



Archaeology  
Data Service

# Quality Assurance

*Kieron Niven, Digital Archivist: Data Standards*  
*6th December 2022*

# Three Broad Types of Quality Assurance:

- Data being submitted is of good quality
- Work undertaken by ADS during archiving of consistent standard
- Systems etc. maintain that quality

## Wider framework of quality assessment

Depending on origins of data:

- from research projects has been through some kind of approval system via funding: DMP and RC approval
- OASIS validation at various levels for commercial projects.
- Commercial units working within an established framework (ClfA, archaeologist requirements)

## Wider framework of quality assessment

ADS guidance for submitting data:

- Guidance on the Selection of Material for Deposit and Archive
  - <https://archaeologydataservice.ac.uk/advice/selectionGuidance.xhtml>
- **ADS Collections Policy: 2.3. Criteria for Evaluating Electronic Datasets.**
- **Assessment and Appraisal Policy**

## Quality Assessment during submission

Checks also carried out more generally when data is being submitted to ADS:

- **ADS-easy** has limits on what can be uploaded (file types) and requirements for metadata
- **OASIS** has limits re. files and metadata (OASIS form)
- **'Traditional' deposition** – assessed by CDM and archivists

# Once data is 'allowed' to enter the archive:

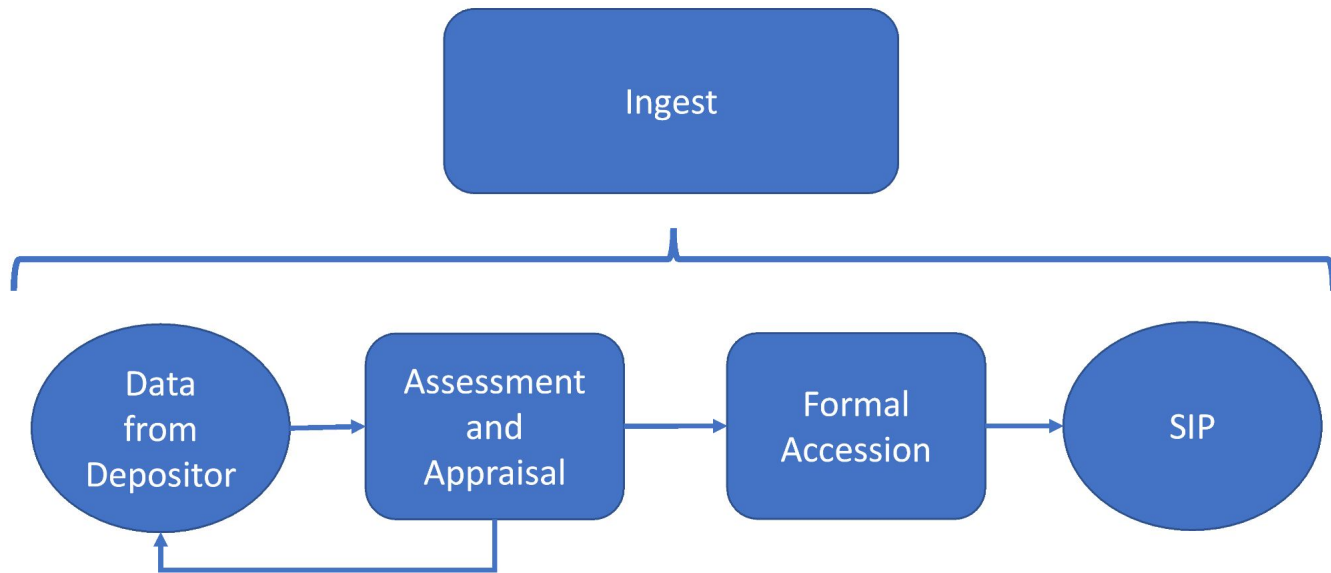
- Two main stages of quality assurance (start and end):
  - Ingest
  - Prior to Release

## Quality Assessment during Ingest

- Covered in ***ADS Assessment and Appraisal Policy***
- Covered in detail in the ***ADS Ingest Manual***
- Ingest process creates the Accession – in OAIS terms: assesses the SIP

## Quality Assessment during Ingest

“...to ensure that data is fit to be accessioned, and that any major issues are identified and wherever possible resolved before the creation of the SIP.”



