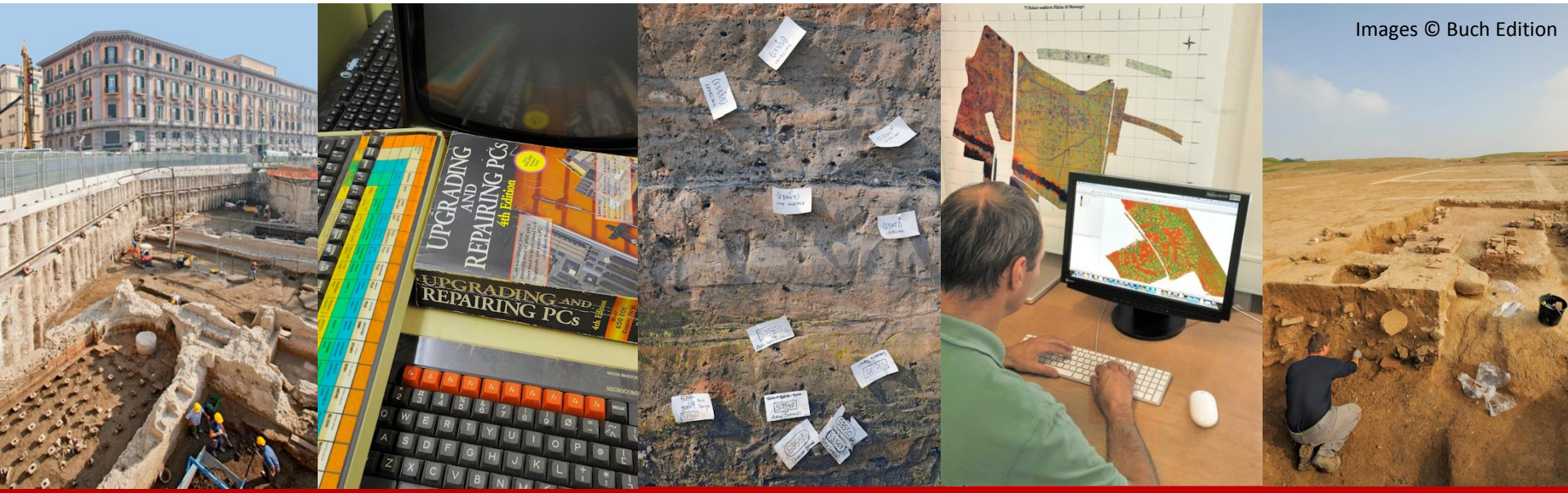
ARCHAEOLOGY DATA SERVICE

What metadata do users want and need, and how much can we get out of them?

Dr Katie Green
Archaeology Data Service
University of York
katie.green@york.ac.uk



Images © Buch Edition

ARCHAEOLOGY DATA SERVICE

The Archaeology Data Service

- Set up in 1996
- Originally part of Arts and Humanities Data Service (AHDS)
- Based at the University of York
- Initially 2 staff – now 13
- Held DSA since 2011
- Won DPC Decennial Award 2012

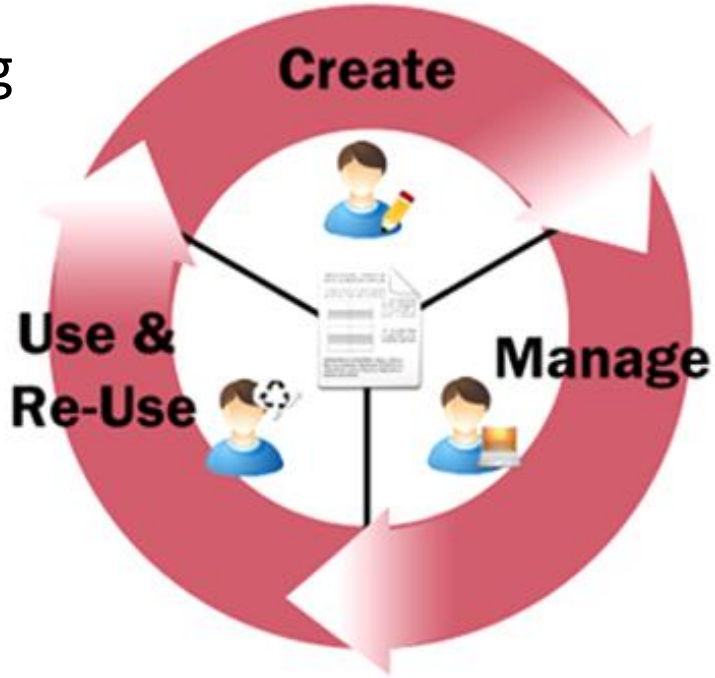
The Archaeology Data Service

Remit:

"Support research, learning and teaching with free, high quality and dependable digital resources"

Key Roles:

- Develop data collections
- Preserve and manage digital data
- Disseminate digital data



ADS resources

Archives

- 23 Journals and Series
- 34,500+ Grey Literature reports
- 700+ Project Archives (1657 collections)
- Six specialist Bibliographies
- 19 Doctoral Theses

