

This PDF is a simplified version of the original article published in Internet Archaeology under the terms of the Creative Commons Attribution 3.0 (CC BY) Unported licence. Enlarged images, models, visualisations etc which support this publication can be found in the original version online. All links also go to the online original.

Please cite this as: Zoldoske, T. 2024 Metadata for Discovery. Planning for an Information Network, Internet Archaeology 65. https://doi.org/10.11141/ia.65.6

Metadata for Discovery. Planning for an Information Network

Teagan Zoldoske

In advance of building a new high speed rail line aimed at connecting the North and South of England, the UK is undertaking the largest archaeological intervention in its history. While the High Speed 2 (HS2) project has created a wealth of new information, before any of that information can be disseminated to the general public, it must first be properly collected, documented, and linked. To this end, data collection is integral to facilitate effective data dissemination and FAIR - Findable, Accessable, Interoperable and Reusable - data to achieve the greatest public value for an archive. The most exciting search results and maps come from what is often seen as scary technical jargon.

This paper will discuss some of the limitations the Archaeology Data Service (ADS) experiences as a digital repository, what is currently being done to maximise the reach of the ADS collections, and what tools have been created to aid both depositors and digital archives alike. Starting at the beginning of the data life cycle, this article will show how large infrastructure projects like HS2 allow the ADS to work with depositors to raise issues about data collection, generation, and description during project development and how collaborative efforts can improve the creation and import of data and metadata into the archive. The ADS has made steps towards keeping our data FAIR but simple, both by streamlining what metadata is essential to foster better discovery and reuse within an archive, and from there, how metadata can be passed to external data catalogues such as ARIADNEplus and The National Archives in the UK.

1. Introduction

A digital archive and the events leading up to its deposit are often seen as two distinct phases in the data life cycle. This demarcation becomes even more pronounced particularly when the archive is deposited with an independent repository. However, the data life cycle does not conclude with archives. Re-use allows the cycle to begin again. In the case of large projects, the intrinsic interconnectivity of the data produced creates immense potential for re-use if planned for from the outset. The High Speed 2 (HS2) rail project is an example of a project that has considered issues related to archiving and re-use from its inception. The scale of the project is immense (the UK's largest ever linear infrastructure project) and the amount of archaeological data that is being produced is equally vast (Solly 2018). As such, HS2 provides a once in a lifetime opportunity to digitally preserve this wealth of archaeology before disseminating it to the wider public so that community groups, researchers, and more can learn about what was discovered. It is the responsibility of the <u>Archaeology Data</u> <u>Service</u> (ADS) to ensure that these data archives are preserved and disseminated.

To effectively manage the amount of information being produced and given to the ADS, the ADS relies heavily on the metadata that is deposited alongside the primary data. Metadata is the information about the data (Niven and Watts 2011). It describes and informs what the data concerns are while improving the findability, accessibility, interoperability, and re-use of the digital asset that the data is associated with (FAIR principles). For example, take the image in Figure 1. What is it? Where was it created and when? The metadata tells us that this is a NW-facing, post-excavation shot of an evaluation trench as a part of works for High Speed 2. It was taken by S.T. for Wardell Armstong on 16 August 2019 at Newyears Green Lane in the London Borough of Hillingdon. The metadata can also tell us more technical information such as the MD5 checksum (an identifier composed of alphanumerics that changes if the underlying data changes in any way, allowing for verification of data integrity), file name, format and format type, file size and more. Metadata is the key to being able to effectively manage and disseminate the archive and is necessary for data findability and interoperability. This is why the ADS has high requirements for its metadata. The more tools that exist to help with incorporating high quality metadata into an archive, the easier it becomes to produce archives with high re-use potential. High Speed 2 has provided an opportunity to evaluate our workflows and develop new tools and dissemination methods due to its data and associated metadata.



Figure 1: An image of a Trench from a Trial Trenching at Newyears Green, Hillingdon, 2019 (HS2 Phase One) [data-set]. York: Archaeology Data Service [distributor] <u>https://doi.org/10.5284/1106774</u>



1.1 The Archaeology Data Service

The Archaeology Data Service (ADS) is the digital repository responsible for archiving and disseminating the archaeological data produced as a part of HS2. Founded in 1996, the ADS is a <u>CoreTrustSeal</u> certified digital repository and ensures the longterm digital preservation of the data entrusted to its care. It does this through the migration of data formats and the continual curation of collections so that the data it holds does not become obsolete (Richards 1997). At the time of writing (2023), the ADS holds over 1.4 million records of archaeology from the United Kingdom and beyond. These records amount to over 45 terabytes of data from over 6,000 individual data archives. Within the 2021-2022 reporting period alone, the ADS received 622 individual archives with 203 of those derived from HS2. Archives contain different types of data, including standard formats like text and images as well as more complex types such as GIS and 3D models, all with their associated metadata.

