TOWARDS
A NATIONAL
COLLECTION

UKRI Arts and
Humanities
Research Council

FIRST REPORT
DISCOVERY PROJECTS

Unpath'd Waters:
Marine and Maritime
Collections in the UK

OCTOBER 2022

Table of Content

Authors

Barney Sloane (Historic England), Julian Richards (Archaeology Data Service, University of York), Fraser Sturt (Centre for Maritime Archaeology, University of Southampton), Ann Coats (University of Portsmouth), Michael Roberts (Centre for Applied Marine Sciences, Bangor University), Vince Gaffney (School of Archaeological and Forensic Sciences, University of Bradford), Stuart Jeffrey (School of Simulation and Visualisation, Glasgow School of Art), Sara Perry (Museum of London Archaeology)

Executive Summary

The UK has a rich maritime heritage, stretching back over 23,000 years. It is impossible to tell the story of our islands without talking about our relationship with the sea. This maritime past is becoming increasingly important. People are more aware of our exploitation of the sea and topics such as colonialism, slavery and immigration.

Unpath'd Waters therefore aims to increase interaction with the UK's maritime heritage by making it easier to research and easier for the public to discover and share stories in new ways. Despite its importance, it is not always easy to study our maritime heritage. Records, maps and objects are scattered across hundreds of different archives, museums, libraries and galleries. A large part of our work will be to develop new ways of making information across these different collections easy to search and find.

This will help everyone – from researchers asking new questions to members of the public having direct access to records. We hope this will encourage more experts from all disciplines to use maritime collections in their own work. To make sure this project has a lasting impact, we will publish all our methods, code and research so anyone can use it in their work and help the future of UK maritime heritage.

Unpath'd Waters will tackle the challenge through eight key activities (we have called them Work Packages). The first will aggregate and assess the character of more than 90 different maritime collections, matching their core attributes and creating a metadata framework which will allow them to be connected. The second will test artificial intelligence and machine learning opportunities to help search these connected collections in new ways. Three connected research activities will then test this framework and these new tools to try to resolve three real-world challenges. *People and the Sea* is looking at how people value wrecks in museums (the *Mary Rose* and the *Holland 1*) compared to some wrecks still on the seabed of the English Channel which have been surveyed and even excavated in part. *Science and the Sea* is using the connected collections to find the identity of wrecks in the Irish Sea and establish their stories and their management needs. *Lands Beneath the Sea* is using collections to build a simulation of now-lost landscapes under the North Sea where prehistoric peoples once dwelled before sea level rise.

Using the outcomes of this research and the collections framework and exploration tools, the project will then reach out to some key sample audiences: cross-disciplinary researchers, visually impaired people, and those who live far from the sea and have little connection with it. These audiences will help us co-design new ways of accessing this connected information so that we can create new portals for the public which are designed around what people themselves want rather than what we think they want. On top of this, we expect to be able to enhance the national inventories of all four home nations – England, Scotland, Wales and Northern Ireland.

Finally, we will prepare a report with the lessons we have learned and recommendations on what works best to help the UK move towards a genuinely National Collection.

Unpath'd Waters is a big project and has a large number of people working on it. It is led by Barney Sloane of Historic England aided by a programme manager and administrative support. A Management Group of experts in digital humanities, marine archaeology and history, ocean science and audience engagement and evaluation leads a consortium of eight universities, three Independent Research Organisations and five collaborating organisations including archaeological trusts, museums and charities.

The consortium is working well and the research programme is proceeding as planned. It is rather too early to claim any major successes, but we have created the necessary metadata framework to connect all the collections in our sample together and are now using that framework to start exploring what those connections can reveal. In creating the framework, we have already made advances in the international 'vocabularies' used to define heritage assets (wrecks, archaeological sites etc), and periods of time. This is vital because without common reference points it is impossible to link collections properly.

We have also developed our project Values Framework – a crucial ethical consideration in establishing genuine representation within the project consortium and with our public audiences whose support will be critical as we move forward with the project.

