Digital Preservation at the Archaeology Data Service

Dr Tim Evans
Deputy Director
Archaeology Data Service

Background

- Accredited digital archive for heritage data in the UK.
- Founded in 1996
- Based at the University of York
- Predominantly England (work with national agencies)
- Data from Industry and Higher Education

For more information
seehttps://archaeologydataservice.ac.uk/about/annualReports.xhtml

Background

- Archaeology generates a lot of data
- Excavation *is* destructive
- Preservation by Record is at the heart of most work
- Moral – and increasingly legal – obligation to ensure data is preserved

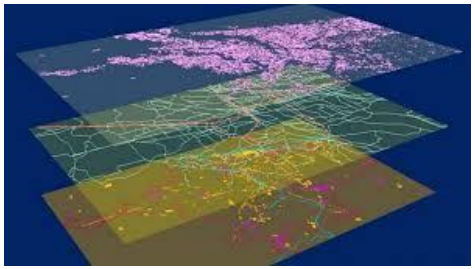
For more information
seehttps://archaeologydataservice.ac.uk/about/annualReports.xhtml

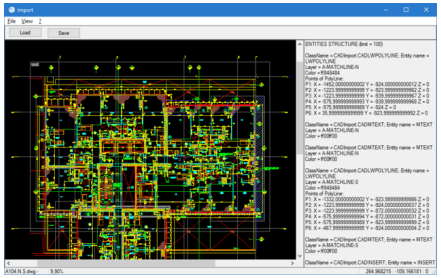
Archaeology can produce…

Data from scientific analysis

```
dvdrental=# select title, release_year, length, replacement_cost from film
dvdrental-#   where length > 120 and replacement_cost > 19.50
dvdrental-#   order by title desc;
         title          | release_year | length | replacement_cost
------------------------+--------------+--------+------------------
 West Lion              |         2006 |    159 |            29.99
 Virgin Daisy           |         2006 |    179 |            29.99
 Uncut Suicides         |         2006 |    172 |            29.99
 Tracy Cider            |         2006 |    142 |            29.99
 Song Hedwig            |         2006 |    163 |            29.99
 Slacker Liaisons       |         2006 |    179 |            29.99
 Sassy Packer           |         2006 |    154 |            29.99
 River Outlaw           |         2006 |    140 |            29.99
 Right Cranes           |         2006 |    153 |            29.99
 Quest Mussolini        |         2006 |    177 |            29.99
 Poseidon Forever       |         2006 |    159 |            29.99
 Loathing Legally       |         2006 |    140 |            29.99
 Lawless Vision         |         2006 |    181 |            29.99
 Jingle Sagebrush       |         2006 |    124 |            29.99
 Jericho Mulan          |         2006 |    171 |            29.99
 Japanese Run           |         2006 |    135 |            29.99
 Gilbert Pelican        |         2006 |    163 |            29.99
 Floats Garden          |         2006 |    146 |            29.99
 Fantasia Park          |         2006 |    131 |            29.99
 Extraordinary Conquerer|         2006 |    122 |            29.99
 Everyone Craft         |         2006 |    163 |            29.99
 Dirty Ace              |         2006 |    147 |            29.99
 Clyde Theory           |         2006 |    139 |            29.99
 Clockwork Paradise     |         2006 |    143 |            29.99
 Ballroom Mockingbird   |         2006 |    173 |            29.99
(15 rows)
```

Databases





CAD + GIS

Background

- ADS archive over 3 million digital objects…
- …In 299 unique formats
- ‘Bespoke’ systems stack not tied to any proprietary or open source preservation system

For more information
seehttps://archaeologydataservice.ac.uk/about/annualReports.xhtml

Background

- Entirely project funded.
- Maintain a core staff of 12.
- Of which 5 are directly/predominantly focussed on Digital Preservation.
- “Curatorial And Technical Staff” or CATs.

