

<b>Project Name:</b>	ForSEADiscovery
<b>Author(s):</b>	Louisa Matthews, Kieron Niven
<b>Period covered:</b>	31 <sup>st</sup> January – 30 <sup>th</sup> March 2018
<b>Origination Date:</b>	28 <sup>th</sup> March 2018
<b>Reviser(s):</b>	-
<b>Derivation:</b>	Project Proposal: 6482_MEDIN_DAC_Application_ADS_Final.pdf
<b>Date of last revision:</b>	17/07/2018
<b>Version:</b>	V1.1
<b>Status:</b>	Agreed
<b>Circulation:</b>	MEDIN
<b>Required Action:</b>	Publication to MEDIN website
<b>File name/Location:</b>	6482_MEDIN_DAC_Project_Report.docx
<b>Approval:</b>	

## 1 Summary

MEDIN funding in 2017/18 enabled ADS to address some long-standing issues with direct access to data, specifically access to marine photogrammetry and video (dive tapes). These data types have typically been difficult to disseminate online due to their size and format. UK data from the [ForSEADiscovery project](#) (which included monitoring records from the Yarmouth Roads protected wreck) was archived and used to investigate options for dissemination and then pilot the preferred options. These comprised a 3D viewer and embedded video so that users can preview datasets before deciding if download is appropriate.

## 2 Introduction

This report summarises the work conducted as part of a data archiving project for the ForSEADiscovery Project. The ForSEADiscovery project (<http://forseadiscovery.eu/>) undertook research on historic shipwreck sites in 2015 and 2016 as part of multi-disciplinary research into timber supply for Iberian shipbuilding during the Age of Discovery. The archiving project has preserved and will make available diverse datasets from dive operations data from commercial companies MA Ltd (UK), and Archeonautica (Galicia, Spain). The data comprised:

- 41 videos of dive footage or topsides diver briefing footage
- 17 photogrammetric models of archaeological wrecks and sites along with primary photographic data (as per ADS [Guide to Good Practice](#))
- Monitoring data from UK designated historic wreck (Yarmouth Roads).

The project covered both UK territorial waters and Spanish territorial waters. MEDIN funding covered the work on the UK component of the material that comprised digital data from underwater fieldwork on the Yarmouth Roads shipwreck site off the North coast of the Isle of Wight. The site is designated under the 1973 Preservation of Wrecks Act as the remains of a late 16th or early 17th century carrack, possibly of Spanish origin, which stranded in Yarmouth Roads, Isle of Wight. The site has been monitored over many years by the Hampshire and Wight Trust for Maritime Archaeology / Maritime Archaeology Trust (<http://www.maritimearchaeologytrust.org/yarmouth>).

## 3 Methods

The project addressed some long-standing issues in relation to improving direct access to data on the ADS website (and via the MEDIN portal).

On receipt of the data from the project, the ADS archivist ingested the material into the ADS preservation systems. Priority was given to the audio-visual and photogrammetry data that were key to the project outcomes. The following is a summary of the assessment and decisions made on the best way of ensuring increased access to the data described in the project design, namely marine photogrammetry and audio-visual material in the form of dive tapes and top-side footage.

### 3.1 Audio-visual Material

Previous ADS experience with disseminating video (currently 120 files over 14 collections) has been largely limited to providing the files as simple downloads. As video files are often large and not easily downloadable, such data, along with other files exceeding a 100Mb limit, are currently compressed to multiple volume ZIP files in order to allow downloading in discreet blocks. While this approach can make direct and easy access difficult, with users potentially spending time downloading material that ultimately may not be relevant to them, the method remains a sustainable and robust approach to basic data dissemination. However, in order to improve access and assessment of such videos, enhancing the current approach by enabling video to be played directly on the ADS website would provide a better way to access data.

The ADS, via Internet Archaeology<sup>1</sup>, has had limited experience providing streaming video. Previously provision was via the University of York video streaming service, which as will be seen below, is not a viable long-term option.