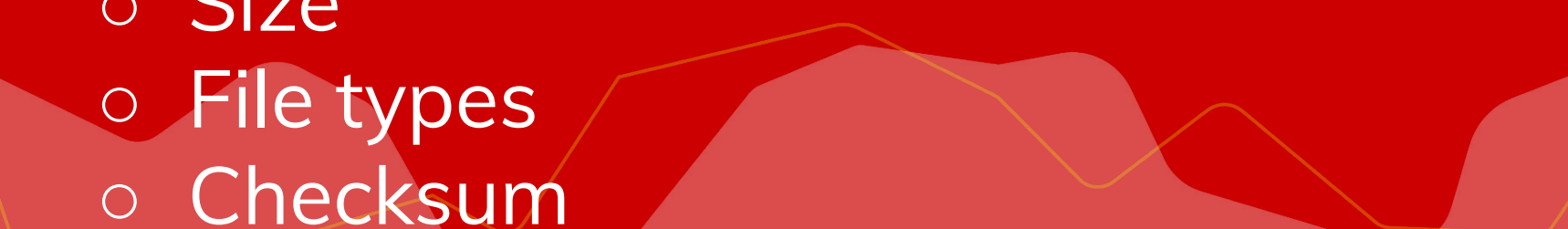


## Quality Assessment during Ingest

Types of checks:

- Correct formats, no virus etc. can be opened.
- Correct metadata – collection and object, and technical.
- No sensitive data or GDPR issues.
- Data is complete and appropriate.

# Once Ingest is complete:

- **Depositors are issued with a receipt**
    - File names
    - Size
    - File types
    - Checksum
- 
- A decorative wavy line in a lighter shade of red, spanning the width of the slide and positioned below the list of receipt items.

# Example Receipt:

	A	B	C	D	E	F	G
1	Deposit Receipt for ADS Collection 1004876						
2	19 Files						
3	Accession Id: 1011917						
4	File & Path	File Size	File Type	MD5 Checksum			
5	OXEBCH19_site_location_plan.pdf	724705	Acrobat PDF/A - Portable Document Format	b7e440197a9726b191cb9a9e7464ef10			
6	OXEBCH19_0003.jpg	2954903	Exchangeable Image File Format (Compressed)	81535951e602a012d415e239e893ee59			
7	OXEBCH19_0004.jpg	2745411	Exchangeable Image File Format (Compressed)	f61c74b32eb75bc929afb4fa3ed22793			
8	OXEBCH19_0006.jpg	2789659	Exchangeable Image File Format (Compressed)	92f3a9559e00155aa2fb85628280e8ab			
9	OXEBCH19_0008.jpg	2867914	Exchangeable Image File Format (Compressed)	edd916b653272c46a16d25576ce65171			
10	OXEBCH19_0012.jpg	2549533	Exchangeable Image File Format (Compressed)	1372dfe2f12e4eb61b475d1b3b9964bd			
11	OXEBCH19_0014.jpg	2629203	Exchangeable Image File Format (Compressed)	f7eeecdd78d68e53639b7424895e9c5e			
12	OXEBCH19_0016.jpg	2536946	Exchangeable Image File Format (Compressed)	02b2f216c85c4442262d4e100583cfd9			
13	OXEBCH19_0017.jpg	2499288	Exchangeable Image File Format (Compressed)	a4810539499742dd4e664e40566e8c5e			
14	OXEBCH19_0_WSI.pdf	2112613	Acrobat PDF/A - Portable Document Format	cf58c3fa4655456d4d894ac78669549d			
15	OXEBCH19_B_All_Records.pdf	1124441	Acrobat PDF/A - Portable Document Format	e2231f4de874adf6f2a977a5eeaf24fc			
16	OXEBCH19_final_drawing.dwg	105959	AutoCAD Drawing	c0ff9c2a771ad0d5f3a73e960c7b75d8			
17	OXEBCH19_section_points.xls	20480	Microsoft Excel 97 Workbook (xls)	69e6e24b10c8d4e396f1ac32141cc684			
18	OXEBCH19_trench_points.xls	20992	Microsoft Excel 97 Workbook (xls)	008a493c5dd490159bb4b4db4c0d1d75			
19	documentation/OA_CAD_layers_spreadsheet.xls	38912	Microsoft Excel 97 Workbook (xls)	300eacf757c32cf27e8f805ed44ba471			
20	documentation/OXEBCH19_ADS_raster_metadata.xlsx	13088	Microsoft Excel for Windows	1ca6d9d5d4819f85f44a9796a22db99b			
21	documentation/OXEBCH19_section_points_spreadsheet_metadata.xlsx	184840	Microsoft Excel for Windows	041c611967c8dd79f4a38e8ddae506c5			
22	documentation/OXEBCH19_trench_points_spreadsheet_metadata.xlsx	184951	Microsoft Excel for Windows	5213d2cefbb71d13707fa6e9fc404265			
23	documentation/OXEBCH19_vector_metadata_spreadsheet.xlsx	189131	Microsoft Excel for Windows	d5df27b1ac61bbe301f1336c298f7342			