ARCHAEOLOGY DATA SERVICE

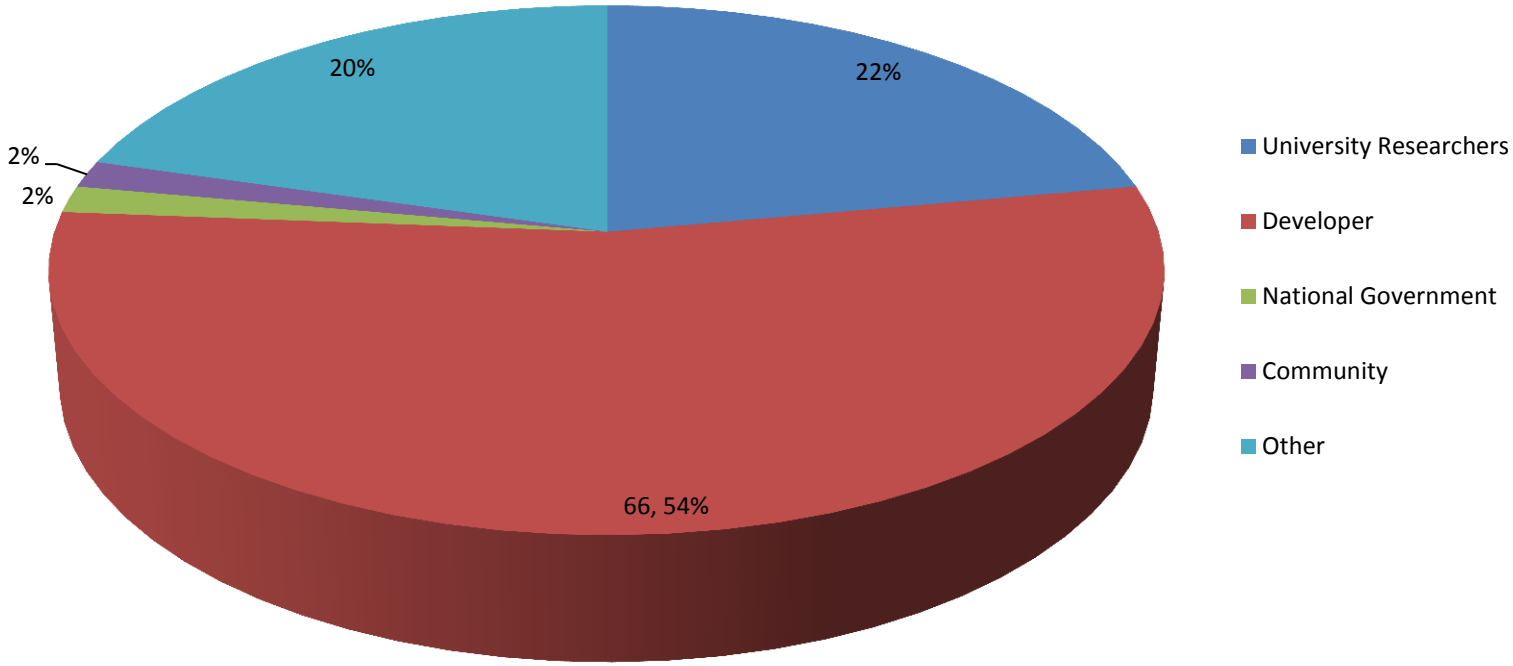
Archaeological data

ADS Archive

- 2.5 million files, 9.9 TB
- 189 formats deposited (not versions)
- 62 preservation formats
- 355 software types used (including versions)

Images © Buch Edition

ADS depositors?



ARCHAEOLOGY DATA SERVICE

What do we ask for?

Guidelines for Depositors

Version 2.0 September 2014.

CONTENTS

- Introduction to the Guidelines
- Why Deposit Data?
- Depositing with the ADS
  - What to deposit
  - How to deposit
  - Costs
- Preparing Collections for Deposit
  - Data Management Plans
  - File Management (Formats, Structure, Naming, Versioning)
  - Metadata
  - Selection and Retention
- File-level Metadata Requirements
  - Documents
  - Databases, Spreadsheets and Statistics
  - Raster Images
  - Geophysics and Remote Sensing
  - CAD and Vector Images
  - Geographical Information Systems
  - Video and Audio
  - Virtual Reality
  - Photogrammetry
- Collection-level Metadata
- Deposit Check List
- Downloads
- Acknowledgements

ARCHAEOLOGY DATA SERVICE

Example of completed metadata sheet for G

Project Title	Hardware/ Operating System	Software	Software Version	Date of Creation		Location (... term should ... new row)
				Start Date	End date	
A GIS for all the burials discovered in the vicinity of York Minster.	PC - Microsoft Windows XP Pro 2002	ESRI ArcGIS	10.1	01-Mar-12	20-Mar-12	UK
						England
						North York
						York
						York Minst

Assessment of Data Quality	Method of Data Capture	Purpose of Data Creation	Filename(s) (each filename should be added as a new row)	Description
High risk of spatial displacement	Digitisation using ArcGIS.	To create a GIS of all known Early medieval sites and finds from the Yorkshire area.	gis2.shp	A brief description of t contents of this group GIS files.
			gis2.shp	A brief description of t
			gis3.shp	A brief description of t