We have also published a project website www.unpathdwaters.org.uk where we will provide updates on the work of the project.

The products of Unpath'd Waters will be held securely and sustainably beyond the lifetime of the project and are being developed to accord with the FAIR principles, so as well as informing the strategy for Towards a National Collection the products of the project will outlast it and be available for others to use.

Abstract

The UK Marine Area covers 867,400 km2, 3.5 times the terrestrial extent. Our marine heritage is extraordinary. Shipwrecks from the Bronze Age to the World Wars bear testimony to Britain as an island nation, a destination for trade and conquest and, in the past, the heart of a global empire. Coastal communities have been shaped by their maritime heritage with stories of loss and heroism. Deeper in time, what is now the North Sea was dry land, peopled by prehistoric communities. Our current land would have been distant uplands above hills and plains and rivers now lost and forgotten.

Numerous collections represent this heritage, covering 23,000 years, including charts, documents, images, film, oral histories, sonar surveys, seismic data, bathymetry, archaeological investigations, artefacts, artworks and palaeoenvironmental cores. These are unconnected and inaccessible. This matter because the story of our seas is of huge interest to the UK public, with millions visiting maritime museums annually, and marine exploitation increasing dramatically for energy, minerals, trade, food and leisure.

To unlock new stories and effect sustainable management, we must join up our marine collections. Unpath'd Waters brings together universities, agencies, museums, trusts and experts to confront this challenge. AI is being applied to innovate searching across collections, simulations to visualise landscapes, and science to identify wrecks and research their artefacts. Unpath'd Waters will deliver management tools to protect our most significant heritage and invite the public to co-design new ways of interacting with the collections. The methods, code and resources created will be published openly so they can used to shape the future of UK marine heritage.

Aims & Objectives

Unpath'd Waters is split into Work Packages (WPs). These are designed to run concurrently and reflexively, informing and being informed by each other, forming a coherent research arc aiming to characterise and organise the collections, resolve obstacles to cross-searching, apply these tools iteratively to real-world marine challenges and develop mechanisms to engage the public in co-designing new access routes.

Throughout this report we refer to the project as UNPATH.

WP1 and 2 are addressing the more technical challenges of connection, access and searching, WP3.1 – 3.3 represent 'test lenses' to work through real-world research demonstrators, and WP4 and 5 are focused on visualisation, audience engagement and evaluation. WP6 is the underpinning programme management work. The five research WPs have the following research objectives.

WP1 Aggregation and Characterisation: *How can we integrate the UK's marine heritage collections?* WP1's objectives are to assess, enhance and aggregate the digital components of the UK's marine heritage collections; review the demands and capabilities of existing digital infrastructures; and enhance these to deliver the UK marine heritage collections to researchers and the public via human and machine-readable interfaces. It will develop appropriate data standards and cataloguing practices to achieve interoperability and deliver sustainable access routes and a primary data archive for UNPATH data outputs. This will lead directly to major improvements in national marine heritage inventories, including Coflein (Wales), Mariner (England), Canmore (Scotland) and, in due course inventories for Northern Ireland and, beyond UK waters, the Isle of Man.

WP2 Discovery: *How can we transform our ability to search those collections?* WP2's objectives are: to harness recent developments in artificial Intelligence and machine learning and build on work done in the Towards a National Collection Heritage Connector Foundation Project, to enrich sparse data, increase interoperability and enable a diversity of search options. It will do this through adopting a low-shot/zero-shot approach to machine learning leveraging the assets created in WP1 via named entity recognition to open linked data. In addition, computer vision (object detection) and increased spatial/temporal awareness/visualisation will be explored within the cases studies detailed in WP3. This will enable distinct modes of discovery and recovery (text, image, 3D shape, sound, spatial, temporal).