1.2 High Speed 2 (HS2)

The High Speed 2 project's goal is to improve rail travel between London and Birmingham in the first phase before continuing into the North of England with links extending into Scotland. As a part of the heritage works prior to construction, High Speed 2 became Europe's largest ever archaeological excavation in 2022. The amount of data produced is expected to be over 15TB for Phase 1 alone.

To date, HS2 has archived over 10TB of archaeological data with the ADS distributed over 300 individual archives. These archives consist of various types of data, which the ADS is not only committing to preserve but also disseminate. To ensure that the data being produced as a part of these infrastructure works preserves the historic environment to its fullest extent, HS2 created a Historic Environment Research and Delivery Strategy (HERDS) covering built heritage, archaeology, and the historic landscape (HS2 2017). The strategy outlines how the digital archive is a headline objective and an integral part of the project and will contribute to the lasting legacy of its historic environment programme. This means that the archives will have elements such as 3D model views, spatial searches created from GIS data, search interfaces based on spreadsheets and databases . With so much data being produced and enhanced interfaces being created, the key to managing and building any of it lies in the metadata sitting behind it.

1.3 Metadata standards

The ADS has specific metadata requirements for all data that are deposited (Archaeology Data Service 2023) and these requirements were built into HS2's technical specifications (see Farshid <u>this issue</u>) prior to starting works. Depending on the data type, the metadata may only need to contain the core metadata fields while some complex data types need additional technical metadata. These requirements were created taking into consideration community standards and are listed in the



ADS's Cataloguing Policy (O'Brien and Evans 2022). Drawing on established metadata standards such as Dublin Core and International Organization for Standardization (ISO) (DCMI-TERMS; International Organization for Standardization 2019), the ADS created consistent requirements for all data that the ADS accepts. Metadata serves as both an aid to understanding the data and as a tool for increasing the discoverability of said data. The ADS's standards ensure that the ADS's archives meet the FAIR principles. With large projects such as HS2, there are often many different companies contracted to contribute to the overall programme of works. This has potential to lead to variations in the data being deposited even if they are all working to the same specification. To ensure that all contractors submit the same standard of metadata and reduce potential variation within the metadata submitted, the ADS's digital archivists individually check all metadata submitted by HS2. This individualised approach allows for any potential keywords, for example, that may have been missed during metadata creation to be added to the archive or individual data to improve the overall quality of the archives. These keywords are matched to thesauri such as FISH Event Types Thesaurus or Archaeological Objects Thesaurus (Scotland) and increase the discoverability of the datasets or individual data.

2. Metadata from the Beginning

Ideally, a method or standardised procedure should be in place to allow for metadata attribution to be assigned to data as they are created and enhanced over time regardless of method of creation. In order for that to happen, it is important to think about how the data will change and be used throughout their life cycle. The data life cycle within a project cannot be seen as a linear path that goes directly from project planning to execution before finally archiving the results in a museum box in the hope of being used one day in the future. It is cyclical in nature and collaboration is needed from the start in order to optimise the data. Data repositories should be involved from the conception of the project, ensuring that preservation and description expertise informs how data are formed and managed right from the start.

A headline objective for HS2 is to create a highly accessible and outstanding archival legacy (objective 3, HS2 2017). This objective focussed HS2 right from the beginning on how to produce data that will in turn create an interesting archive that would be accessible and purposeful for everyone (professional and lay users alike). Focusing on the outcomes from the beginning means that file formats and data types are agreed upon. Repositories are responsible for the data they care for and in the ADS's case, forever. Therefore to ensure that data can be properly cared for, thought must be given to the stability and suitability of file formats while also considering community standards. As software and hardware technologies progress, the risk of older files becoming obsolete increases. To a certain extent, this is unavoidable, and when these older files become inaccessible using modern software, the ADS migrates these older files to newer formats . To minimize the number of migrations and risk of software obsolescence, not all file formats are acceptable for archival purposes and knowing that can change how the data are created and exported thus simplifying the archival process.