Jim | Beam | Archaeology

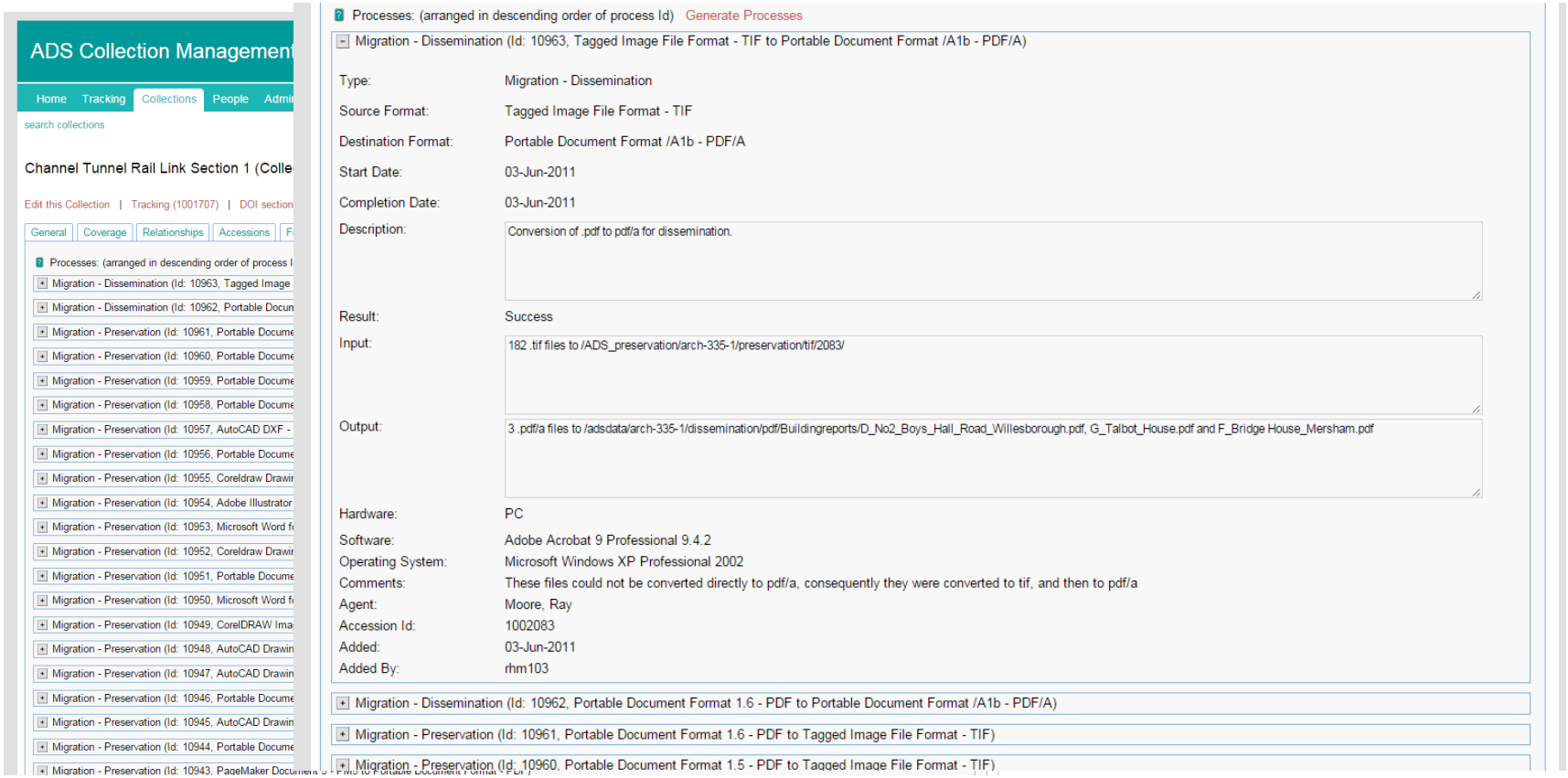
ARCHAEOLOGY DATA SERVICE

Example of a metadata sheet for Geophysics files deposited with the ADS

All Survey Types	Survey Name		The Big Field, Yorkshire.
	Survey Index		1234
	Description		A geophysical survey of the Big Field, Yorkshire.
	Survey Purpose		To find roman fort
	Location Coordinates (expressed in Lat/Long/OSGB)		123450 678900
	Location		UK, England, North Yorkshire, York
	Survey Duration	Start date	31-Jul-13
		End date	02-Aug-13
	Copyright Holder	Name	Brian Harvey
		Organisation	Suffolk County Council
	Solid Geology		Carboniferous Limestone
	Drift Geology		River terrace deposits
	Land Use		Arable
	Survey Type		Resistivity
	Instrumentation	Type	RMIIS
		Name	Resistivity Meter
	Area Surveyed (include unit of measurement)		100m2
	Grid Size (include unit of	Length	100m
		Width	50m
	Method of Coverage		Single Traverse
	Traverse Separation		2m
	Reading Interval		0.25
	Sampling Position		0.5
	Line Sequence		parallel
	Resolution		1
	Survey Direction		SSW
	Description of File Formats		Description text if needed
	File names (place a comma between each name)		grid1.txt, grid2.txt, grid3.txt
	Additional Remarks		Description text if needed
Electro-magnetic	Coil Configuration		
	Recorded Component		
Ground Penetrating Radar	Antenna Information		
	Time Delay		
	Time Sampling Resolution		
	Time Span		
	Average subsurface velocity		
	Average subsurface note		
Magnetometer	Magnetic north		
	Instrument drift		
Resistance	Electrode configuration		Dipole-Dipole
	Electrode spacing		0.5
	Multiple configurations		information from this sequence
Maritime Sonar	Average water velocity		
	Sonar frequency		
	Beam width at nadir		

http://archaeologydataservice.ac.uk

What else do we collect?



What do we get?

- Metadata completeness

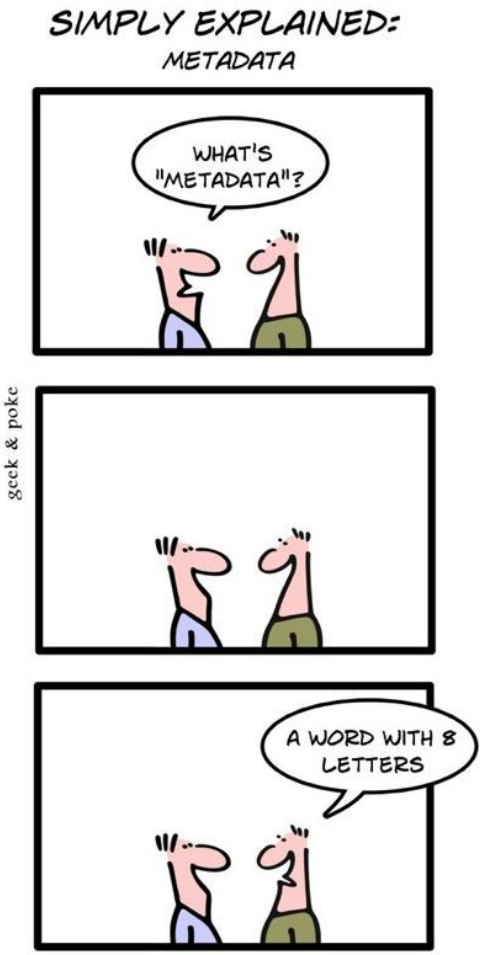
73%

- Amount of correspondence

One to two emails

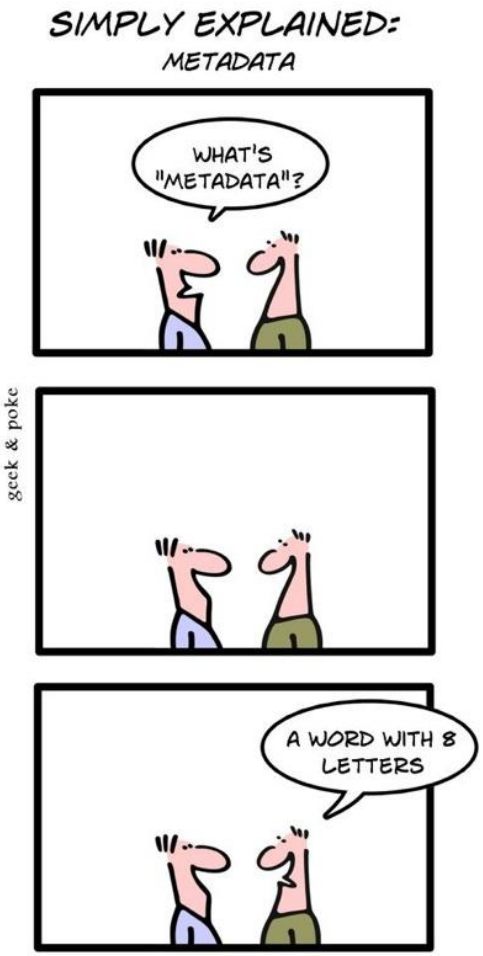
- Time take to get complete data

A couple of days to YEARs!



