25 years of trends in digital data deposition at the ADS

CAA 2021 Digital Crossroads, 16/06/2021

Teagan Zoldoske
Digital Archives Assistant, ADS
teagan.zoldoske@york.ac.uk

Olivia Foster & Kieron Niven

ARCHAEOLOGY DATA SERVICE

UNIVERSITY *of York*

CORE TRUST SEAL

THE ARCHAEOLOGY DATA SERVICE

- Based at the University of York
- 1996 - present
- CoreTrustSeal Certified
- Provide Open Access data:
  - 1.4m metadata records of UK archaeology
  - \> 62,000 UK reports
  - \> 2000 international project archives
  - \>22Tb of data
  - \>2,000,000 files

Limitations of the data the holds

ADS and the data it holds

The ADS holds ~50% of all archaeological grey literature produced in England since 1990 and in some regions it is currently receiving over 90%.
- Between 2013-2018, c.36000 grey literature created within OASIS
- Only 1% of that had a corresponding digital archive

Digital archives may not have the data in the original form
- Appraisal process prior to being deposited
- ADS accepts only certain file formats
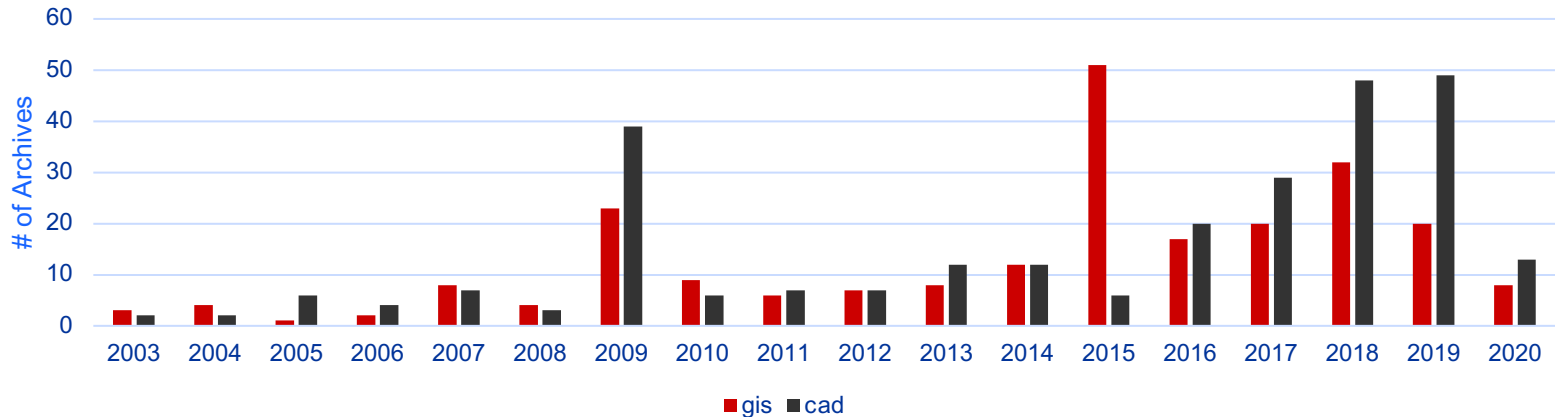- GIS or CAD may be deposited as static raster maps or site plans

Most archives are from the commercial archaeology sector

Looking at the CAD and GIS data

The GIS and the CAD data, trends

The ADS holds over 600 collections with CAD and/or GIS data
- Over 8,500 individual files since 1997
- 15% more CAD data than GIS
- 2020 saw a decrease in GIS and CAD files being created by 60-70%

Why is there less GIS data?