As a result of these complexities, HS2 involved the ADS from the outset to assist with scoping the project. This involved discussing how and what kinds of data would be produced and what the ADS could and would do with the data. This has allowed for open communication about what would be needed to build search interfaces for these collections throughout the life of the project and improves each archive's reuse potential from the moment the data was created. Through such discussions, preferred file formats and required additional information, such as keywords to describe the data and vocabularies for these keywords, were decided upon right at the beginning.

2.1 Limitations

There are limitations however. No matter what the ideal is, in the end the data must be usable and practical for the duration of their active use within a project. For instance, many context drawings are still made by hand. It would be impractical to demand that they all be digital from the beginning; sites may be remote and lacking electricity or network connectivity. From experience, metadata creation gets more difficult and time-consuming the more time has passed since the data's creation and the more 'hands' the data has passed through. Data is often created by multiple people and even more may contribute towards documenting or processing the data before it ever gets to the archival stage in the life cycle. With so many people working on the project data, there is also a risk that details might get lost. When basic details such as the subject of a photograph are missing, some of the value can be permanently diminished: after an excavation it can be impossible to tell the significance of a seemingly empty trench.

3. Metadata for Discovery

Good metadata can be the key to not only understanding but also locating data. Data becomes findable due to its associated metadata. Keywords that clearly describe the data allow for that data to be discovered in searches giving the opportunity for reuse. Keywords with the same vocabularies can be linked to regional and national records and thus increase the archives' interoperability. An example of this is the inclusion of ADS data in external catalogues such as the <u>ARIADNEplus Portal</u>, which provides a central access point to data from twenty-three European countries. Metadata from the ADS is mapped to key terms from the <u>Getty Thesaurus of</u> <u>Geographic Names</u> (TGN) which acts as a link allowing data from the ADS to be found both within the ADS's website or from the <u>ARIADNEplus Portal</u>.

Q, AAT search	
Burial Q. 60	
Dural Incords	^
Londa registers	
Local Stocs Local Local Region (Inc. 1997)	
Vordal Votos Vordal	
Constant	
datale banks	
Interary masks	
	>
22 #	asults
burials (see http://worab.oxfty.edu/au/300263485)	^
Broader: Ceremones Cultural events	
Narrower: 🗣 rebuild	
Note: Acts of interment, usually but not necessarily ceremonial, including preparation of the corpse and its placement in the grave, and all accompanying ritual practices,	1
any.	
Hierarchy	
Terms	
	~
ç.	,

Figure 2: Vocabulary mapping to AAT by ARIADNEplus

Figure 2 shows <u>ARIADNEplus's vocabulary mapping tool</u> which was designed for interoperability. By searching for a vocabulary term, one thesaurus can be mapped to another thesaurus to create a link which widens the reach of the data or archives attached to that term. The 'Evidence Thesaurus' from the UK's Forum on Information Standards in Heritage (FISH) is used with the term 'Burial' as an example vocabulary term. When this term is searched, a list of potential matches are displayed and can be selected from, in this case, the Getty's <u>TGN</u> term 'Burials'. The ADS uses this kind of mapping within its internal management systems to allow for mapping to external catalogues (such as ARIADNEplus, TNA Discovery, MEDIN) which in turn amplifies the reach of the ADS's data.

Keywords do not solely increase the reach and findability of data from an archive out to external catalogues but can also be used within an archive itself. Figure 3 shows an example of how keywords assigned to individual images allow a refinement tool to filter images within a HS2 collection. This refinement allows for specific images to be found from potentially thousands of images within an archive. Refinement via keyword is not limited to images, however, and can be applied to larger and more complex searches across a single archive or all of the HS2 archives. It even has the potential to link files to files, for example images to context drawings, if the proper links between the data are also documented and submitted within the metadata.



	Kiln	×	Search Images		
D	Human remains	^			
	Iron				
	Kiln				Into
	Land Drain		ADS_raster_metadata.csv (717 kB)	CSV	Download
F	Layer				
1	Level				
	Limestone				
	Linear		East facing oblique shot of north extention		Info
	Medieval Building		rying oven (1m x 2m scale) name: 1C20LOWAR_C1_0568.jpg (4 MB)	JPG	
	Metal				Download
	Mortar				
	Needle Holder		North facing shot of oven/ kiln (1034) (2m		Info
	Oven		e)	2	
	Pit		name: 1C20LOWAR_C3_0646.jpg (4 MB)	JPG	Download
	Pitchfork				

Figure 3: Use of image refinement for the sorting of images within the dissemination page. Example from High Speed Two Ltd., Archaeological Research Services Ltd (n.d.) from Archaeological Recording at Deserted Medieval Village of Lower Radbourne, Warwickshire, 2021-2022 (HS2 Phase One).