WP3.1 People and the Sea: *How can we enhance the significance of submerged and displayed wrecks?* WP3.1 will use a sample of submerged Protected Wrecks (and nearby unprotected wrecks) in the English Channel (The Needles, Mary Rose and Holland 5) and displayed wrecks (Mary Rose and Holland 1) by employing new stakeholder processes for elevating the significance of underwater wrecks in relation to displayed wrecks. Approaches will test the capacity of connected collections, including wreck site surveys, recovered artefacts, documentary records and scientific samples, to engage audiences in co-creating new narratives and innovative ways of engaging with the wrecks. WP3.1 aims to evaluate these novel approaches in enhancing public availability of wreck databases, serving as proof of concept for accessing other wreck sites and developing strategies for Protected and unprotected wrecks.

WP3.2 Science and the Sea: *How can we use collections to identify located wrecks and improve their management?* Using an Irish Sea pilot, WP3.2 aims to test connectivity between disparate marine scientific collections to identify wrecks and enable more effective heritage management of shipwrecks through better understanding of their condition and integrity. WP3.2 aims to link a collection of high-resolution multibeam

sonar surveys from >100 incorrectly identified or unknown wreck sites in the Irish Sea to other collections (shipping and naval archives, vessel plans, insurance records, personal correspondence and photographs etc), to transform the value of heritage assets whose research access is limited typically to vessel names and location of loss data. WP3.2 will also test integration of site-related geophysical and hydrodynamical/sediment transport datasets and marine ecological data to enable site-specific assessments on localised site conditions, integrity and sustainability of the national collection of shipwrecks.

WP3.3 Lands beneath the Sea: *How can we open access to complex collections representing submerged ancient landscapes?* WP3.3 will pioneer a method of accessing complex digital scientific collections relating to the underlying processes that created the archives representing the inundated prehistoric North Sea landscape. These processes link users, data and unifying concepts between different datasets, including evidence for climate change and rising sea levels. As process-led access models are largely untried, a bespoke computer simulation will be created to allow users to experience the dramatically changing world which created the data collections, providing a more comprehensible, and interactive method of engaging with them. WP3.2 aims to enhance researcher and public access to an otherwise entirely occluded prehistoric landscape and to support far more accurate predictive models of likely archaeological potential in an area set to see massive exploitation for renewables.

WP4. Designing, Connecting, Immersing: *How can we co-design and co-create research and engagement access for new audiences?* WP4 will address issues of diversifying audiences, improving cross-collection and cross-disciplinary research by co-designing novel interfaces with integrated maritime datasets through the use of immersive systems, through three user-group case studies (see WP5). Objectives are to test: (1) What new affordances do integrated maritime datasets offer cross-disciplinary researchers? (2) Can modes of navigation be co-designed that counter the complexity of multiple datasets, from multiple sources and in multiple formats (2D/3D) which work for both research and public audiences? (3) Can we reach audiences that do not traditionally engage with these data but could directly participate in the design of access modes that encourage their future engagement? (4) Can the technical challenges posed by User Generated Pathways through the UNPATH datasets be meaningfully addressed via multi-modal immersive interfaces (i.e. delivered as full immersive systems, but also via e.g. desktop and mobile devices)?

WP5. Engaging and Evaluating Audiences: *How can we ensure that the audiences are truly representational, empowered and their inputs formally evaluated?* WP5 will underpin WPs 3 and 4 by aiming to identify, engage with, and evaluate impacts on users. WP5 aims to: (1) discover shared values for the project (and to guide audience-focused work) in line with the [Digital Culture Charter](), and develop metrics to assess project accountability to the values; (2) conduct audience mapping, to specifically identify and articulate a programme to ensure representation of three key target user-groups: non-coastal (inland) communities who do not identify with the sea; visually-impaired publics; and cross-disciplinary natural science and cultural researchers; (3) facilitate co-design with these audiences (qv WP4) and assess the efficacy and inclusiveness of the co-design approach; (4) facilitate testing and evaluating of the pilot designs that emerge through the interaction of WPs3 & 4 with our target audiences; and (5) coordinate and evaluate an exhibition programme of the final designs, including enrolment of other museums partners and industrial and third sector collaborating organisations into the programme to connect with their existing initiatives.