The past five years have seen steady increases in new CAD data created, unlike GIS

GIS vs CAD
- GIS has more open source options than CAD
- GIS can store more information then CAD
- CAD can be easier to show 2D drawings, i.e., section drawings
- Different requirements for archival
- Hardware/software costs
- Staff training and implementation costs

Archiving the data & looking forward

The ADS limits the data and file types it accepts
- Help ensure long term accessibility
- Stability and industry standards
- Proprietary and open software file types

Mass migration of CAD data in 2013 and more expected

GIS will need to be watched

Both CAD and GIS data are better then pdfs of the same information but England currently prefers CAD and will likely continue to use both equally for some time to come

Archiving advantages and disadvantages

CAD
file format

Preferred File Format	Accepted File Format
AutoCAD (2018 or later) .dwg	**AutoCAD (2017 or earlier)** .dwg
Scalable Vector Graphics .svg	**Drawing Interchange Format** .dxf

Any supporting documentation, i.e., metadata, data tables, codes, abbreviations or terminologies

C<small>AD</small>

ADS vector metadata	CSV	3 Kb

60510 35 West St Test Pit Evaluation, Local Grid survey AutoCAD drawing. (60510_Wilton_TPEval_Local_Grid)	DWG	29 Kb
	DXF	156 Kb
	PDF	14 Kb
60511- 60513 West St Wilton, Excavation and Evaluation survey AutoCAD drawing. (60511-13_Wilton_survey)	DWG	33 Kb
	DXF	169 Kb
	PDF	6 Kb
60515- 60516 West St Wilton, Excavation Watching Brief survey AutoCAD drawing. (60515-16_Phased_postex_plan)	DWG	212 Kb
	DXF	1 Mb
	PDF	567 Kb
60517 35 West St / Wilton Auto's (41-43 West St) Phases of archaeological work concordance drawing (60517_Phases_of_work_concordance)	DWG	44 Kb
	DXF	221 Kb
	PDF	166 Kb
60515 Excavation, Wilton Auto's (41-43 West St), phased site survey drawing (60515_Wilton_phased_survey)	DWG	189 Kb
	DXF	965 Kb
	PDF	53 Kb

G_{IS} file format

Data Type	Preferred File Format	Accepted File Format
GIS (General)	**ESRI Shapefile** .shp (+ .shx + .dbf and other associated files) **Geography Markup Language** .gml	**ArcInfo Interchange** .e00 **MapInfo Interchange Format** .mif + .mid
GIS (Geodatabases)		**Delimited text and ESRI Shapefile** .csv + .shp **GeoJSON** .geojson
GIS (Raster)	**Geo-referenced TIF Image** .tif (+ world file: .tfw) or GeoTIFF	**ERDAS Imagine files** .img (+ .rrd, aux.xml, img.xml) **ESRI GRID ascii** .asc/.grd **ESRI GRID binary** .adf **JPG World** .jpg + .jpw (.rrd, .aux, .xml) **Keyhole Markup Language** .kml **PNG World** .png + .pgw (.rrd,.aux,.xml)