4. Tools for Bringing Metadata into Archives

Even with proper planning, however, it can be difficult to ensure that all links within an archive exist. Given the number of contractors working under HS2 all sending their data on to both HS2 and the ADS, several areas that had the potential for things to go wrong were identified before the first archive was submitted. For established institutions, changing workflows to account for new requirements or even trying to export data from their software could be time consuming and allow for errors to be introduced into the workflow. To try and address these problems, the ADS created a number of tools as a part of its work with HS2. These tools aimed at making the creation, depositing, and archiving of datasets with their corresponding metadata easier for both the depositor and the digital archivist. Work on these tools, however, did not stop once the first archives were deposited and as new opportunities to help smooth the process appeared, new tools were developed to help. High Speed 2 is providing the ADS an opportunity for these tools to be developed and tested before becoming integrated into the ADS's systems and guidance for the benefit of the archaeology sector more widely.



4.1 Metadata mapping

The metadata behind any data deposited with the ADS is the key towards generating any of the ADS's webpages and the system that connects it all is the ADS's Object Management System (OMS). It is for this reason that all metadata deposited with the ADS needs to be mappable to the OMS. Figure 4 shows how data deposited with the ADS appears on the ADS's website while Figure 5 shows the metadata associated with that data. The metadata displayed in Figure 5 is a combination of depositor created metadata and ADS enhanced metadata. For example, the ADS does not require depositors to list file sizes as a metadata requirement; that is generated at the time of deposit. If the archivist working on the dataset sees an opportunity to enhance the metadata then such edits may also be included within the info page (Figure 5) in line with the ADS's Assessment and Appraisal Policy (Evans and Green 2023). Most often the metadata the ADS receives is in one or more spreadsheet files.



Title: P7231822 fence posts overview Filename: 1C22TBCAM_P7231822.jpg (2 MB)



Figure 4: A file available for download and an additional info button. Example from High Speed Two Ltd., INFRA Archaeology (n.d.) Data from an Archaeological Monitoring at Three Bears Cottage, South Heath Cutting, Buckinghamshire, 2021 (HS2 Phase One)

	Object ID	2188334									
	Object Title	P7231822 fence posts overview									
	File name	1C22TBCAM_P7231822.jpg									
	Format	IPG									
	Format Type	IDEC Eile Interchange Format	1.01								
	Pormat type	JEG File Interchange Format	JPEG File Interchange Format 1								
	Checksum	3e8160dda7ba481f8f98df5943b5f									
	Created	23-Jul-2021									
	Last Modified	05-Jul-2022									
	File size	2 Mb									
					Download						
	Data Type	Image									
	Title	P7231822 fence posts overview									
	Creator	INFRA Archaeology	INFRA Archaeology								
	Copyright	© HS2 Ltd.									
		LOCATION 3 Bears Cottage									
	Location	LOCATION									
		LOCATION	Buckinghamshire								
		LOCATION	Great Missenden								
	Coordinates	Latitude Longitude	51.723010 -0.711050								
	Subject	MONUMENT/OBJECT Excavation									
	Subject	MONUMENT/OBJECT Environment									
Period MIDAS Uncertain											
Ass	ociated Metadata:										
	Image metadata			CSV	1 Kb						

Figure 5: The file's info page showing some of the associated metadata held in the ADS's database. Example from an Archaeological Monitoring at Three Bears Cottage, South Heath Cutting, Buckinghamshire, 2021 (HS2 Phase One)

There are a number of templates available for depositors to provide their metadata to the ADS, including core and data type specific metadata templates. However, the level of detail that is added to these templates is variable. Some depositors feel that the minimum requirements from the ADS and HS2 are sufficient for their collection while others provide extremely rich metadata. To make it easier for depositors, the ADS developed a metadata mapping tool which allows for both ADS- and depositorcreated templates to be imported into the ADS's OMS (see Figure 6). In this example, the depositor's original fieldname column contains the metadata terms that are being mapped to the ADS's database fields. Example data line 1 shows what metadata is being inserted into the ADS's OMS and there is the opportunity for extra processing if necessary. Ideally, by allowing the depositors to choose the way that the metadata is deposited, depositors will be able to meet the ADS's requirements for metadata in a way that works for their workflows and encourage them to provide richer metadata where possible.

