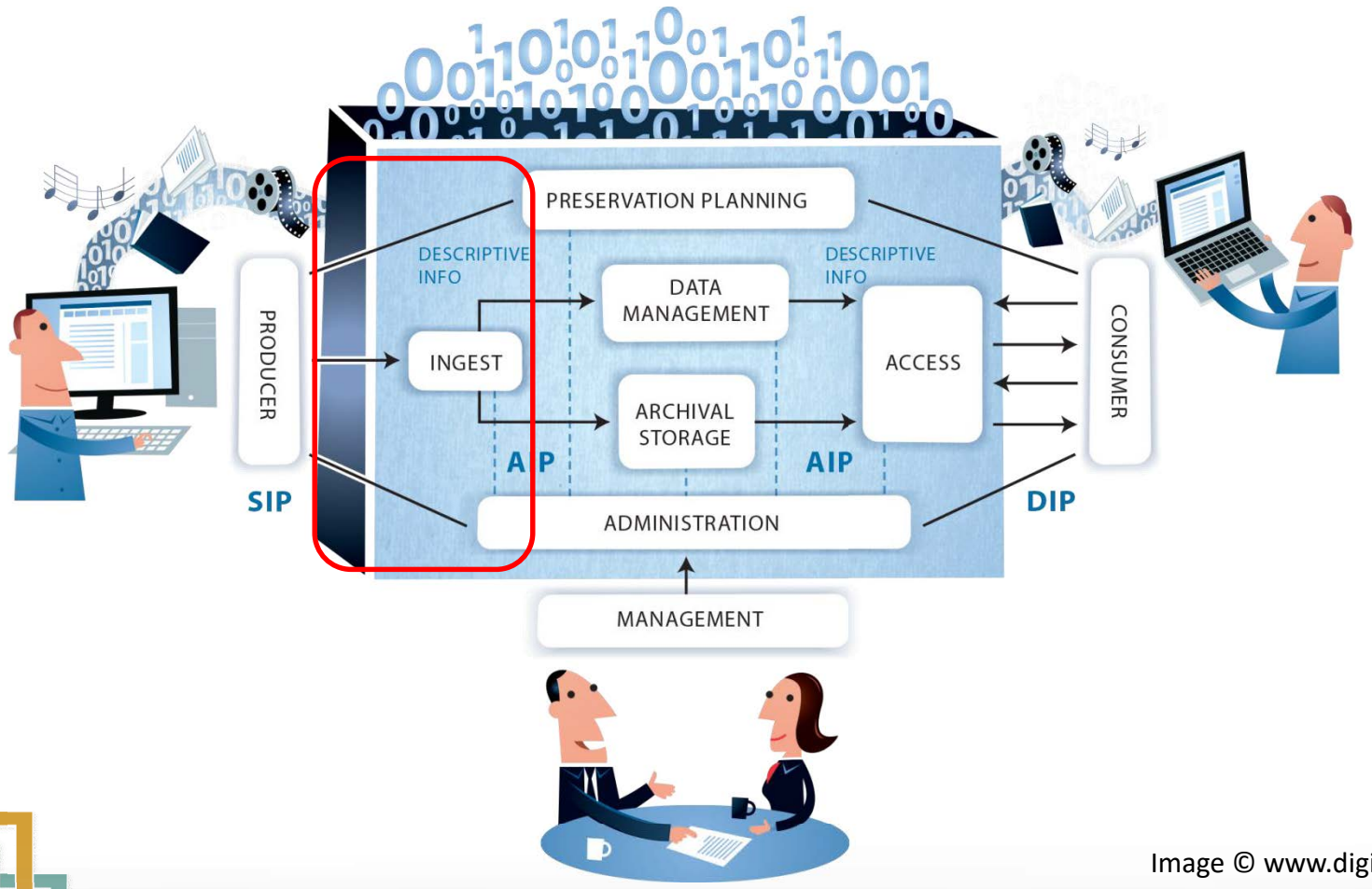
Image © www.digitalbevaring.dk

Deposit Evaluation

- Intellectual content & potential interest in their re-use
- Viability of data: management, preservation and dissemination
- Are we a suitable archive?
- Authority to deposit the data
- Material is ‘complete’
- Digital form in preferred file format – consult repository websites
- Sufficient project documentation and file Metadata