Partnership structure

The partnership structure for Unpath'd Waters is set out in Fig 1 schematically.

Historic England (Barney Sloane) is the Principal Investigator (PI), supported directly by a Programme Manager (TaNC-funded) and an Administrator (HE funded). An external Advisory Board provides the PI with an independent sounding board.

The PI chairs a Management Group made up of the lead Co-Investigators from each Work Package (WP), the Programme Manager, Administrator and the Communications lead (HE-funded). This Group monitors progress, addresses cross-project issues and provides reports down to each WP.
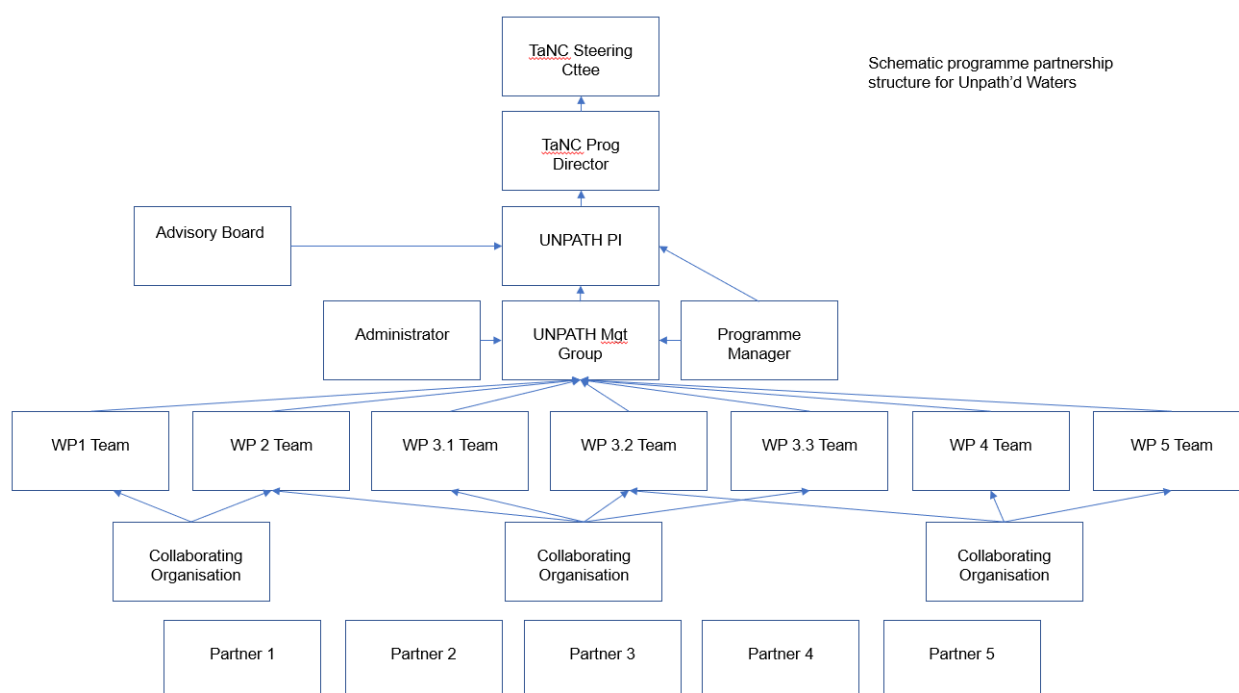

Fig 1: Schematic programme partnership structure for Unpath'd Waters

Work Package participants are:

- WP1: Lead: University of York (Archaeology Data Service). Co-Is: MEDIN, Historic Environment Scotland, Historic England; Collaborating Orgs (CO): Wessex Archaeology
- WP2: Lead: University of Southampton (UoS). Co-Is: Historic England. COs: RCAHMW
- WP3.1: Lead: University of Portsmouth (UoP). Co-Is: MOLA. COs: Mary Rose Trust, Nautical Archaeology Society, Maritime Archaeology Trust.
- WP3.2: Lead: Bangor University (BU). Co-Is: University of Ulster (UoU); COs: RCAHMW
- WP3.3: Lead: University of Bradford (UoBr). Co-Is: Historic England
- WP4: Lead: Glasgow School of Art (GSA). Co-Is: MOLA, University of St Andrews (UStA); Cos: Wessex Archaeology and Maritime Archaeology Trust
- WP5: Lead: MOLA. Co-Is: National Maritime Museum, Glasgow School of Art
- WP6: Lead: Historic England (HE).