Original fieldname	Database field	Extra processing	Example data line 1			
Flename	Table: OM5_OBJECT_REPRESENTATION : Field: FILE_NAME v	ENTATION ; Field: FILE_NAME V None V				
Caption	Table: OM5_OBJECT : Field: TITLE v	None v	P7231822 fence posts overview			
Subject Keyword 1	Table: OMS_OBJECT_DC_SUBJECT : Field: SUBJECT v	None v	Excavation			
Subject Keyword 2	Table: OMS_OBJECT_DC_SUBJECT ; Field: SUBJECT v	None v	Environment			
Period Term 1	Table: OM5_OBJECT_DC_COVERAGE_PERIOD : Reid: PERIOD v	None v	Uncertain			
Creator Organisation	Table: OMS_OBJECT_ACTOR_LINK : Field: ORGANISATION	CREATOR ~	INFRA Archaeology			
Copyright Holder Organisation	Table: OMS_OBJECT_ACTOR_LINK : Field: ORGANISATION ~	COPYRIGHT	HS2 Ltd			
Location	Table: OMS_OBJECT_DC_COVERAGE_LOC ; Field: DESCRIPTION v	None	3 Dears Cottage, Great Missenden, Buckinghamshire, England			
Longitude	Table: OMS_OBJECT_DC_COVERAGE_COORD ; Field: EASTING v	Choose a coordinate type 🔍	-0.71105			
Latitude	Table: OMS_OBJECT_DC_COVERAGE_COORD : Field: NORTHING v	Choose a coordinate type 🛛 👻	51.72301			
Creation Date	Table: OMS_OBJECT : Field: CREATED_YEAR v	Choose a date format	23/07/2021			
Software	Table: OMS_OBJECT_REPRESENTATION ; Field: SOFTWARE	None	Olympus Tough Camera			

Figure 6: The ADS's metadata mapping tool with data showing how the metadata from Figure 5 was imported into the OMS. Figure from an Archaeological Monitoring at Three Bears Cottage, South Heath Cutting, Buckinghamshire, 2021 (HS2 Phase One)

4.2 Metadata comparison

Prior to importing metadata into the OMS, the metadata needs to be checked against the data submitted within an archive. In most instances, all the metadata is listed within one or more spreadsheets. For larger, more complex collections that may have thousands of files within hundreds of folders, the sheer number of files can make it difficult to ensure that all the data has a corresponding metadata entry. To help aid with this check, Figure 7 shows the metadata comparison tool that was created to compare the files in a collection with the metadata submitted along with it. This tool allows for all files within a folder and its subfolders to be compared to a single metadata file. It produces a report detailing if any files are missing from the folder or metadata. From there, the archivist can provide the list to the depositor for them to correct. In time, this tool will be included in the ADS's ingest method so that these issues can be flagged as they are being deposited before an archivist begins to work on the collection.



Figure 7: A mock-up of what would be flagged using the metadata comparison tool with data. Figure used data from High Speed Two Ltd., MOLA Headland Infrastructure, Headland Archaeology Ltd (n.d.) Data from a Trial Trenching at Park Street Burial Ground, now Park Street Gardens, Birmingham, West Midlands, 2017-2018 (HS2 Phase One)

4.3 Deposit appraisal

Building on the metadata comparison tool, the deposit appraisal tool checks the files themselves. Instead of requiring an archivist to check each individual file that has been deposited during the accessioning process, this tool runs all the files through a checking system that compares the file names to the ADS's requirements for file names as listed in the ADS Ingest Manual (Archaeology Data Service 2022). The tool then provides an at-a-glance list of the files and any renames that may be required, as shown in Figure 8, as well as a detailed output file. Figure 9 shows an example output file which includes: file counts, file formats, directories, duplicate files, empty files (files that the ADS received but do not actually contain any data within them, a form of data corruption), and a breakdown of file naming issues. This comparison tool performs initial checks of the data to bring attention to potential issues. The time saved, compared to checking by hand, is immense. The HS2 archive shown in Figure 9 contains over 139,000 files equalling more than 3TB of data and would have taken days to check by hand, but the initial checks took minutes. While currently only available to the ADS's archivists, in the future this tool will be integrated into the ADS's deposit system for potential depositors to use.