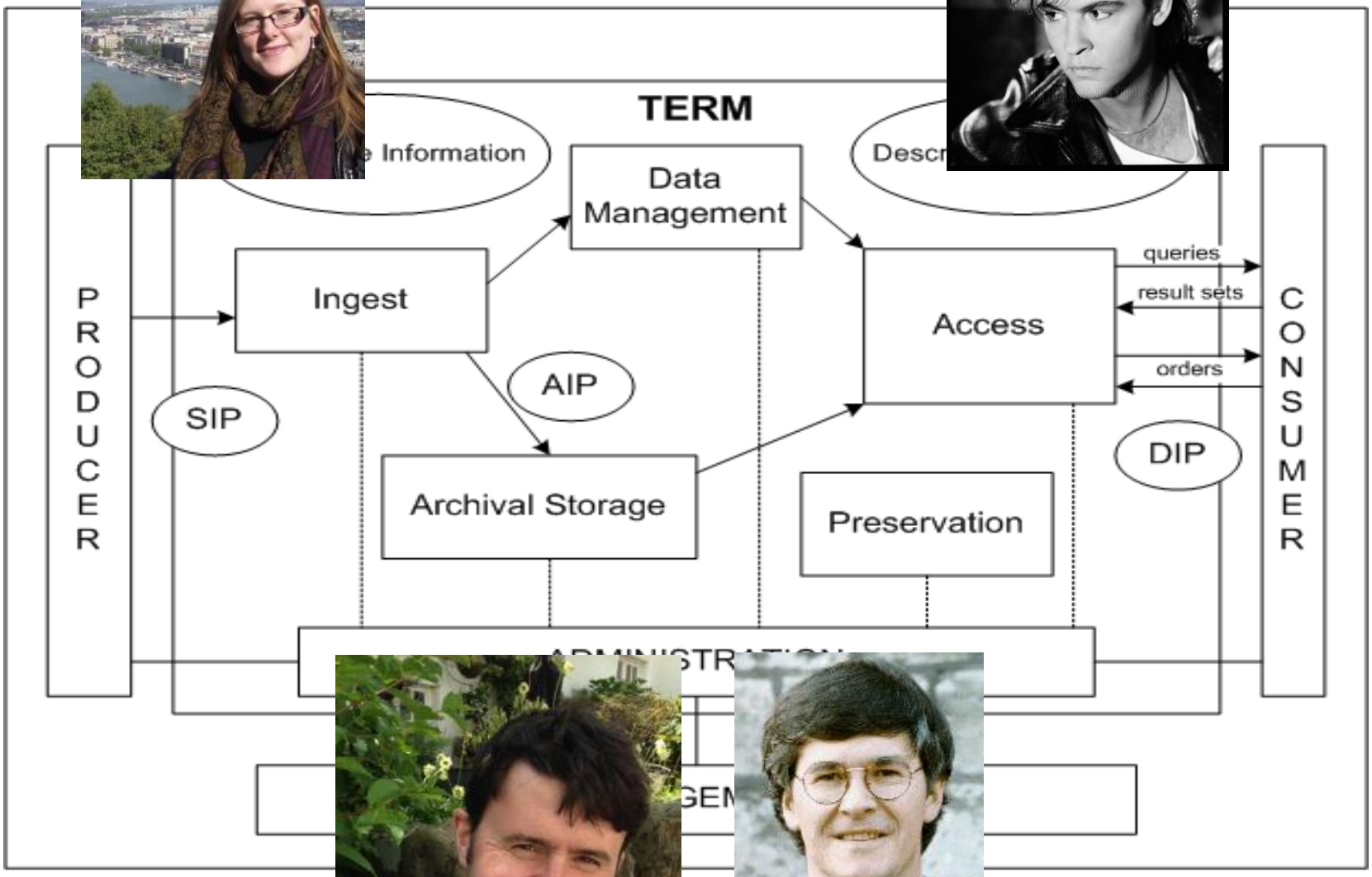


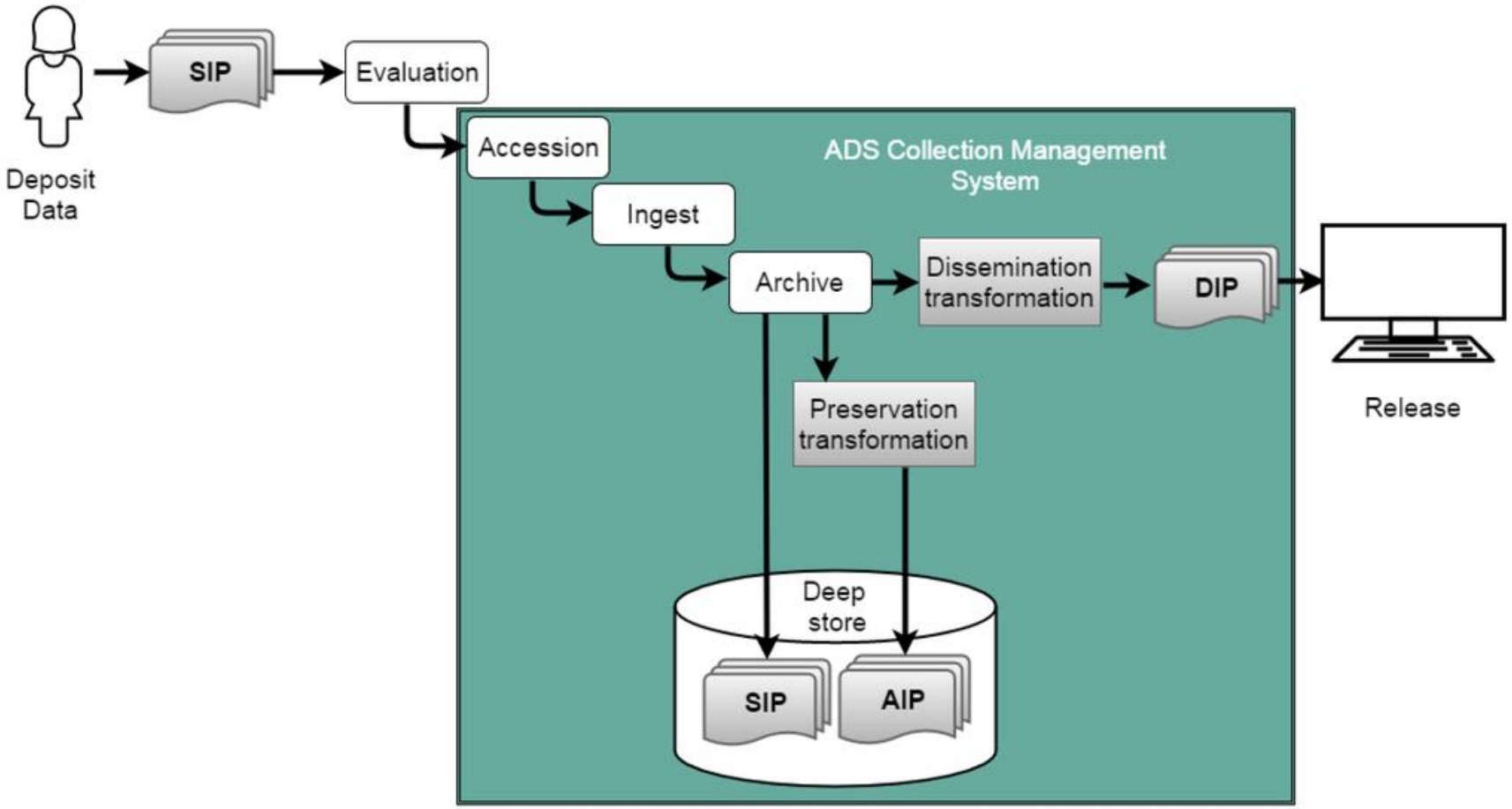
Fig. 1. Major function ode Committee for Space Data Systems (CCSDS), C r-Ar nodology Abstract Standard. (OAIS). Wh com e Data System Standards.