Any supporting documentation, i.e., metadata, data tables, codes, abbreviations or terminologies
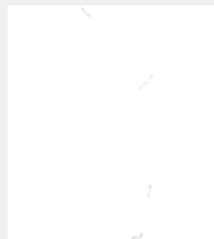
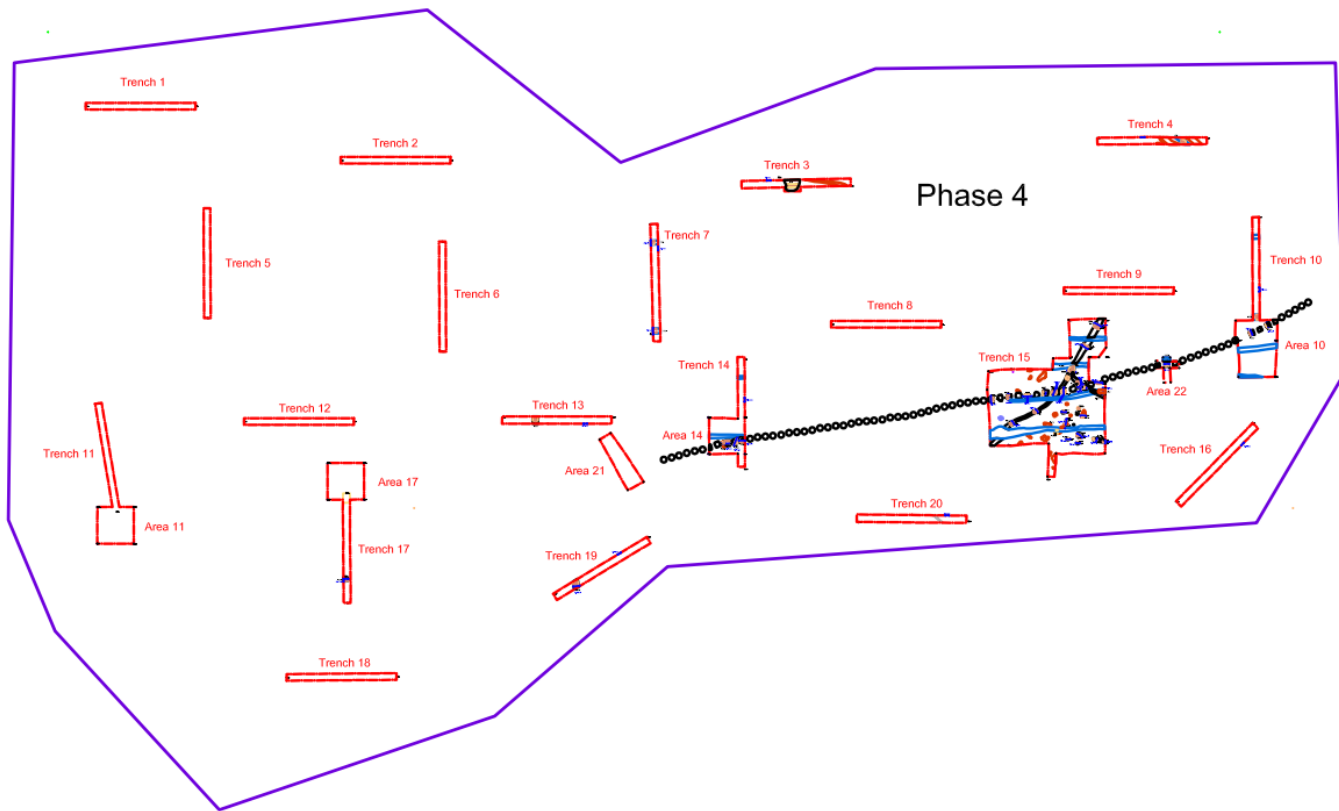| GIS metadata		CSV	1 Kb
PREVIEW	Ends of evaluation trenches, as points	ZIP	22 Kb
PREVIEW	Sampled archaeological features	ZIP	8 Kb
PREVIEW	Map sample trenches	ZIP	1 Kb

GIS

GIS
software

CAD example

Wessex Archaeology (2021) *Site and Post-Excavation data from multiple fieldwork investigations at 35 West St, Wilton / Wilton Autos, 41-43 West St, Wilton, Wiltshire 2007-2020* [data-set]. York: Archaeology Data Service [distributor] https://doi.org/10.5284/1084798

CAD example

Phase 4

Trench 1, Trench 2, Trench 3, Trench 4, Trench 5, Trench 6, Trench 7, Trench 8, Trench 9, Trench 10, Area 10, Trench 11, Area 11, Trench 12, Trench 13, Trench 14, Area 14, Trench 15, Area 22, Trench 16, Area 17, Area 21, Trench 17, Trench 18, Trench 19, Trench 20

GIS example

Ends of evaluation trenches, as points
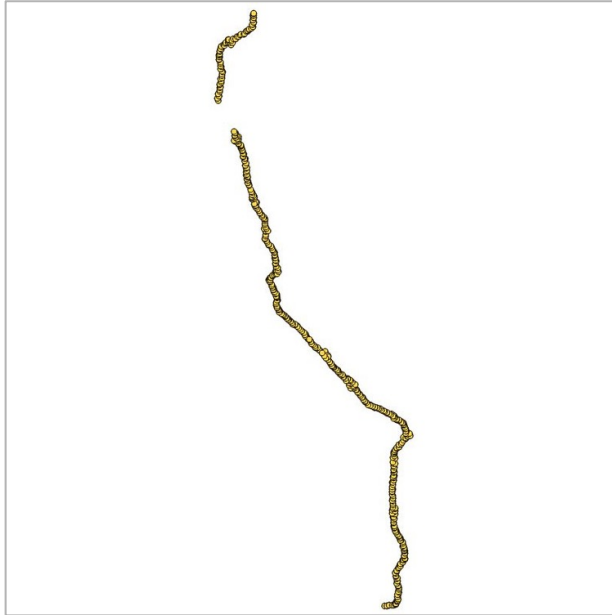
Areas of suspected archaeology