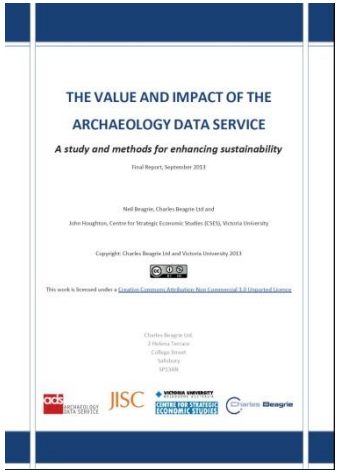
Depositor Feedback

- It takes too long
- I didn’t create the data so I don’t know how to fill out the metadata correctly
- It cost too much to pay people to create
- My projects finished, I am now working on the next project, and don’t have time
- It takes too long
- I don’t understand what I should be entering
- I don’t know what software created the data
- I don’t see the point
- I don’t want to
- It takes too long

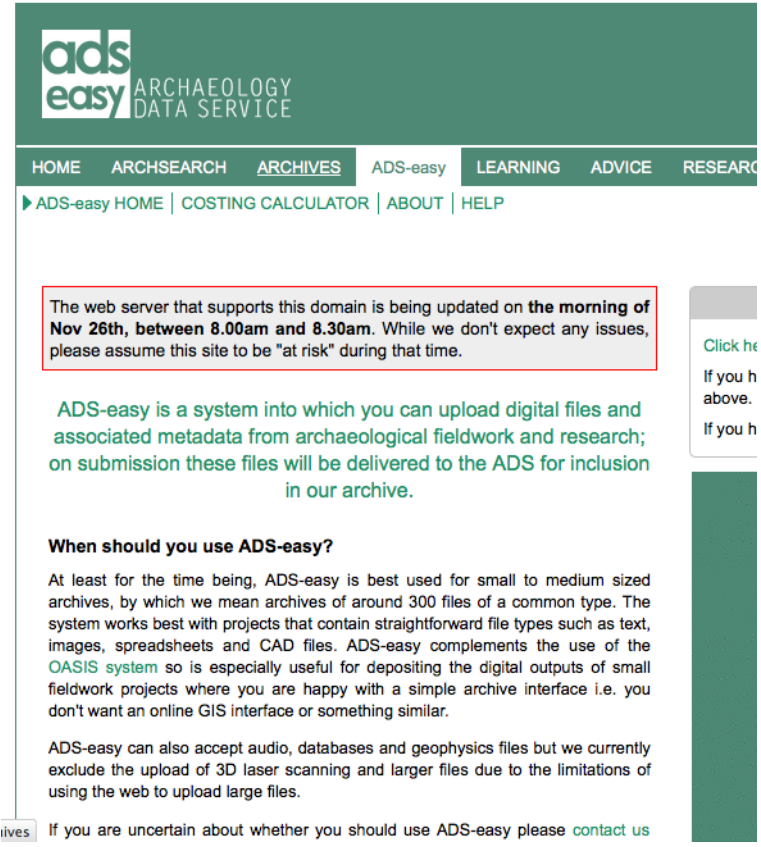


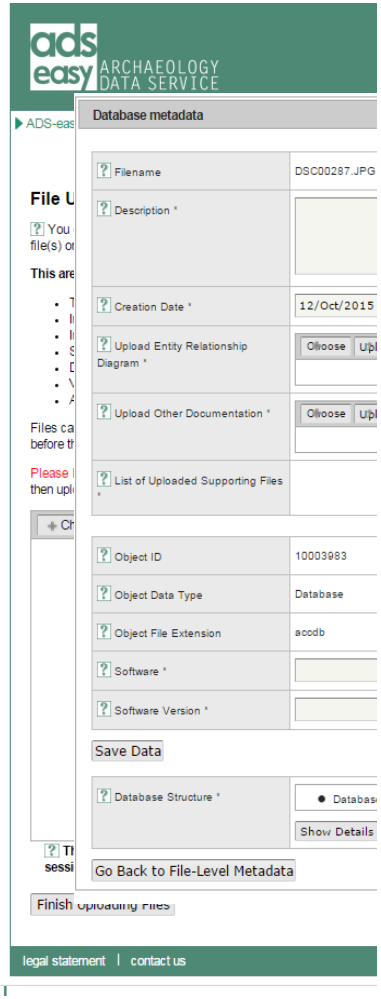
ADS Impact Study

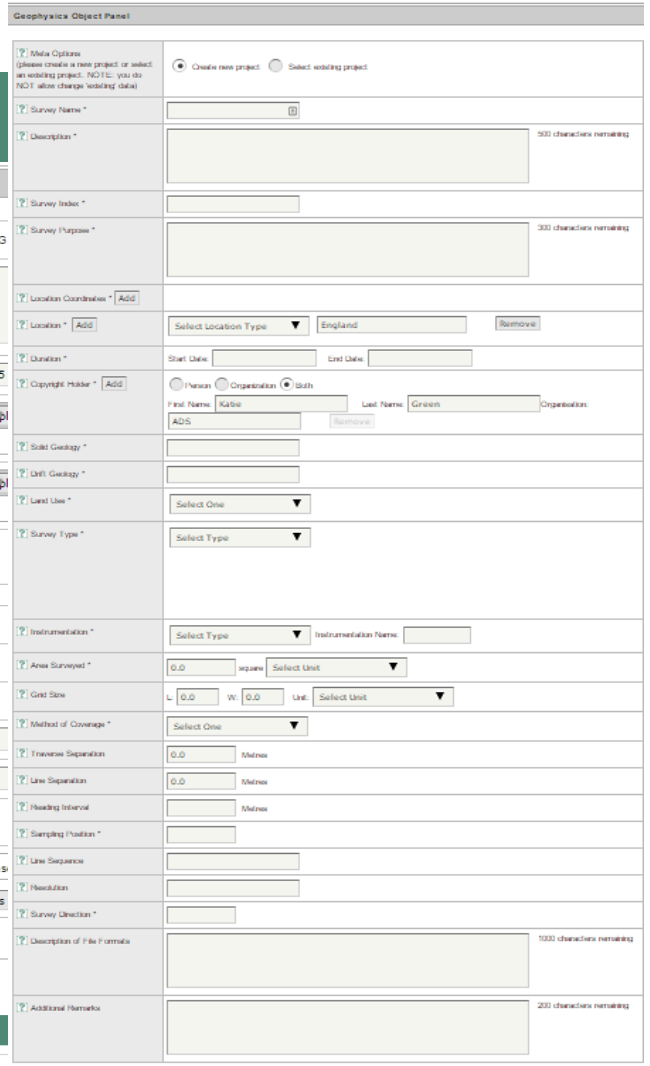
- Depositors reported a wide range of preparation and deposit times, from as little as **15 minutes** to more than **1,000 hours**.
- The **mean** reported preparation and deposit time was **82 hours** (median 15 hours) [N=65].
- One might expect initial deposits of the data to take longer than updates, but that was not the case for our respondents.
- Initial deposits took and average of **79 hours** (median 15) and updates an average 95 hours (median 15).