# After Ingest – Processing the Data

Ingest allows us to proceed with:

- Normalisation
- Loading metadata

...without worry of further changes.

## Quality Assessment during Data Normalisation and Migration

- Usually carried out on a *Data Type* basis
- Data Procedure guides for each data type: specify formats, etc.
- Also specify *Significant Properties* and elements to be aware of.
  - Key to check that these aspects are retained during migration.

# Example of Sig. Properties:

1	Surface form	Edge	Edge form	Edge subtype	Abrasion	Type	Class
2	P Flat	B1a	rectangular	sharp arrises	O none	Surfaces FC1	One flat surface
3	C Convex	B1b	rectangular	rounded arrises	L low	Surfaces FC2	One convex surface
4	Cv Concave	B1c	rectangular	top arris sharp, base rounded	L-M	Surfaces FC3	One concave surface
5	u undulating	B1d	rectangular	base arris sharp, top rounded	M moderate	Surfaces FC4	Two flat surfaces at right angles
6	w worn	B2a	concave	acute arrises	M-H	Surfaces FC5	One flat & 1 curved surfaces at right angles
7	PC plano-convex	B2b	concave	rounded arrises	H high	Surfaces FC6	3 flat surfaces at right angles forming a square corner
8	PCv plano-concave	B2c	concave	slight bevel to top	vH very high fragile, fragmented	Surfaces FC7	2 flat // shaped surfaces [= Morris's Slabs (Briq)]
9	S irregular surface	B2d	concave	slight bevel to base	F	Surfaces FC7a	2 flat // shaped surfaces [= Morris's Slabs (Briq)] with edge
10	Surface finish	B3a	convex	sharp arrises	~ no record	Surfaces FC8	largely amorphous, some hint of utilisation or moulded surface
11	1 Smooth, very even and regular	B3b	convex	rounded profile, rounded arrises		Surfaces FC9	indeterminate, amorphous
12	2 even, fairly regular	B3c	convex	sharp/fairly sharp upper, well rounded/bevelled lower angle		Lining LN	Lining
13	2b Wiped	B3d	convex	sharp/fairly sharp lower, well rounded/bevelled upper angle		Lining LN1	lining of clay structure
14	3a Finger smoothed - grooves	B3e	convex	bulbous on both sides		Lining LN2	lining of turf structure
15	3b Finger smoothed - dimples: fir	B3f	convex	bulbous on top side only		Lining LN3	Lining of stone structure
16	3c Pie-crust - finger-thumb press	B3g	convex	bulbous on lower side only		Lining FU LN	Furnace lining / wall surface (surface vitrifié/scorifié)
17	4 Roughly moulded, irregular finish	B4a	bevelled	sharp arrises, undercuts		Structural Str1	(flat) exterior surface with interwoven wattles on interior
18	5 impressed bonding surface with wattle impressions, wattle framework	B4b	bevelled	sharp arrises, base projects		Structural Str2	(flat) exterior surface with // wattles on interior
19	5c	B4c	chevron	sharp arrises		Structural Str3	(flat) exterior surface with wattles of uncertain construction form on interior
20	5st st-stems/straw	B4d	bevelled	rounded top arris		Structural Str4	(flat) exterior surface with large timber impressions on interior
21	5r r-bedrock/natural	B5a	acute	lentoid, flat base, curved top		Structural Str5	2 surfaces, one smooth, one rough daubed/plastered bedding/bonding face (eg final render)
							Flat even surface on one side, rough surface sheared at interface or with stone or other impressions on opposite

# Quality Assessment during Data Normalisation and Migration

- Running DROID
  - File identification. Can highlight non-standard files etc.
  - Creates fixity value for each file, entry in OMS
  - Rerunning DROID flags a changed fixity value

## Quality Assessment during Data Normalisation and Migration

- Creating the catalogue record (CMS metadata)
- Elements of quality assurance here:
  - Using controlled vocabs.
  - Verifying locations.
  - Coordinates.