Birmingham Archaeology (2017) Wixoe to Kirtling Green Pipeline. Archaeological Evaluation [data-set]. York: Archaeology Data Service [distributor] https://doi.org/10.5284/1046266
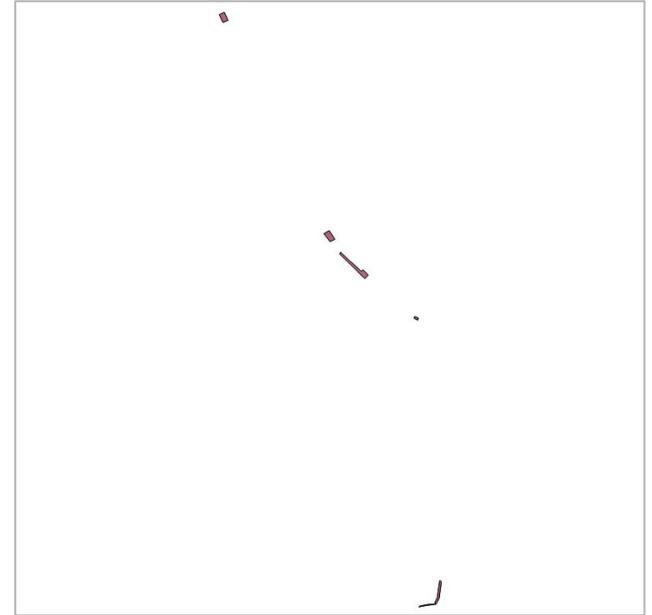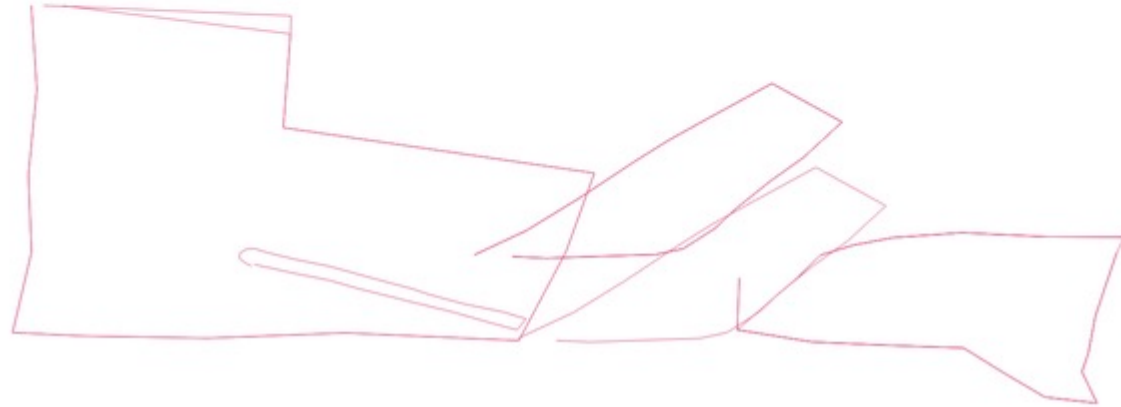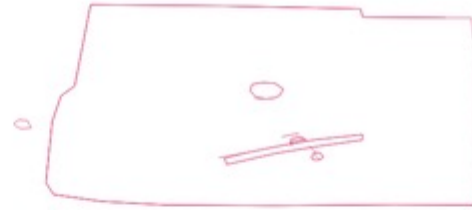
G**IS**

example

Simona Denis (2021) Site Data from an Archaeological Watching Brief at Station Road and to the rear of The Moors, Thatcham 2019-2020 [data-set]. York: Archaeology Data Service [distributor] https://doi.org/10.5284/1085016

Concluding thoughts

The information that CAD and GIS hold is important

CAD and GIS are all parts of an archive

GIS archives often have CAD deposited as well

Standards and file stability change as hardware and software change

Archaeology Data Service

http://archaeologydataservice.ac.uk

Department of Archaeology

The King's Manor

University of York

York, YO1 7EP

Thank you!

Teagan Zoldoske

Digital Archives Assistant, ADS
teagan.zoldoske@york.ac.uk