Archiving



Submission Information Package (SIP)

- **Virus check**
- Media and file **readability** check
- Data resource **integrity** check
- Check file **formats** suitable for deposit
- Documentation **completeness** check
- Data **validation** and **consistency** checks
- Web interface text check
- Copy to data server

Submission Information Package (SIP)

- Authenticate original version
- Replace spaces with underscores
- Log details of SIP in **Collections Management System**
- Create **checksums**
- Run **Droid** to generate file level metadata
- Store **licence**
- Scan paper documentation
- Acknowledge receipt of data
- Store original media

Archiving

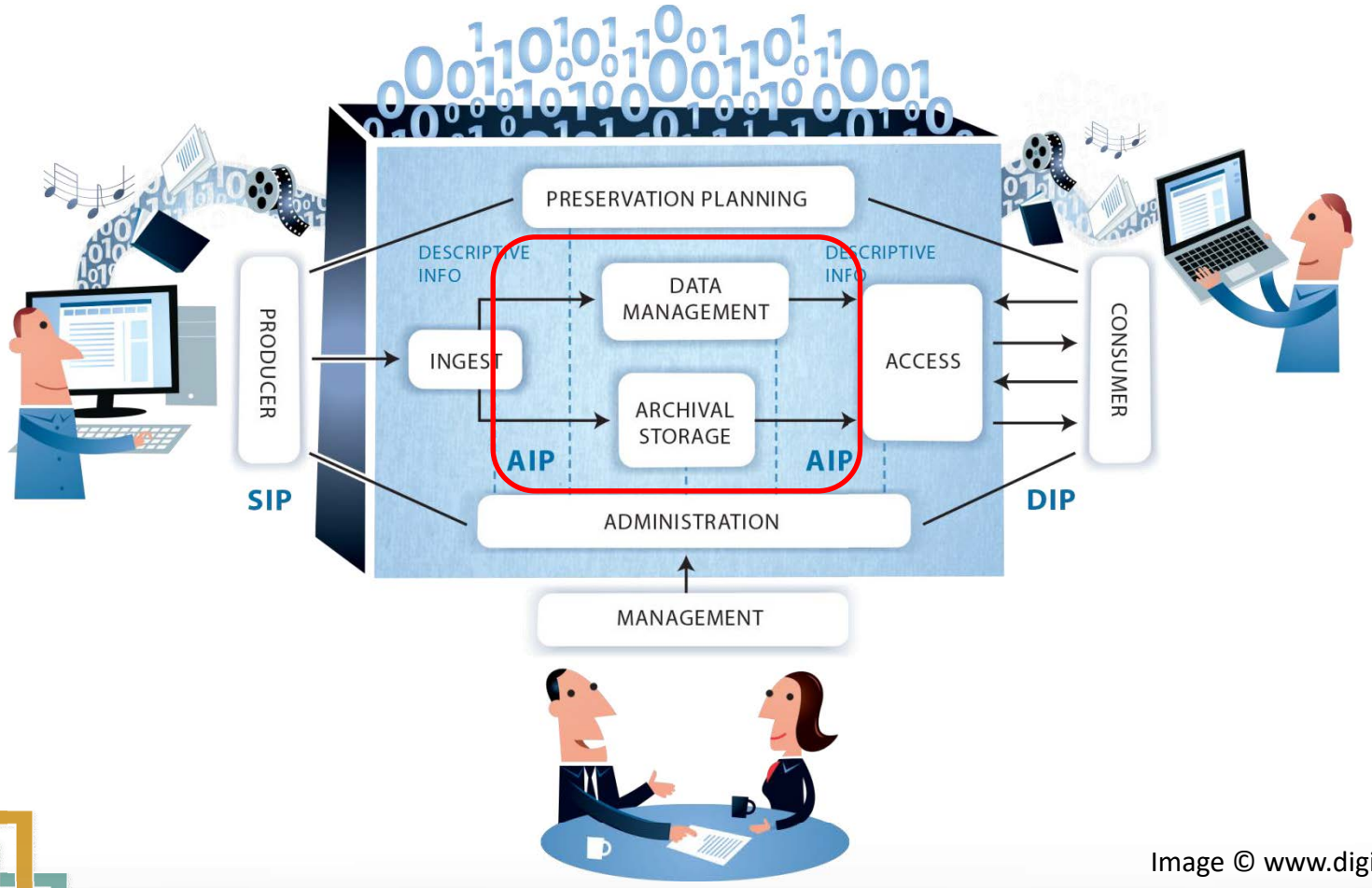
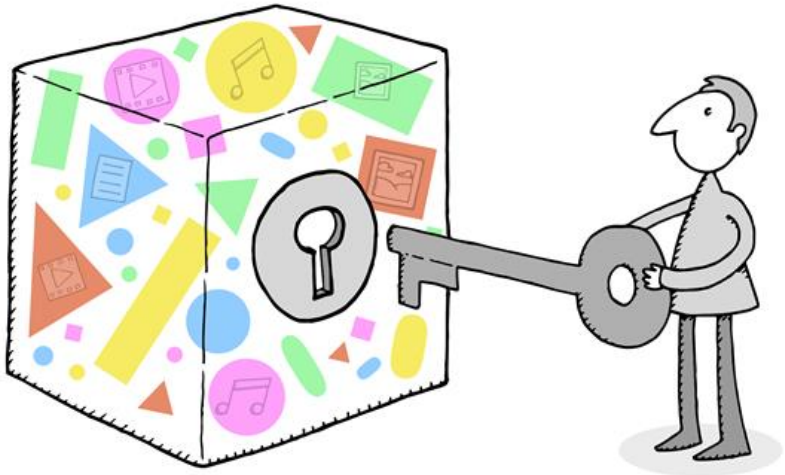