### 3.1.1 Options

#### 3.1.1.1 Downloads

As with all ADS archives, it is necessary to continue with the current practice of providing basic download access to 'full resolution' data. In the case of audio-visual data this involves the dissemination of uncompressed MP4 files under 100Mb and to ZIP compressed multiple volume video files over 100Mb.

#### 3.1.1.2 Online Video

Initial enquiries to the University of York IT Service indicated that the existing streaming video service<sup>2</sup> is still available although a number of issues were highlighted. Firstly the service is maintained at a departmental level (shared with the larger Archaeology Department) and therefore unlikely to have the capacity to host ADS datasets in the long term. Additional concerns were highlighted that the centrally-managed service may also pose issues of compatibility between the ADS website and the university's Web CMS. Finally, and most significantly, it was also highlighted that the University of York is looking to discontinue the streaming service in the near future and recommended YouTube as a possible streaming alternative. The latter service has been investigated by the ADS in the past and, along with other and other third-party streaming services, been found to present significant problems in terms of intellectual property rights issues for many archive depositors. For example, Youtube asks for a 'worldwide, non-exclusive, royalty-free, transferable licence (with right to sub-licence) to use, reproduce, distribute, prepare derivative works of, display, and perform that Content...' which is not acceptable for many data owners. Additionally, there are no guarantees as to the longevity of hosting, or permanency of urls, which can cause significant issues in the long term if content is deleted or moved, and makes these types of services unsuitable.

More widely, the proposed discontinuation of the University of York streaming service and recommendation of YouTube also points to the wider movement, resulting from security concerns around Flash, to move away from flash-based video towards newer HTML5 based implementations. In line with this wider trend, and the options available to the ADS within the university IT provision and via third-party providers, the decision was made to implement online playable video via the W3C standard HTML5 video on the ForSEADiscovery archive pages alongside the standard downloadable MP4 files.

### 3.1.2 Implementation

The implementation of HTML5 on the ADS archive pages has been via the inclusion of the HTML5 video tag<sup>3</sup> within archive pages. Limited customisation and styling has been undertaken to ensure that the display of video files fits into the existing ADS website style and layout.

In line with existing ADS guidelines for downloadable files, files for online playing have been reduced, as far as possible, to within the 100Mb file size limit. This allows for files to be easily played and assessed within the archive web pages and, if required, the full resolution version can then be downloaded where required. To fit within this size limit, large video files have had their frame size reduced using the free Handbrake software<sup>4</sup>. In the case of the ForSEADiscovery video files this involved reducing frame size to around ¼ of the original whilst retaining the original aspect ratio (e.g. 1920x1080 original files were reduced to 480x270). The MPEG 4 video files were additionally assessed for conformance with the H.264 video and AAC audio codecs used by HTML5 although no conversions were required for the ForSEADiscovery datasets.

## 3.2 Photogrammetry

ADS mechanisms for disseminating large datasets resulting from photogrammetric surveys and Structure from Motion (SfM) modelling have largely been limited to providing simple data downloads. Such datasets usually consist of a large 3D models (e.g. an OBJ file and associated textures, usually compressed into single or multiple volume ZIP file) along with

<sup>1</sup> T. Williams 2008 'The landscapes of Islamic Merv, Turkmenistan: Where to draw the line?', Internet Archaeology 25.  
<https://doi.org/10.11141/ia.25.1>

<sup>2</sup> <https://www.york.ac.uk/it-services/services/streaming/>

<sup>3</sup> [https://www.w3schools.com/tags/tag\\_video.asp](https://www.w3schools.com/tags/tag_video.asp)

<sup>4</sup> <https://handbrake.fr/>

anywhere between thirty to several hundred source raster images from which the model has been created. As with audio visual data, a similar issue exists in that it is often difficult for users to discern the utility of a large 3D model before downloading it. Additionally there is the issue of being able to access large quantities of source photography at a suitable level of granularity, one which balances the need to access with relative ease specific images against also being able to download large volumes of files for potential reuse and reprocessing.