## Quality Assessment: Finalising and Releasing the archive

- Internal check by all staff when archiving work is finished (just web pages).
- Depositor is asked to 'sign off' work in many cases.
- Most archives are allocated to another members of the CATS to 'AIP check'.

# AIP Checking Tool

- Functions as a validator across the database and file store.
- During the archiving process, the archivist can check most elements using the tool.

## File Checks

Filenames and extensions	Result
Spaces	0
Spaces/characters in FILE_PATH (i.e. directories if not picked up in file name checks)	0
Filenames: other characters (see Rep.Ops)	0
Full stops - files (excludes shp.xml)	0
Full stops - directories	0
Uppercase file extensions	0

Admin files	Result (Admin Directory)	Result (CMS Upload)
Licence files	0	1
Deposit receipt	0	0
Email files	0	0

**Note:** Files uploaded to the CMS without a 'note' will not appear in the *Result Upload* check.

Accession IDs	Result
CMS:	12082
Filestore (directory names):	

## OMS Checks

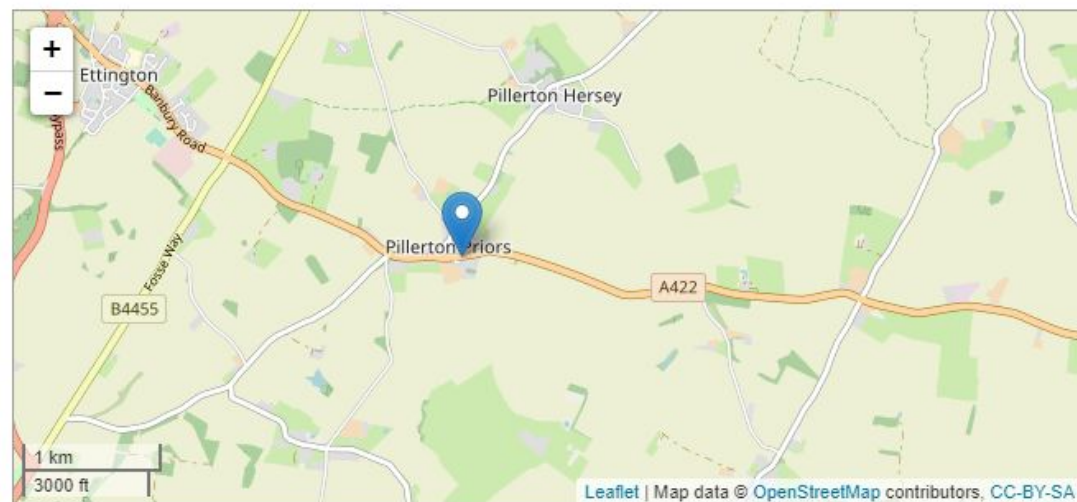
Unmatched Objects	Result
No. of unmatched objects	0
Single-file objects (ignores Admin data type)	0
Check that these files have the correct parent-child relationships.	0
No. of unrelated documentation files (object ID)	0
Files with no related documentation objects [file object_id]	0
Undefined PREMIS REL TYPE (parent_id:child_id)	0
Groups	Result
No. of Groups	No groups defined

## CMS Checks

General	Result
ACDM / AO-CAT?	MISSING
Coverage	Result
TGN	No TGN
Website top level	MISSING
Place	MISSING
LC Subject Heading	MISSING

## Bounding Box and Points

Bounding Box		
None defined		
Points		
COORD TYPE	EASTING/LONG	NORTHING/LAT
LL	-1.57224276	52.12588944



# Finishing the Process

Once signed off:

- **a *manifest* file is created and stored with the data.**
  - Provides a fixed, 'isolated' record of the data to check against.
- **Data is also sent to AWS storage**
- **All staff notified of release.**



Archaeology  
Data Service

# Quality Assurance

*kieron.niven@york.ac.uk*



## Archaeology Data Service

Department of Archaeology

University of York

The King's Manor

Exhibition Square

York, YO1 7EP



[www.archaeologydataservice.ac.uk](http://www.archaeologydataservice.ac.uk)



[help@archaeologydataservice.ac.uk](mailto:help@archaeologydataservice.ac.uk)