Image © www.digitalbevaring.dk

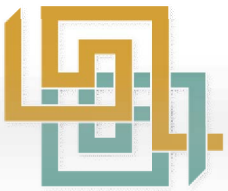
Archival Information Package (AIP)

- Check **licence**
- Check copyright and confidentiality clearance
- **Consistency** checks
- Selecting **preservation** and **dissemination** file formats
- Develop a **conversion** plan
- **Convert** the files
- **Validate** file conversion
- **Metadata** update
- Create and store **checksums** for the AIP
- Submit AIP for checking



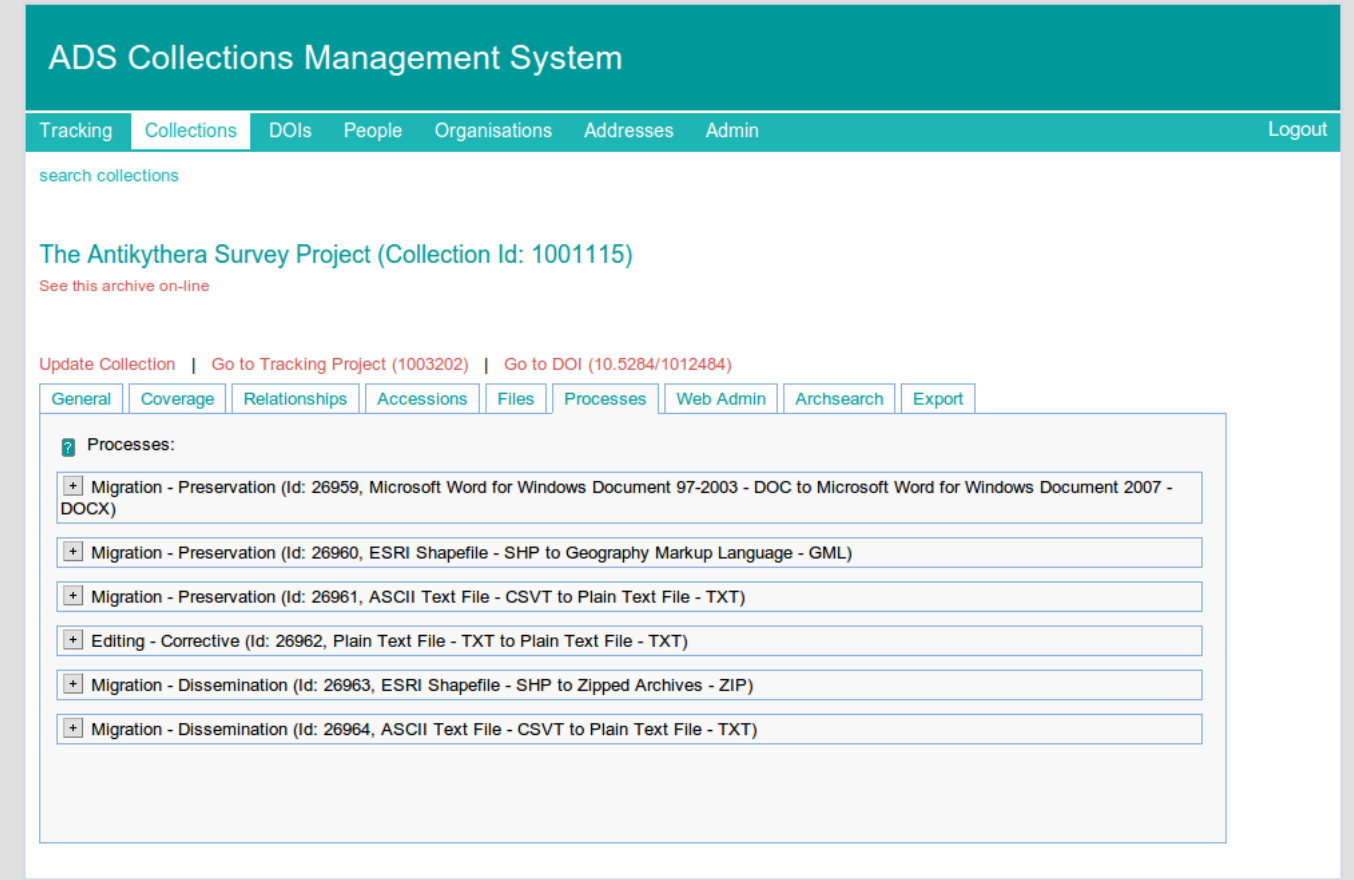
Information Packages

Delivery	Preservation	Presentation	Notes
Uncompressed Baseline TIFF v.6 **.tif**	Uncompressed Baseline TIFF v.6 **.tif**	Portable Network Graphics **.png** or Joint Photographic Expert Group **.jpg**	Any EXIF & IPTC metadata will also need preserving
Portable Network Graphics **.png**	Uncompressed Baseline TIFF v.6 **.tif**	Portable Network Graphics **.png**	""