File: Submitted		New file names	
IC195MAEX C22020 apf 1374 photogrammetry photogram Mig. 7800,1956 8 Mig. 7800,1956 8 Mig. 7805,1956 9 Mig. 7805,1956 9 IC195MAEX (2019), and the standard decision also 0 Mig. Societion bri 1 IC195MAEX (2020), pr. 1374, Calibration, pr. 9 IC195MAEX (2020), pr. 1374, Calibration, pr. 9 IC195MAEX (2020), pr. 1374, Genero 10 IC195MAEX (2020), pr. 1374, Genero 10 IC195MAEX (2020), pr. 1374, Shotogrammetry, mod IC195MAEX (2020), pr. 1374, Shotogrammetry, mod IC195MAEX (2020), pr. 1374, generosin	peoply metadota.cov 7 g, reports pat 8 el_metadota.cov 12 photo_metadota.cov 13 gg, metadota.cov 14	IC195MAE(, C22020, ypl, 1374, photogrammetry, pl M6, 7383, gp 6 M6, 7383, gp 5 M6, 7383, gp 5 M6, 7383, gp 1 M6, 7383, gp 1 M6, 7383, gp 2 IC195MAEX, spraidsheet, metadata, decision also 0 M8, bocston bit 1 IC195MAEX, C21020, spl, 1374, Calamaton.cs; 9 IC195MAEX, C22020, spl, 1374, GCPs.cs; 11 IC195MAEX, C22020, spl, 1374, photogrammetry, pl IC195MAEX, C22020, spl, 1374, photogrammetry, pl	histography_metadata.cov 7 sing_reports.pdf 8 codel_metadata.cov 12 thisphoto_metadata.cov 13 ttings_metadata.cov 14
		Dist de	and a second

Figure 8: Mock up of what would be flagged using the deposit appraisal tool. Figure from an Archaeological Recording at St Mary's Church, Stoke Mandeville, Aylesbury, Buckinghamshire, 2020-2022 (HS2 Phase One)

Date: Input Directory: Output File: Directories: Files: Directories and Files: Total Filesize:	10/01/2023 (14:34) A:/HS2 Data/Central/MOLA/1C19SMAEX A:/HS2 Data/Central/MOLA/1C19SMAEX\output_2023_01_10.log 15,091 139,316 154,407 3,287,909,033,190 bytes											
Extensions:	csv: jpeg: tif: cpg:	19,042 77,820 1,672 11	x1sx: JPG: tiff: db:	9 9,728 4,690 11	docx: jpg: tfw: dbf:	1 253 4 11	txt: png: xml: prj:	4 6,743 4,770 11	pdf: mtl: qmd: qml:	2439 2,368 10 11	MP4: obj: shx: shp:	2 2,378 11 11
Spaces (includes folders A:/HS2 Data/Central/HOLA) - 5 /1C195M	AEX\1C19	SMAEX_Fe	male_h	ad.jpg							
Uppercase JPG = 7646 A:/H52 Data/Central/HOLA/ICI95MAEX\ICI95MAEX_Site_Overall.JPG 												
<pre>Multiple Extensions (or Full Stop) = 1671 A:/HS2 Data/Central/MOLA/1C195MAEX\Raster\Archaeological_drawings\plan_1C195MAEX_121.56.tif </pre>												
Commas = 1 A:/HS2 Data/Central/HOLA/1C195MAEX\Raster\Site_photos\photo_1C195MAEX_5,672.3PG												
Parenthesis = 1 A:/H52 Data/Central/HOLA/1C195MAEX\Raster\Archaeological_drawings\plan_1C195MAEX_528(2)_modified.tif.points												
Duplicate Files = 3403 A:/H52 Data/Central/MOLA/IC195MAEX\IC195MAEX_plate1.JPG A:/H52 Data/Central/MOLA/IC195MAEX\hs2lparc-509737\Raster_data\site_photos\IC195MAEX_plate1.JPG 												
Empty Files = 8 A:/HS2 Data/Central/HOLA 	/101954	AEX\hs21	parc-509	737\Rast	er_data\	site	_photos\	photo_10	195MAEX	2665.JPG		

Figure 9: Mock-up of an output file using the deposit appraisal. Only one example file for each issue is shown. Figure from an Archaeological Recording at St Mary's Church, Stoke Mandeville, Aylesbury, Buckinghamshire, 2020-2022 (HS2 Phase One)

5. Conclusions

The High Speed 2 project chose the ADS to help create a legacy of highly accessible and outstanding archives of the historic environment. To make that a reality, the ADS has focussed on metadata deposition to ensure the data it receives from HS2 can be disseminated in an interesting and useful way to everyone. The ADS has created multiple tools to aid in depositing and accessioning both data and metadata. These tools allow the depositor freedom to choose the metadata template that works best for them and check that the metadata and data match and have no errors. Once the metadata has been added to the ADS's OMS, the ADS can in turn create crosssearches, interactive maps, curated collections, integrated publications, and more.