Addressing Feedback





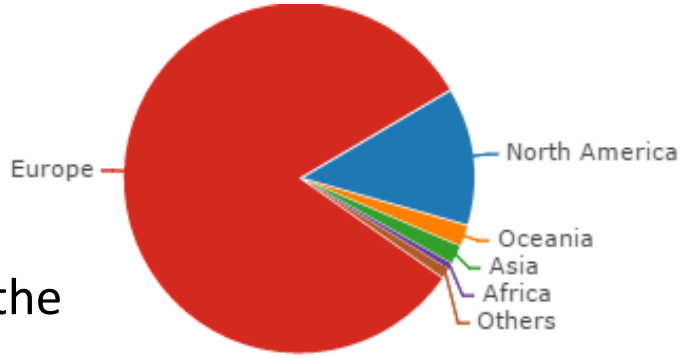


ADS-easy feedback

- It takes too long
- It cost too much to pay people to fill out form
- My projects finished, I am now working on the next project, and don't have time
- I didn't create the data so I don't know how to fill out the metadata correctly
- I don't know what software created the data
- **Too much metadata!**

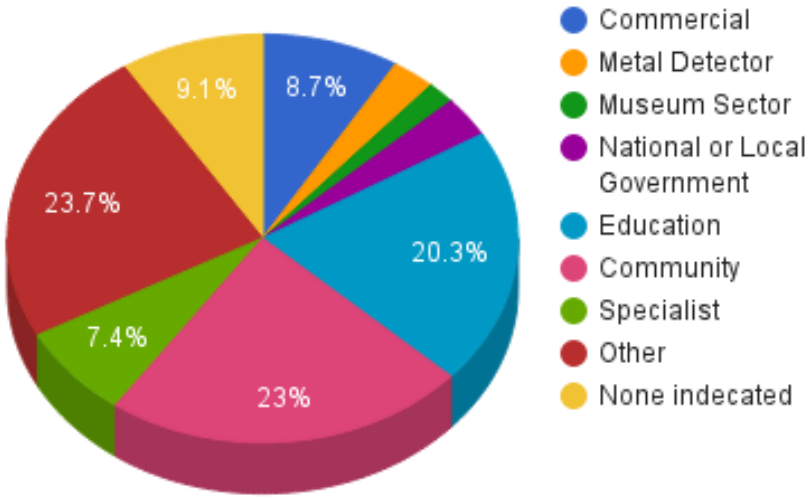
ARCHAEOLOGY DATA SERVICE

Who are our designated community?

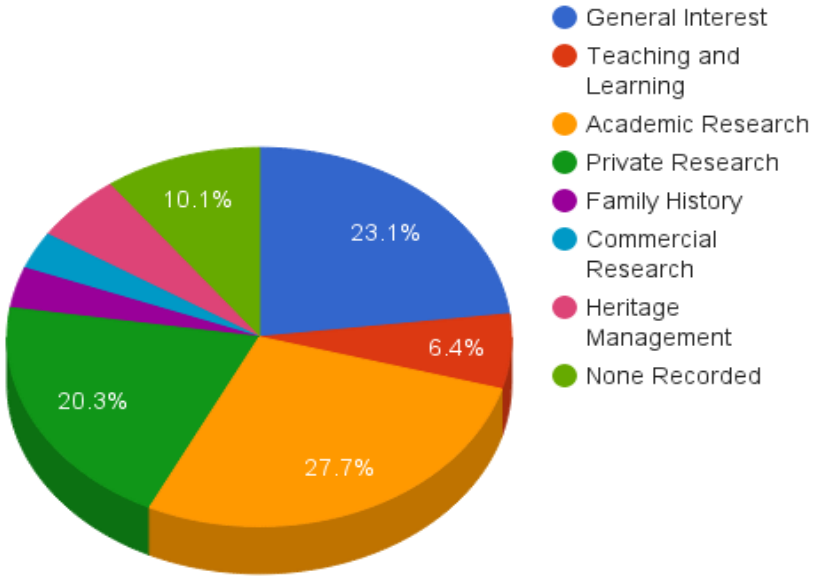


Historic environment professionals and members of the general public interested in the historic environment

Registered Users



ADS Primary Use



What do users want to know?

- What is the data about?
- How can I re-use data?
- Who created the data?
- How was data created?
- How old is the data/when was it updated?

Discovery

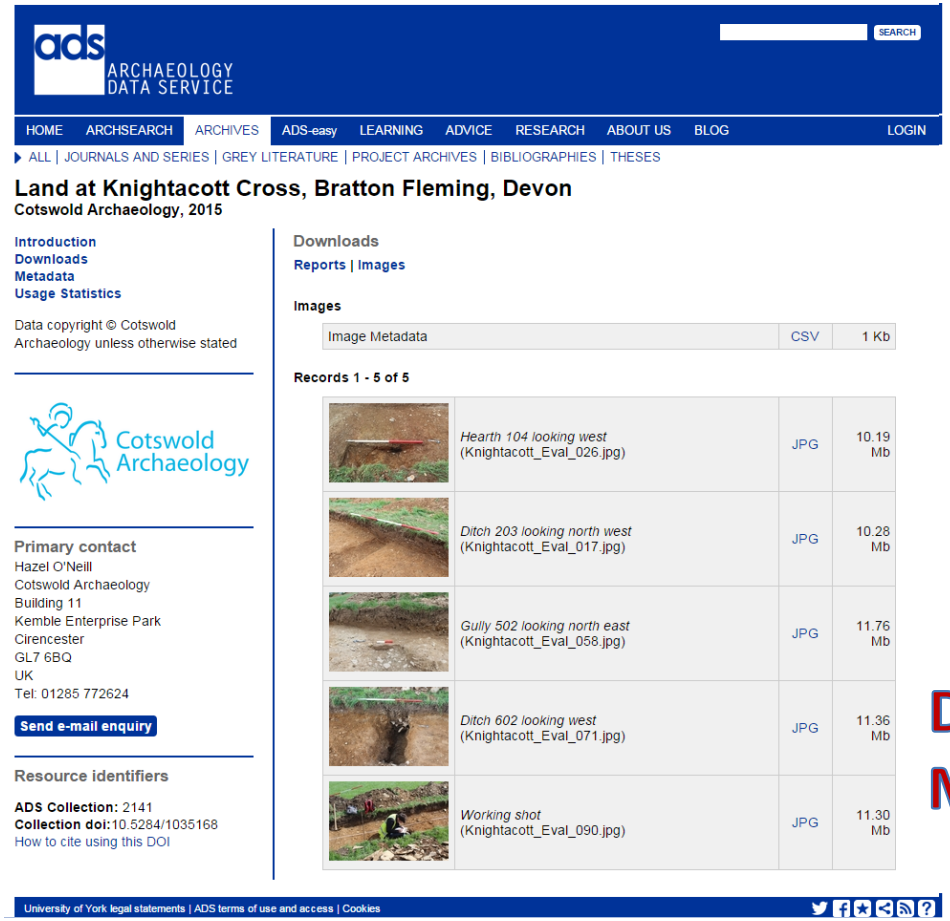
Rights

Creation

Event

ARCHAEOLOGY DATA SERVICE

What do we present to users?