Joint Photographic Expert Group **.jpg/ .jpeg**	Uncompressed Baseline TIFF v.6 **.tif**	Joint Photographic Expert Group **.jpg / .jpeg**	""
Graphics Interchange Format (Compuserve) **.gif**	Uncompressed Baseline TIFF v.6 **.tif**	Portable Network Graphics **.png**	""
Bit-Mapped Graphics Format (Microsoft) **.bmp**	Uncompressed Baseline TIFF v.6 **.tif**	Portable Network Graphics **.png**	""
PhotoCD **.pcd**	Uncompressed Baseline TIFF v.6 **.tif**	Portable Network Graphics **.png**	""
Photoshop (Adobe) **.psd**	Uncompressed Baseline TIFF v.6 **.tif**	Portable Network Graphics **.png**	""
CorelPaint **.cpt**	Uncompressed Baseline TIFF v.6 **.tif**	Portable Network Graphics **.png**	""
Adobe Digital Negative **.dng**	Adobe Digital Negative **.dng** and **.tif**	Adobe Digital Negative **.dng** and Joint Photographic Expert Group **.jpg**	""
JPEG2000 **.jp2 / .jpx**	Uncompressed Baseline TIFF v.6 **.tif**	JPEG2000 **.jp2 / .jpx**	""