Partner organisations are:

- Cadw (Welsh Government)
- Historic Environment Division, Department of Communities Northern Ireland
- Manx Heritage, Isle of Man's heritage department
- Marine Management Organisation
- Lloyd's Register Foundation
- Protected Wrecks Association
- British Geological Survey

Staffing structure

WP1: ADS Co-I Julian Richards leads WP1. MEDIN Co-I Clare Postlethwaite leader of the MEDIN data-providing organisations, is facilitating data access, and ensuring compatibility with the MEDIN. HES Co-I Peter McKeague supports Scottish (Canmore) inventory input. HE Co-I Paul Jeffery supports English marine inventory input and oversees RA Frankie Lau on UNPATH/Mariner integration. ADS RA Holly Wright supports RA Jamie Geddes on the evolving UNPATH ontology and metadata standards; and provides day-to-day management of WP1.

WP2 UoS marine and archaeological computing expert Co-I Fraser Sturt leads WP2, manages RA Jack Pink and links to CO input, and leads on spatiotemporal data discovery tools. UoS data science Co-I Adrienne Chapman leads on application and integration of AI methods for search and discovery.

WP3.1 UoP social/cultural maritime historian Co-I Ann Coats leads, coordinating wreck collections, managing the core UoP RAs, coordinating CO input. MOLA education/outreach Co-I Caroline Barrie-Smith oversees RA Sherman in the development and evaluation of co-creative public engagement and ensuring links to WPs 4 and 5. UoP RAs Garry Scarlett, Robert Inkpen, Tarek Teba, Karen McBride are creating visualisations of wreck sites and investigating access to collections. UoP RA Claire Bailey-Ross, a digital heritage engagement impact specialist, supports a programme of immersive/viewer resources. MOLA RA Andy Sherman, community archaeologist, supports development and evaluation of co-creative public engagement approaches.

WP3.2 BU Marine geoscientist Co-I Michael Roberts leads and manages WP3.2, coordinating wreck and environment input from UoU marine geoscientists Co-Is Wes Forsythe and Rory Quinn; supervising RA Innes McCartney and Kathy Blakey in linking collections; and coordinating CO input.

WP3.3 UoBr landscape archaeologist Co-I Vince Gaffney leads, overseeing UoBr RAs Phil Murgatroyd and Simon Fitch in integrating diverse data sets and related software development programmes required for modelling; contributes to the visualisation development in WP4; and ensures academic outputs. RA Murgatroyd is providing agent modelling/simulation expertise to design and develop software and supporting material to provide user access to the landscapes. RA Fitch, marine landscape modelling expert, is integrating required archaeological, geophysical, topographic and core collections/data sets.

WP4 GSA heritage data visualisation expert Co-I Stuart Jeffrey leads WP4, managing development and delivery of the UNPATH Explorer. He will also coordinate RAs (tbc) and COs and liaise with WPs 3.1 and 5. UStA community archaeologist Co-I Tom Dawson will support the co-design and evaluation documentation process and will also develop Explorer content including integration of maritime heritage AV content.

WP5 MOLA audience engagement and evaluation expert Co-I Sara Perry leads, providing a values framework and ensuring digital equity and access across UNPATH, and supporting Co-I Barrie-Smith (WP3.1) and Co-I Jeffrey (WP4); NMM museum/collections expert Co-I Andrew Choong will facilitate access to museum partners and their target audiences; coordinate final dissemination/ exhibitionary activities with these supervising RA Lawrence Northall, and support Co-I Perry in delivering other outputs of WP5.