Preservation via migration

- Every file accessioned preserved as a standardised format e.g. all raster images as uncompressed TIFF, databases as delimited text…
- This is called “normalisation”
- Formats used based on a mixture of technical consideration (e.g. compression = data loss), but also judgements on longevity of format and ease of establishing future migrations.

ARCHAEOLOGY DATA SERVICE

Operational workflow



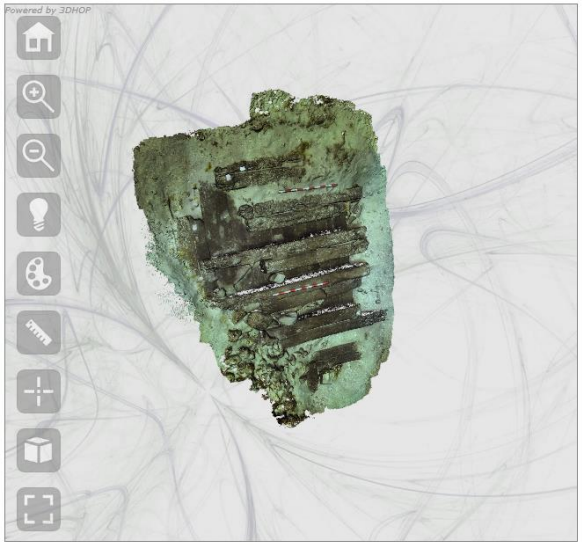
- Increasingly, our work is focussed on understanding what we have, where it is, actively managing risks to that content
  - Checksums: is it the same?
  - Obsolescence: do we need to migrate
  - = understanding what we've done and what we need to do

1,003,695	Original	kap020434.jpg	/ADS_preservation/arch-1591-1/original/3306/2013-08-05/jpg/...
1,003,696	Original	kap020435.JPG	/ADS_preservation/arch-1591-1/original/3306/2013-08-05/orig...
1,003,697	Dissemination	kap020435.JPG	/adsdata/arch-1591-1/dissemination/jpg/origimages/kap0204...
1,003,698	Dissemination	kap020435.jpg	/adsdata/arch-1591-1/dissemination/jpg/fullsized/kap020435....
1,003,699	Preservation	kap020435.tif	/ADS_preservation/arch-1591-1/preservation/tif/fullsized/kap0...
1,003,700	Preservation	kap020435.tif	/ADS_preservation/arch-1591-1/preservation/tif/origimages/ka...
1,003,701	Original	kap020435.jpg	/ADS_preservation/arch-1591-1/original/3306/2013-08-05/jpg/...
1,003,702	Original	kap020436.JPG	/ADS_preservation/arch-1591-1/original/3306/2013-08-05/orig...
1,003,703	Dissemination	kap020436.JPG	/adsdata/arch-1591-1/dissemination/jpg/origimages/kap0204...
1,003,704	Dissemination	kap020436.jpg	/adsdata/arch-1591-1/dissemination/jpg/fullsized/kap020436....
1,003,705	Preservation	kap020436.tif	/ADS_preservation/arch-1591-1/preservation/tif/fullsized/kap0...
1,003,706	Preservation	kap020436.tif	/ADS_preservation/arch-1591-1/preservation/tif/origimages/ka...
1,003,707	Original	kap020436.jpg	/ADS_preservation/arch-1591-1/original/3306/2013-08-05/jpg/...
1,003,708	Original	kap020437.JPG	/ADS_preservation/arch-1591-1/original/3306/2013-08-05/orig...
1,003,709	Dissemination	kap020437.JPG	/adsdata/arch-1591-1/dissemination/jpg/origimages/kap0204...
1,003,710	Dissemination	kap020437.jpg	/adsdata/arch-1591-1/dissemination/jpg/fullsized/kap020437....
1,003,711	Preservation	kap020437.tif	/ADS_preservation/arch-1591-1/preservation/tif/fullsized/kap0...
1,003,712	Preservation	kap020437.tif	/ADS_preservation/arch-1591-1/preservation/tif/origimages/ka...
1,003,713	Original	kap020437.jpg	/ADS_preservation/arch-1591-1/original/3306/2013-08-05/jpg/...
1,003,714	Original	kap020438.JPG	/ADS_preservation/arch-1591-1/original/3306/2013-08-05/orig...
1,003,715	Dissemination	kap020438.JPG	/adsdata/arch-1591-1/dissemination/jpg/origimages/kap0204...
1,003,716	Dissemination	kap020438.jpg	/adsdata/arch-1591-1/dissemination/jpg/fullsized/kap020438....
1,003,717	Preservation	kap020438.tif	/ADS_preservation/arch-1591-1/preservation/tif/fullsized/kap0...
1,003,718	Preservation	kap020438.tif	/ADS_preservation/arch-1591-1/preservation/tif/origimages/ka...
1,003,719	Original	kap020438.jpg	/ADS_preservation/arch-1591-1/original/3306/2013-08-05/jpg/...
1,003,720	Original	kap020439.JPG	/ADS_preservation/arch-1591-1/original/3306/2013-08-05/orig...
1,003,721	Dissemination	kap020439.JPG	/adsdata/arch-1591-1/dissemination/jpg/origimages/kap0204...
1,003,722	Dissemination	kap020439.jpg	/adsdata/arch-1591-1/dissemination/jpg/fullsized/kap020439....
1,003,723	Preservation	kap020439.tif	/ADS_preservation/arch-1591-1/preservation/tif/fullsized/kap0...
1,003,724	Preservation	kap020439.tif	/ADS_preservation/arch-1591-1/preservation/tif/origimages/ka...
1,003,725	Original	kap020439.jpg	/ADS_preservation/arch-1591-1/original/3306/2013-08-05/jpg/...
1,003,726	Original	p0000020.jpg	/ADS_preservation/arch-1591-1/original/3306/2013-08-05/jpg/...
1,003,727	Dissemination	p0000020.jpg	/adsdata/arch-1591-1/dissemination/jpg/fullsized/p0000020.jpg
1,003,728	Dissemination	p0000020.jpg	/adsdata/arch-1591-1/dissemination/jpg/origimages/p000002...
1,003,729	Preservation	p0000020.tif	/ADS_preservation/arch-1591-1/preservation/tif/fullsized/p000...
1,003,730	Preservation	p0000020.tif	/ADS_preservation/arch-1591-1/preservation/tif/origimages/p0...
1,003,731	Original	p0000020.tif	/ADS_preservation/arch-1591-1/original/3306/2013-08-05/orig...