ARCHAEOLOGY DATA SERVICE

SEARCH

HOME | ARCHSEARCH | ARCHIVES | ADS-easy | LEARNING | ADVICE | RESEARCH | ABOUT US | BLOG | LOGIN

ALL | JOURNALS AND SERIES | GREY LITERATURE | PROJECT ARCHIVES | BIBLIOGRAPHIES | THESES

Land at Knightacott Cross, Bratton Fleming, Devon
Cotswold Archaeology, 2015

Introduction
Downloads
Metadata
Usage Statistics

Data copyright © Cotswold Archaeology unless otherwise stated

Cotswold Archaeology

Primary contact
Hazel O'Neill
Cotswold Archaeology
Building 11
Kemble Enterprise Park
Cirencester
GL7 6BQ
UK
Tel: 01285 772624

Send e-mail enquiry

Resource identifiers

ADS Collection: 2141
Collection doi:10.5284/1035168
How to cite using this DOI

Downloads

Reports | Images

Images

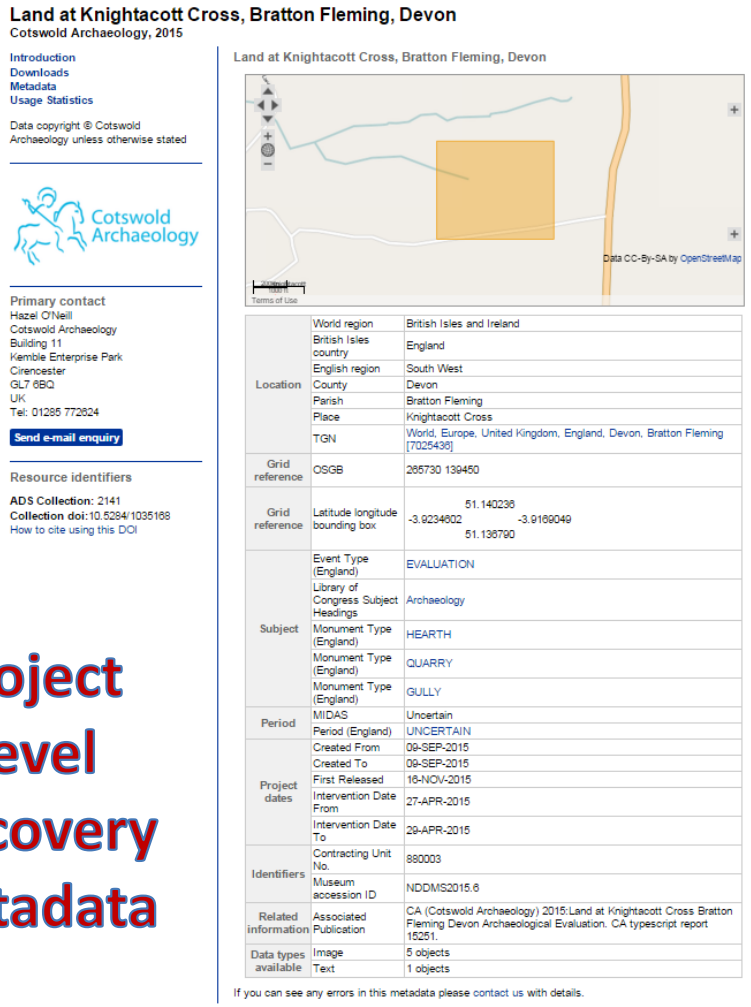
Image Metadata	CSV	1 Kb

Records 1 - 5 of 5

	Hearth 104 looking west (Knightacott_Eval_026.jpg)	JPG	10.19 Mb
	Ditch 203 looking north west (Knightacott_Eval_017.jpg)	JPG	10.28 Mb
	Gully 502 looking north east (Knightacott_Eval_058.jpg)	JPG	11.76 Mb
	Ditch 602 looking west (Knightacott_Eval_071.jpg)	JPG	11.36 Mb
	Working shot (Knightacott_Eval_090.jpg)	JPG	11.30 Mb

University of York legal statements | ADS terms of use and access | Cookies

Project Level Discovery Metadata



Land at Knightacott Cross, Bratton Fleming, Devon
Cotswold Archaeology, 2015

Introduction
Downloads
Metadata
Usage Statistics

Data copyright © Cotswold Archaeology unless otherwise stated

Cotswold Archaeology

Primary contact
Hazel O'Neill
Cotswold Archaeology
Building 11
Kemble Enterprise Park
Cirencester
GL7 6BQ
UK
Tel: 01285 772624

Send e-mail enquiry

Resource identifiers

ADS Collection: 2141
Collection doi:10.5284/1035168
How to cite using this DOI

Land at Knightacott Cross, Bratton Fleming, Devon

Location	World region	British Isles and Ireland
	British Isles country	England
	English region	South West
	County	Devon
	Parish	Bratton Fleming
	Place	Knightacott Cross
	TGN	World, Europe, United Kingdom, England, Devon, Bratton Fleming [7025436]
Grid reference	OSGB	265730 139450
Grid reference	Latitude longitude bounding box	51.140236 -3.9234602 -3.9169049 51.136790
Subject	Event Type (England)	EVALUATION
	Library of Congress Subject Headings	Archaeology
	Monument Type (England)	HEARTH
	Monument Type (England)	QUARRY
	Monument Type (England)	GULLY
Period	MIDAS	Uncertain
	Period (England)	UNCERTAIN
Project dates	Created From	09-SEP-2015
	Created To	09-SEP-2015
	First Released	16-NOV-2015
	Intervention Date From	27-APR-2015
	Intervention Date To	29-APR-2015
Identifiers	Contracting Unit No.	880003
	Museum accession ID	NDDMS2015.6
Related information	Associated Publication	CA (Cotswold Archaeology) 2015:Land at Knightacott Cross Bratton Fleming Devon Archaeological Evaluation. CA typescript report 15251.
Data types available	Image	5 objects
	Text	1 objects

If you can see any errors in this metadata please contact us with details.