WP6 HE heritage management specialist PI Barney Sloane is providing overall leadership (WP6) of the consortium, finances, programme and delivery; managing input of HE Co-Is Paul Jeffery (Partner coordination) and Antony Firth (marine strategy lead); and overseeing RAs Lau (UNPATH/Mariner interface);

Abi Morris, UNPATH programme manager; and RA(TBC), an evaluation specialist, who will help evaluate UNPATH using HE's Public Value Framework.

Collaborating organisation staff are: Wessex Archaeology (Graham Scott): Senior Maritime Technical Specialist providing WA datasets (WP1) and input on wrecks (WP3.1 and 5); Maritime Archaeology Trust (Julie Satchell and Brandon Mason): working on Needles wrecks (WP3.1) and audience engagement (WP5); Mary Rose Trust (Alex Hildred, Alastair Miles and Hannah Matthews) providing Mary Rose collections and audience access (WP3.1); Nautical Archaeology Society (Mark Beattie-Edwards) providing data and mapping for Holland 1 and 5 wrecks (WP3.1); Royal Commission on the Ancient and Historical Monuments of Wales (Julian Whitewright) providing access to Coflein datasets (WP1, 3.2).

Overall programme

The following table provides a summary of the project's programme

UNPATH WORK PACKAGE 1: Aggregation & characterisation	Start	Finish
Ontology Development	Nov 21	Nov 22
Data Auditing	Nov 21	Jun 23
Portal Development	Nov 21	Aug 24
Data Uploads	May 22	July 24
Standards and Good Practice Guide	Mar 23	Aug 24
Data Archiving	Sep 23	Aug 24
UNPATH WORK PACKAGE 2: Discovery	**Start**	**Finish**
User Needs / Audience Mapping	Nov 21	Apr 22
Development	May 22	Oct 22
Explore use of NLP, zero shot and object detection on national datasets and WP case studies	Oct 22	Sep 23
Embed extended search option and publish	Sep 23	Jul 24
UNPATH WORK PACKAGE 3.1: People & the sea	**Start**	**Finish**
Data collection, review and analysis of sites and study areas	Nov 21	Nov 22
Prototype development	Jul 22	Apr 23
Working with audiences	Dec 22	Oct 23
Public & schools' engagement	Mar 23	Jul 24
Analysis and evaluation	May 23	Jul 24
UNPATH WORK PACKAGE 3.2: Science & the sea	**Start**	**Finish**
Identification of wreck sites, relevant scientific data sets, historical archives	Dec 21	Aug 22
Creation of metadata catalogues, Unpath science data toolkit & site ID methodology	May 22	Oct 22
Assessment of collection connectivity and methodology testing link data/archives/collections approaches development	Nov 22	Jul 24
Recommendations for candidate Protected Wrecks	Dec 23	May 24
Analysis and evaluation	Dec 23	Nov 24
UNPATH WORK PACKAGE 3.3: Landscapes beneath the sea	**Start**	**Finish**
Prototype terrain model	Nov 21	Nov 22
Climate-sensitive temporal model	Nov 22	Sep 23
Prototype integrated landscape model, tech testing and user engagement	Oct 23	Dec 23
Beta version of final product, tech testing and public engagement	Nov 23	Jun 24
Version 1.0 of software package	May 24	Sep 24
Version 1.1 of final product and design document/report	Oct 24	Nov 24
UNPATH WORK PACKAGE 4: Designing, Connecting, Immersing	**Start**	**Finish**
Define co-design methodology & recruitment	Oct 22	May 23

Audience engagement and co-design activities	May 23	Mar 24
Create Unpath Explorer technical workplan & technical framework	Nov 22	Sept 23
Curated Pathways co-design	Sep 23	Nov 23
UNPATH Explorer re-casting for multiple dissemination modes	Nov 23	Mar 24
Multi-modal output testing and evaluation	Mar 24	Sep 24
UNPATH WORK PACKAGE 5: Engaging and evaluating audiences	**Start**	**Finish**
Values Framework	Nov 21	Jun 22
Audience mapping	Apr 22	Dec 22
Co-design research & methodology development	Oct 22	Dec 22
Co-design sessions	Dec 22	Nov 23
Report on user analyses	Nov 23	Feb 24
Specify and test audience evaluation methodology	May 22	Dec 23
Roll out audience evaluations & report on outcomes	Dec 23	Oct 24
Agree outline exhibition plans	Oct 22	Mar 23
Exhibition development	Dec 22	May 24