Migration

Requires active and continuous management



Archiving

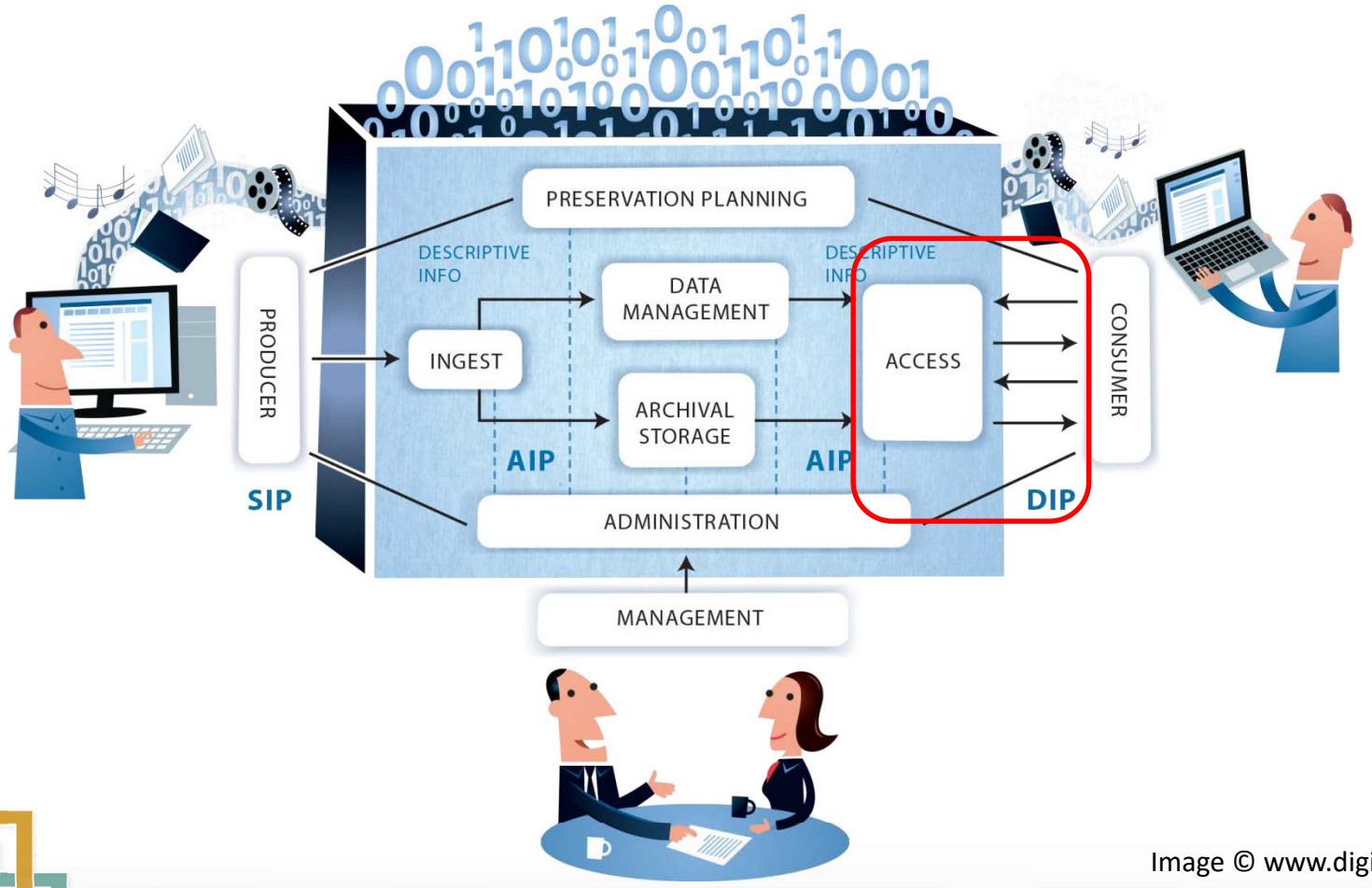


Image © www.digitalbevaring.dk

Dissemination Information Package (DIP)