University of York legal statements | ADS terms of use and access | Cookies

Metadata collection alternatives

- NLP on text
- Image recognition
- Autogenerate some technical metadata

From Internal Pages
archaeologydataservice.ac.uk/archives/view/cottamb_2015/query.cfm 39%
archaeologydataservice.ac.uk/archives/view/cottamb_2015/ 18%
archaeologydataservice.ac...6d6574616461746 12%
archaeologydataservice.ac.uk/archives/view/cottamb_2015/metadata.cfm 12%
archaeologydataservice.ac.uk/archives/view/cottamb_2015/results.cfm 6.1%
Others 12%
/archives/view/cottamb_2015/downloads.cfm
50 pageviews
Incoming traffic
33 from internal pages
0 from internal searches
0 from search engines
0 from websites
0 from campaigns
0 direct entries
Outgoing traffic
16 to internal pages
0 internal searches
23 downloads
0 outlinks
24 exits
10 page reloads
To Internal Pages
Downloads
78% /catalogue/adsdata/arch-2117-1/dissemination/csv/Cottam_full_database.csv
22% /catalogue/adsdata/arch-2117-1/disse...umentation/spreadsheet_metadata.pdf
Exits

What do users need?

ads ARCHAEOLOGY DATA SERVICE

HOME MAP

ADS 3D V

Introduction

Las Cuevas site is l
Las Cuevas Resear
Chiquibul Forest Re
Belize. This site is a
located approximate
Caracol (for more in
site see Moyes et al
et al. 2014).

Primary contac

Dr Fabrizio Galeazz
Department of Arch
University of York
King's Manor
York
YO1 7EP
England

Send e-mail enquiry

Resource identifiers

ADS Collection: 1077
Collection doi:10.5284
How to cite using this D

1-Project Description

Project Name	Virtual Amarna Project
Name of monument, survey area, or object	NA
Monument/Object Number	8761
Monument/Object Description	Small clay mould - Tututkamun
Survey Location	Amarna, Egypt
Survey Date(s)	17-Mar-2008
Survey Conditions	Indoors
Scanner Details	Konica Minolta VIVID 9i; mm; Serial No: 1001198
Company/Operator Name	Center for Advanced Spatial Technologies,Christopher Goodmaster
Control data collected?	No

4-Mesh Metadata

Pre-Meshing Metadata

Name of Pre-Mesh Dataset	8761_GRE.txt
Number of Points in File	53032
Overlap Reduction	Y
Smoothing	N
Subsampling	N
Color Editions	N
Point Deletion Summary	Overlap reduction was computed in Polyworks software. Following overlap reduction, floating data points were also deleted. Data remnants from overlap reduction were also deleted as necessary.

Polygonal Mesh Metadata:

Name of Mesh Dataset	8761_hi.obj
Holes Filled	Y
Smoothing	Y
Color Editions	Y
Healing/despiking	Y
Total Triangle Count (post editing, predecimation)	120848
RGB Color Included	Y
Data Reduction	N
Coordinate System Adjustment	Y
CS Adjustment Matrix	Cannot export from software
Additional processing notes	Mesh was color corrected in Rapidform XOR.

Decimated Polygonal Mesh Metadata:

Name of Decimated Mesh Dataset	8761_lo.obj
Total Original Triangle Count	120848
Decimated Triangle Count	25000
RGB Color Preserved from original dataset	Y

e the scan artifacts. All additional lighting/color adjustments on the

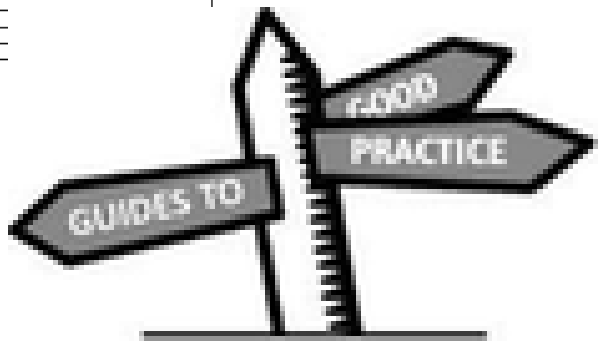
0.153

8

iet, images

Belize)

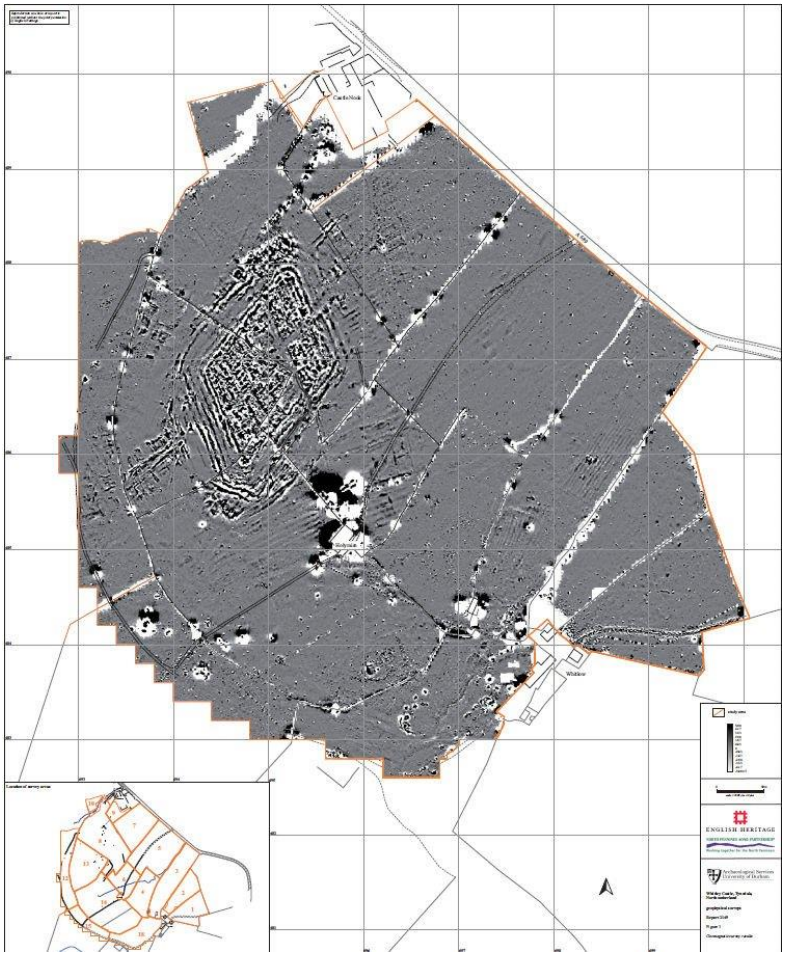
p of Structure 1,
A. Unit 9 is 3 by 3
o the centre of the
was excavated
ıster floors were
ialeazzi et al.



http:/

What do users need?