The ADS has previously investigated tools for the previewing of this material within the scope of specific large-scale research-funded projects however it has not been possible to implement consistent use of these tools across all relevant archives as costs have proven prohibitive. This project has allowed the ADS to investigate the potential for a more efficient implementation of such tools whilst also assessing updated functionality and ensuring that adequate documentation is created to allow ongoing use.

### 3.2.1 Options

While 3D data is now widely available on the internet it is difficult to find cases where entire 3D datasets (including both the raw image data and final 3D models) are disseminated online. The MEDIN Portal appears to contain no information on such datasets with '3D' or 'photogrammetry' data types absent from the advanced search options.

### 3.2.2 3D Models

No current provision for the dissemination of 3D models exists within the University of York and third-party options are largely focussed upon services such as Sketchfab. As with YouTube, while Sketchfab claims no ownership rights over user content it does claim a "perpetual, irrevocable, sub-licensable (through multiple tiers) right and license to use, reproduce, publicly display, distribute and adapt" user content which, again as with YouTube, the ADS feels would be unacceptable for many data depositors. In light of this, and of existing experience and implementation, it was decided to revisit the open source 3DHOP software package for the online display of 3D models. 3DHOP had been utilised previously by the ADS on stand-alone projects, but had relied on an outside specialist to implement and advise on installation and use. Since then, application has been upgraded, and its implementation has become more streamlined and easier to document. This means that the tool is now more cost-effective to utilise than before, and therefore more sustainable. It was decided to investigate the use of 3D hop for dissemination of the ForSEADiscovery data.

### 3.2.3 Images

Many archaeological project archives held by the ADS contain images and these are disseminated in a standardised gallery format (individual image with gallery thumbnail and low resolution preview). This approach has been previously adopted unchanged for the AHRC funded ACCORD archives and while it provides excellent options in terms of being able to assess individual images (via thumbnails) and to download single files it works less well where access to hundreds of files is required. A more recent ADS photogrammetric archive (the Crystal Palace Iguanodon condition survey) has adopted an almost opposite approach where large sections of the project's image collection have been packaged together into multiple volume ZIP files. While this allows relatively easy access (fewer 'clicks') to the whole image set it also requires all ZIP volume files to be downloaded even if just one image file is required. A logical solution to these issues is to create a 'middle option' allowing images to be stored together in independent groups within a logical structure.

### 3.2.4 Implementation

#### 3.2.4.1 3DHOP

The implementation of 3DHOP required a detailed assessment of the previous ADS use cases along with an assessment of new features and options available in the current version of the software (version 4.1). As with the implementation of HTML5 video, limited customisation and styling was performed in order to fit the software in with the current ADS website. In contrast with previous implementations it was decided to use an expanded toolset version of the software to allow more onscreen manipulation of models without the need to download either the file or more complex software. The newer, updated version of the viewer also featured greater browser compatibility than previous versions alongside better handling of models.

#### 3.2.4.2 Images

The 'middle option' for disseminating photogrammetry images required the straightforward compression of image files into sub-100Mb batches. This differs from the previous multiple volume ZIP approach in that that each ZIP batch is

independent. Additionally, naming these batches using image file name ranges has helped identify what batch is relevant, and is aided by project metadata files describing the images. The resulting download pages for each photogrammetry model allows the 3DHOP model, the OBJ file, and all image files to be presented together within a relatively compact space.

## 4 Conclusion

The ForSEADiscovery project data has been ingested and preserved in the ADS systems. At present the AV and Photogrammetry data are presented in a draft interface (used to test the streaming and 3D interfaces) which can be found at [http://archaeologydataservice.ac.uk/archives/view/forseadiscovery\\_2018/](http://archaeologydataservice.ac.uk/archives/view/forseadiscovery_2018/).

The collection will be formally released in due course and may be referenced as follows:

Nigel Nayling, Garry Momber, Miguel San Claudio, Ana Crespo Solana (2018) **ForSEADiscovery** [data-set]. York: Archaeology Data Service [distributor] <https://doi.org/10.5284/1048329>