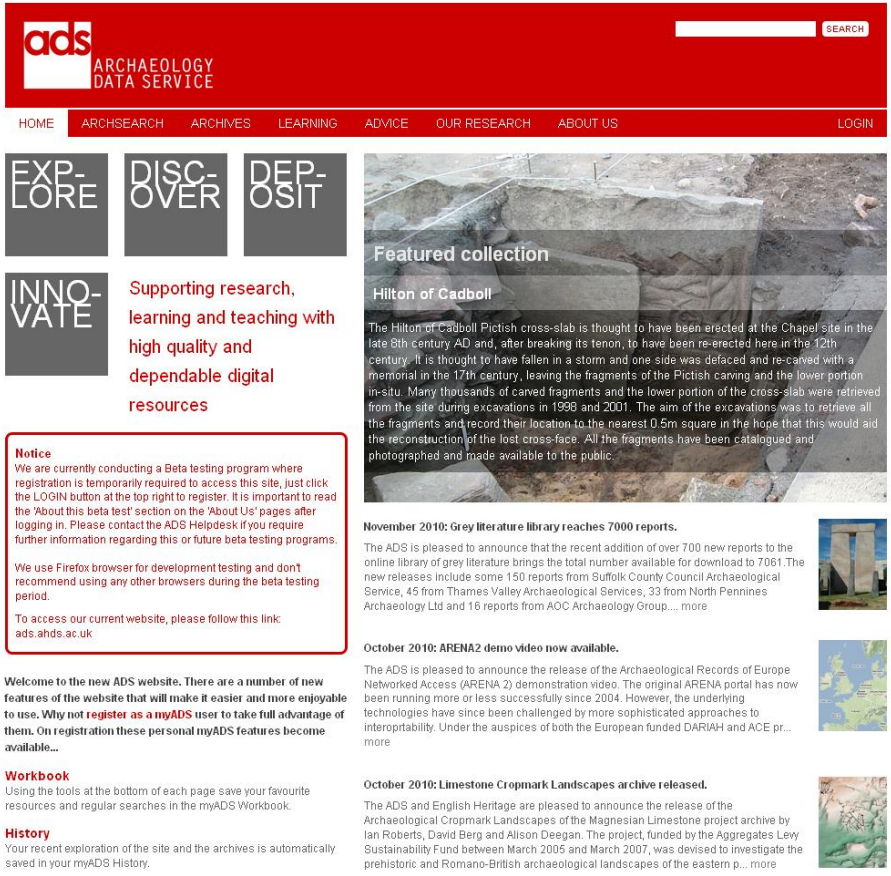
- **Convert** the files
- **Validate** file conversion
- Create web **interface**
- Allocate permanent urls / **DOIs**
- Pre-Release interface
- Make any depositor changes
- Release Archive
- Publicise Archive

Dissemination

How do ADS disseminate data?

All archives are **freely** available through the web interface.

http://archaeologydataservice.ac.uk/



Dissemination

ADS Resources

1.3 million metadata records

Archives

- 22 Journals and Series
- 35,000+ Grey Literature reports
- 900+ Project Archives
- Six specialist Bibliographies
- 19 Doctoral Theses