FILE_TYPE	FILE_TYPE_VER	PRONOM_ID
JPEG File Interchange Format	1.02	fmt/44
Exchangeable Image File Format (Compressed)	2.1	x-fmt/390
Exchangeable Image File Format (Compressed)	2.1	x-fmt/390
JPEG File Interchange Format	1.02	fmt/44
Tagged Image File Format	<null>	fmt/353
Tagged Image File Format	<null>	fmt/353
JPEG File Interchange Format	1.02	fmt/44
Exchangeable Image File Format (Compressed)	2.1	x-fmt/390
Exchangeable Image File Format (Compressed)	2.1	x-fmt/390
JPEG File Interchange Format	1.02	fmt/44
Tagged Image File Format	<null>	fmt/353
Tagged Image File Format	<null>	fmt/353
JPEG File Interchange Format	1.02	fmt/44
Exchangeable Image File Format (Compressed)	2.1	x-fmt/390
Exchangeable Image File Format (Compressed)	2.1	x-fmt/390
JPEG File Interchange Format	1.02	fmt/44
Tagged Image File Format	<null>	fmt/353
Tagged Image File Format	<null>	fmt/353
JPEG File Interchange Format	1.02	fmt/44
Exchangeable Image File Format (Compressed)	2.1	x-fmt/390
Exchangeable Image File Format (Compressed)	2.1	x-fmt/390
JPEG File Interchange Format	1.02	fmt/44
Tagged Image File Format	<null>	fmt/353
Tagged Image File Format	<null>	fmt/353
JPEG File Interchange Format	1.02	fmt/44
Exchangeable Image File Format (Compressed)	2.1	x-fmt/390
Exchangeable Image File Format (Compressed)	2.1	x-fmt/390
JPEG File Interchange Format	1.02	fmt/44
Tagged Image File Format	<null>	fmt/353
Tagged Image File Format	<null>	fmt/353
JPEG File Interchange Format	1.02	fmt/44
Exchangeable Image File Format (Compressed)	2.1	x-fmt/390