From the very beginning, HS2 set out with the intention of creating archives that people would seek out and actively re-use. The ADS's goal as a digital repository is to support access to, innovation of, and research on the archives it holds. The archaeology and heritage data produced as a part of HS2 has always been intended for the public's benefit. Each archive is created to aid in that goal and as Julian Richards, the director of the ADS said, "The single most useful thing you can do to ensure the long-term preservation of your data is to plan for it to be re-used." That all begins at the beginning.

Bibliography

Archaeology Data Service 2022 ADS Ingest Manual, Archaeology Data Service. <u>https://archaeologydataservice.ac.uk/app/uploads/2023/02/ADS_ingest_ma</u> <u>nual_v7.pdf</u> [Last accessed: 31 May 2023]

Archaeology Data Service 2023 *Data requirements table*. *Instructions for Depositors*. <u>https://archaeologydataservice.ac.uk/help-guidance/instructions-for-depositors/files-and-metadata/</u>[Last accessed: 31 May 2023].

Evans, T. and Green, K. 2022 ADS Assessment and Appraisal Policy, Archaeology Data Service. <u>https://archaeologydataservice.ac.uk/app/uploads/2023/02/assessment_an_d_appraisal_v1-4.pdf</u> [Last accessed: 31 May 2023]

Green, K., Niven, K. and Field, G. 2016 'Migrating 2 and 3D Datasets: Preserving AutoCAD at the Archaeology Data Service', *ISPRS International Journal of Geo-Information* **5**(4). <u>https://doi.org/10.3390/ijgi5040044</u>

High Speed Two Ltd., Archaeological Research Services Ltd 2023 Data from an Archaeological Recording at Deserted Medieval Village of Lower Radbourne, Warwickshire, 2021-2022 (HS2 Phase One) [data-set], York: Archaeology Data Service [distributor] <u>https://archaeologydataservice.ac.uk/archives/view/1005014/</u>

High Speed Two Ltd., Wardell Armstrong Archaeology and Cultural Heritage, Archaeology Wales 2023 *Data from a Trial Trenching at Newyears Green, Hillingdon,* 2019 (HS2 Phase One) [data-set], York: Archaeology Data Service [distributor] <u>https://doi.org/10.5284/1106788</u>

HS2 2017 Historic Environment Research and Delivery Strategy: Phase One. <u>https://assets.publishing.service.gov.uk/government/uploads/system/uploads/a</u>





ttachment_data/file/642655/hs2_phase_one_historic_environment_research_and_de livery_strategy.pdf [Last accessed: 31 May 2023].

International Organization for Standardization [ISO] 2019 ISO 15836-2:2019 Information and documentation - The Dublin Core metadata element set - Part 2: DCMI Properties and classes. <u>https://www.iso.org/standard/71341.html</u> [Last accessed: 31 May 2023].

Niven, K. and Watts, J. 2011 Project metadata. Guides to Good Practice. <u>http://guides.archaeologydataservice.ac.uk/</u> [Last accessed: 31 May 2023].

O'Brien, J. and Evans, T. 2022 ADS Cataloguing Policy, Archaeology Data Service. <u>https://archaeologydataservice.ac.uk/app/uploads/2023/04/ads_cataloguin</u> <u>g_policy_v3-0.pdf</u> [Last accessed: 31 May 2023]

Richards, J.D. 1997 'Preservation and re-use of digital data: the role of the Archaeology Data Service', *Antiquity* **71**(274), 1057-1059. <u>https://doi.org/10.1017/S0003598X00086014</u>

Solly M. 2018 10,000 Years of British History to Be Unearthed in Excavations in Advance of Planned Rail Line. <u>https://www.smithsonianmag.com/smart-news/excavation-planned-rail-line-yields-trove-archaeological-treasures-spanning-10000-years-british-history-180970739/</u> [Last accessed: 31 May 2023]