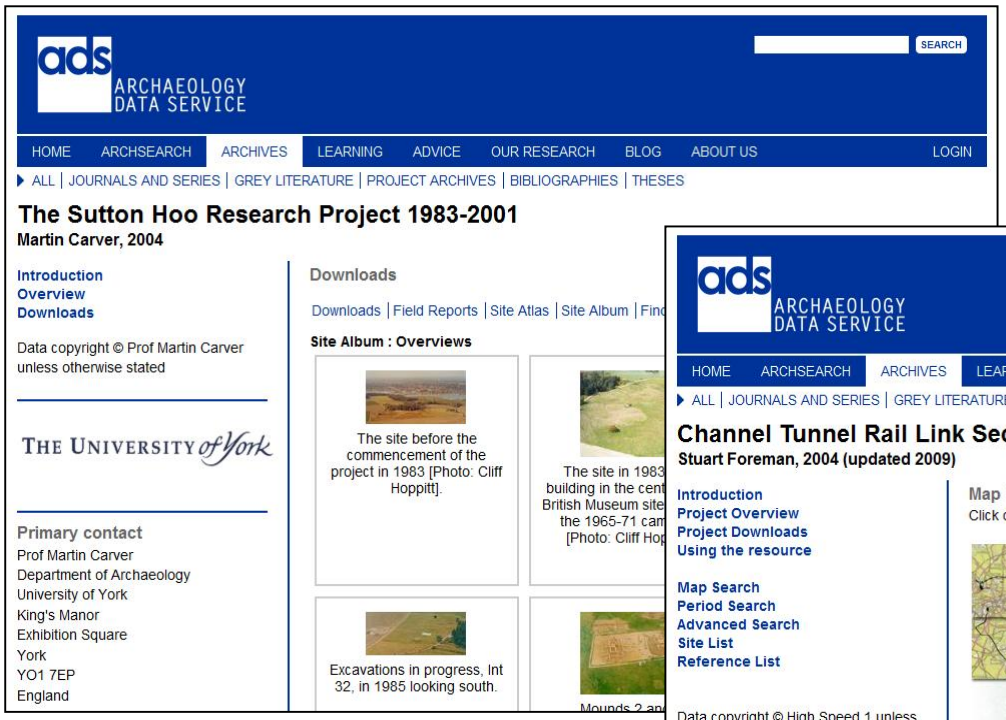
Specialised Websites

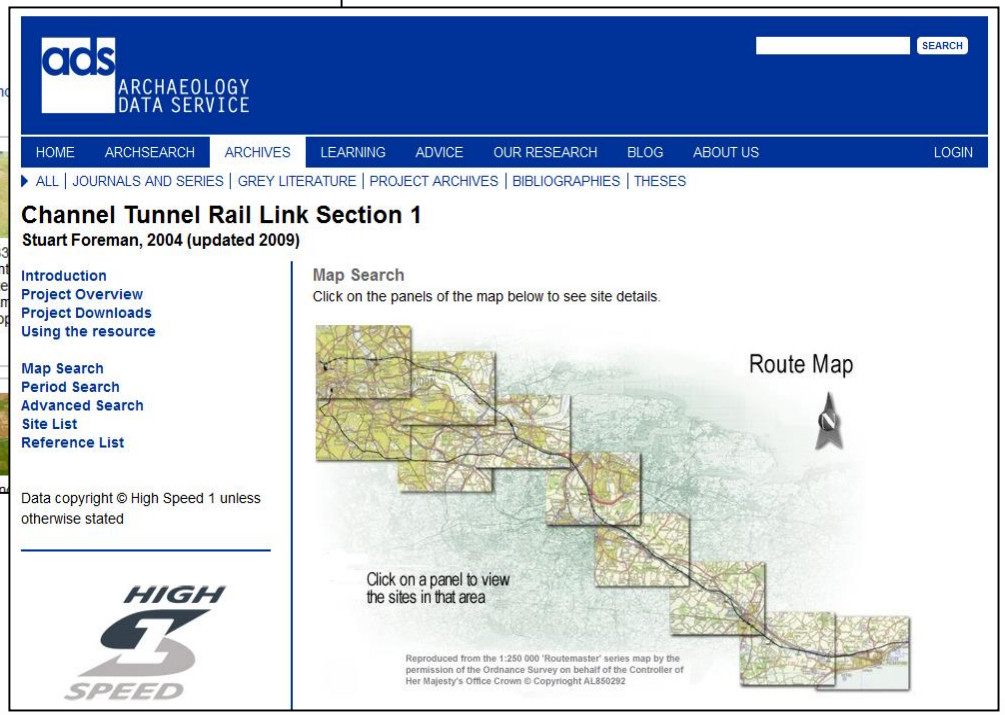
- England's Rock Art
- Image Bank

Dissemination

Project Archives

University-led research projects



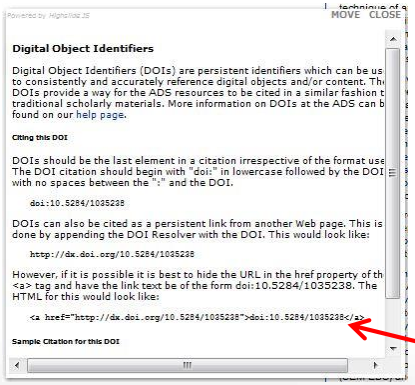


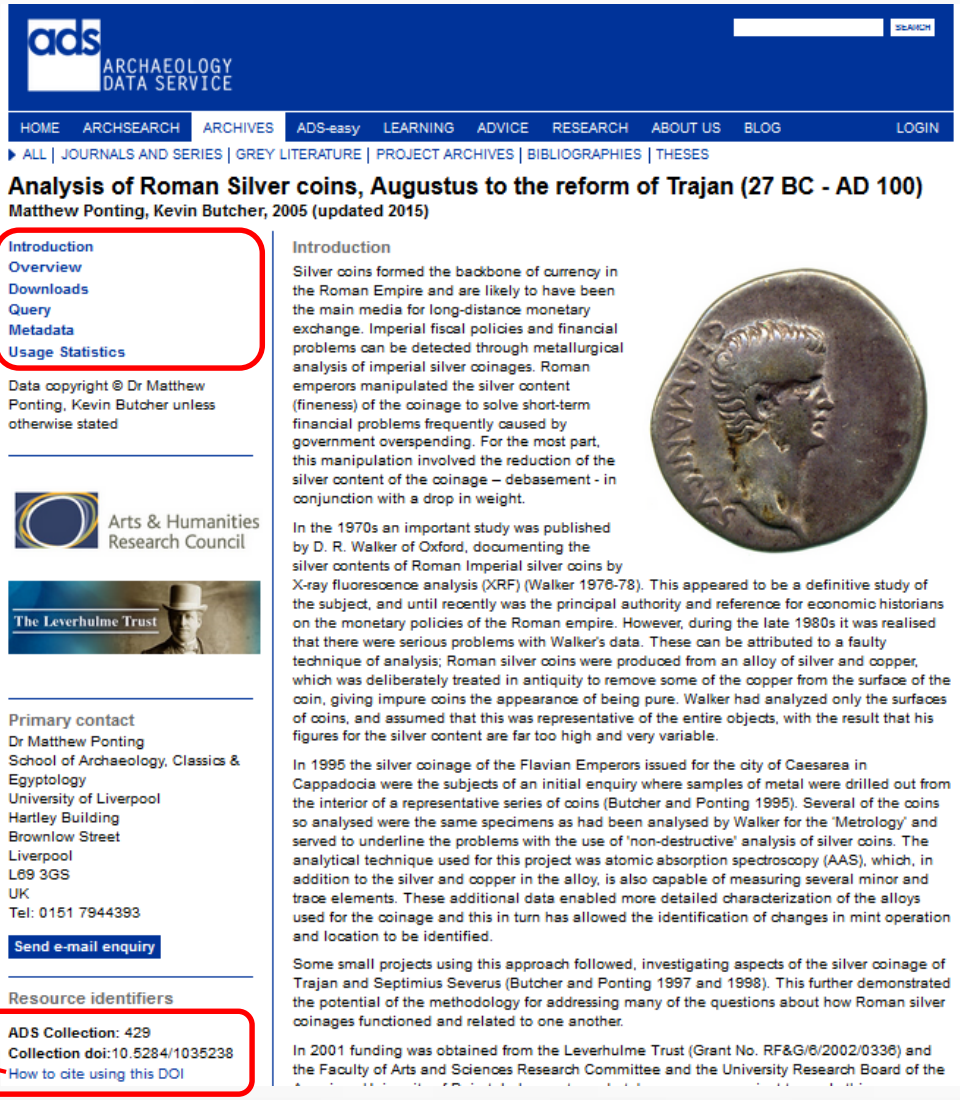
Major infrastructure programmes

Dissemination

Project Archives

- Introduction
- Overview
- Downloads and/or Search
- Metadata
- Usage Statistics
- Resource Identifiers





Acknowledgements

ARIADNE is a project funded by the European Commission under the Community's Seventh Framework Programme, contract no. FP7-INFRASTRUCTURES-2012-1-313193.

These training materials are based on teaching materials for research data management in archaeology created by Lindsay Lloyd Smith (2011) in the JISC-funded DataTrain project based at Cambridge University Library with contributions from the Archaeology Data Service.