PREMIS

- ADS are increasingly using the PREservation Metadata: Implementation Strategies
  - Objects: levels of information that have relationships to one another
  - Preservation Events: The things we do
  - Actors (us, the tools we use)



AIP DIP SIP

AIP DIP SIP

AIP DIP SIP

Metadata

OBJ + MTL +Textures

Source images (lots of raw)

REP_ID	PURPOSE	FILE_NAME	FILE_SIZE	CHECKSUM	OBJECT_ID
3,935,118	Original	Bayo_04_obj.jpg	2,646,743	d82278660faeb15979ce47a2408033fd	1,640,995
3,935,119	Original	Bayo_04_obj.mtl	591	85d7f17485d895259dc47f2b8c7770bd	1,640,995
3,935,120	Original	Bayo_04_obj.obj	62,254,861	52a77f1a869f06d225930d380a92df09	1,640,995
3,935,121	Original	Bayo_04_obj1.jpg	2,286,790	e81f43d7b8b7977ef8d0595617aae964	1,640,995
3,935,122	Original	Bayo_04_obj2.jpg	1,851,007	182af6c0d14222d8cfa9f9cf12644fdc	1,640,995
3,935,123	Original	Bayo_04_obj3.jpg	2,237,707	b92b8b585a1306524c53fc2a42f53a48	1,640,995

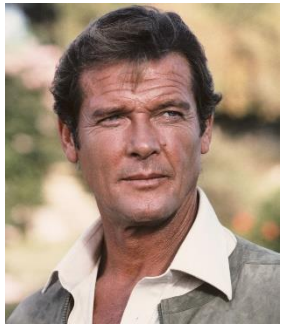
process.validation

process.virus check

Is Documented In

process.normalization

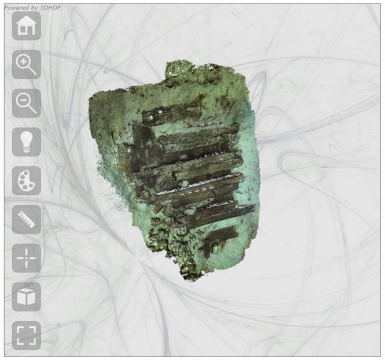
process.migration



agent.actor.r_moore

agent.software.xnview

Is Source of



PREMIS

- Ensuring the ‘complexity’ of objects and their metadata is retained
- Supporting ongoing repository management activities
- Retaining the history of digital object

Internal + External Assessment

- DPC RAM
- CoreTrustSeal (CTS)
- Encourage reflection, and assessment of where we can improve
- Often feed into developing Policy and ensuring workflows are documented

DigitalPreservationCoalition

Collaboration (1)

- With the DPC – training for staff, being part of a community and sharing best practice
- National Heritage agencies – giving the same message!

DigitalPreservationCoalition

ads ARCHAEOLOGY DATA SERVICE

100 CBHC RCAHMW

HISTORIC ENVIRONMENT SCOTLAND | ÀRAINNEACHD EACHDRAIDHEIL ALBA

Historic England

Collaboration (2)

- Internationally!
- Sharing metadata: Ariadne https://t.co/Ri32qb8sjq?amp=1

- Working to establish best practice: SEADDA https://www.seadda.eu/

SEADDA
Saving European Archaeology from the Digital Dark Age

ARIADNE plus

ads ARCHAEOLOGY DATA SERVICE

Thanks!

tim.evans@york.ac.uk

http://archaeologydataservice.ac.uk/blog/

https://twitter.com/ADS_Update

help@archaeologydataservice.ac.uk