- More than they realise?



Creation Event

Who are our metadata users?

Migrating data

A Database of Upper Palaeolithic Sites based on Wymer and Bonsall, 1977 (1001020)

Object Id	Object Title	Data Type	Created Date	
1001566	LaterUpperPalaeolithic-Site	Spreadsheet	22-Jan-2010	LaterUpperPalaeolithic-Site.csv LaterUpperPalaeolithic-Site.csv LaterUpperPalaeolithic-Site.csv
1001565	LaterUpperPalaeolithic-Reference	Spreadsheet	22-Jan-2010	LaterUpperPalaeolithic-Reference.csv LaterUpperPalaeolithic-Reference.csv LaterUpperPalaeolithic-Reference.csv
1001564	LaterUpperPalaeolithic-Object	Spreadsheet	22-Jan-2010	LaterUpperPalaeolithic-Object.csv LaterUpperPalaeolithic-Object.csv LaterUpperPalaeolithic-Object.csv
1001563	LaterUpperPalaeolithic-Museum	Spreadsheet	22-Jan-2010	LaterUpperPalaeolithic-Museum.csv LaterUpperPalaeolithic-Museum.csv LaterUpperPalaeolithic-Museum.csv
1001562	LaterUpperPalaeolithic-Excavator	Spreadsheet	22-Jan-2010	LaterUpperPalaeolithic-Excavator.csv LaterUpperPalaeolithic-Excavator.csv LaterUpperPalaeolithic-Excavator.csv
1001561	LaterUpperPalaeolithic-Compiler	Spreadsheet	22-Jan-2010	LaterUpperPalaeolithic-Compiler.csv LaterUpperPalaeolithic-Compiler.csv LaterUpperPalaeolithic-Compiler.csv
1001560	LaterUpperPalaeolithic-Collection	Spreadsheet	22-Jan-2010	LaterUpperPalaeolithic-Collection.csv LaterUpperPalaeolithic-Collection.csv LaterUpperPalaeolithic-Collection.csv
1001559	EarlierUpperPalaeolithic-Site	Spreadsheet	21-Jan-2010	EarlierUpperPalaeolithic-Site.csv EarlierUpperPalaeolithic-Site.csv EarlierUpperPalaeolithic-Site.csv
1001558	EarlierUpperPalaeolithic-Reference	Spreadsheet	21-Jan-2010	EarlierUpperPalaeolithic-Reference.csv EarlierUpperPalaeolithic-Reference.csv EarlierUpperPalaeolithic-Reference.csv
1001557	EarlierUpperPalaeolithic-Object	Spreadsheet	21-Jan-2010	EarlierUpperPalaeolithic-Object.csv EarlierUpperPalaeolithic-Object.csv EarlierUpperPalaeolithic-Object.csv
1001556	EarlierUpperPalaeolithic-Museum	Spreadsheet	21-Jan-2010	EarlierUpperPalaeolithic-Museum.csv EarlierUpperPalaeolithic-Museum.csv EarlierUpperPalaeolithic-Museum.csv
1001555	EarlierUpperPalaeolithic-Excavator	Spreadsheet	22-Jan-2010	EarlierUpperPalaeolithic-Excavator.csv EarlierUpperPalaeolithic-Excavator.csv EarlierUpperPalaeolithic-Excavator.csv
1001554	EarlierUpperPalaeolithic-Compiler	Spreadsheet	21-Jan-2010	EarlierUpperPalaeolithic-Compiler.csv EarlierUpperPalaeolithic-Compiler.csv EarlierUpperPalaeolithic-Compiler.csv
1001553	EarlierUpperPalaeolithic-Collection	Spreadsheet	21-Jan-2010	EarlierUpperPalaeolithic-Collection.csv EarlierUpperPalaeolithic-Collection.csv EarlierUpperPalaeolithic-Collection.csv
1001552	DatabaseNotes-PalaeolithicSites	Text	22-Jan-2010	DatabaseNotes-PalaeolithicSites.doc DatabaseNotes-PalaeolithicSites.docx DatabaseNotes-PalaeolithicSites.pdf
1001567	PalaeolithicSites_ADS_Metadata	Text	04-Jan-2011	PalaeolithicSites_ADS_Metadata.docx

Processes: (arranged in descending order of process Id) Generate Processes

Migration - Dissemination (Id: 10963, Tagged Image File Format - TIF to Portable Document Format /A1b - PDF/A)

Type:	Migration - Dissemination
Source Format:	Tagged Image File Format - TIF
Destination Format:	Portable Document Format /A1b - PDF/A
Start Date:	03-Jun-2011
Completion Date:	03-Jun-2011
Description:	Conversion of .pdf to pdf/a for dissemination.
Result:	Success
Input:	182 .tif files to /ADS_preservation/arch-335-1/preservation/tif/2083/
Output:	3 .pdf/a files to /adsdata/arch-335-1/dissemination/pdf/Buildingreports/D_No2_Boys_Hall_Road_Willesborough.pdf, G_Talbot_House.pdf and F_Bridge House_Mersham.pdf
Hardware:	PC
Software:	Adobe Acrobat 9 Professional 9.4.2
Operating System:	Microsoft Windows XP Professional 2002
Comments:	These files could not be converted directly to pdf/a, consequently they were converted to tif, and then to pdf/a
Agent:	Moore, Ray
Accession Id:	1002083
Added:	03-Jun-2011
Added By:	rhm103

Migration - Dissemination (Id: 10962, Portable Document Format 1.6 - PDF to Portable Document Format /A1b - PDF/A)

Migration - Preservation (Id: 10961, Portable Document Format 1.6 - PDF to Tagged Image File Format - TIF)

Migration - Preservation (Id: 10960, Portable Document Format 1.5 - PDF to Tagged Image File Format - TIF)

What metadata do users want and need, and how much can we get out of them?

- Know who your users are
- User don't always know what they need
- One size does not fill all users/data types
- People will always complain
- Do whats best for the archive

Follow us on Twitter:
@ADS_Update

Like us on Facebook:
http://www.facebook.com/archaeology.data.service

Katie Green

katie.green@york.ac.uk

Archaeology Data Service
Department of Archaeology
The King's Manor
University of York
York
YO1 7EP