Events & Consultations

Subject	Date	Number of attendees/responses
WP5 Values framework development		
Initial brainstorming workshop	8 March 2022	13
Consultation on draft values	Completed w/c 4 April 2022	13
Consultation with the Consortium on refined values	Completed 30 April 2022	76
WP5 Audience Mapping		
Consultation and data gathering with Unpath'd Waters Consortium	April - July 2022	76
Working with audiences		
WP3.1 Pilot student hackathon – task was to visualise biological research data, creating a visual immersive/interactive using genetic data from the Mary Rose	July 2022	15
Events		
Unpath'd Waters Consortium event	November 2022	76
Talks & Conferences:		
WP3.2 Science & the Sea (University of Bangor): public lecture by MJR at a Bangor U3A group meeting	25 March 2022	40
WP5 Audiences & Evaluation (MOLA): Transforming Collections Project Values Exploration Presentation, University of the Arts London	12 April 2022	50
WP3.1 People & the Sea (University of Portsmouth): Naval Dockyards Society Conference, National Museum of the Royal Navy Portsmouth: presentation by Antony Firth: Placing Warships: Reconnecting vessels and dockyards	9–11 June 2022	59 in-person; 11 online delegates
WP3.2 Science & the Sea (University of Bangor): presentations by MJR and JW at the Research Framework for the Archaeology of Wales mini conference	27 June 2022	40
WP1 Aggregation & Characterisation (ADS, University of York): Hackathon, Linked Pasts Conference, York	29 November -1 December 2022	25

Research approach

The research design of the project is linear overall: An Incubator Lab in the first third of the programme is discovering, aggregating and characterising the collections, developing search and access tools, and setting the research basis for the three marine themes. The project will move into an Innovation Lab, where we apply the tools in real-world scenarios to test feasibility, resolve barriers, then employ these outputs to explore genuine innovation in co-design and co-creation capability to permit consumption, enhancement and evolution of the connected collections through visualisation, immersion and exhibition engagement. Impact Delivery, the final phase, will complete the work's evaluation and development of the formal documentation and final reporting outputs.

The seven Work Packages are progressing through these three phases, with iterative connections to ensure that lessons learned within one WP can be transmitted quickly to other WPs, providing robust course-correction flexibility. This design also ensures that the project Research Associates, and in particular the Early Career Researchers, will benefit from a full spectrum viewpoint and see their roles within the context of the entire programme. Further, this design integrates insights from industrial, third sector and heritage management organisations at each step of the project to ensure real-world applicability.

WP1 has undertaken a partner data structure survey and survey of existing data infrastructures, feeding development of the UNPATH core metadata model and core data standards. The range of collections included in this survey (and the resulting ontology, of which an initial version is complete) is considerable, covering 94 different collections which range from records of artefacts from shipwrecks, through to seismic datasets from the oil and gas industry which reveal submerged landforms. WP1 will enhance existing portals with aggregation of UK marine data (national bodies – Coflein, Canmore, Heritage Gateway; ADS ArchSearch; MEDIN, ARIADNE); and build upon work done by ADS in the EU-funded VENUS project to provide up-to-date data standards guidance in the ADS Open Access world-renowned Guides to Good Practice. FAIR and CARE principles sit at the heart of the UNPATH approach (Co-I Richards and Co-I Perry were PI and Co-I respectively on the TaNC *Making it FAIR* project).

WP2 is trialling Natural Language Processing (NPL) and computer vision-based enhancement and segmentation of sparse data; and will undertake audience/stakeholder behaviour/engagement analysis (defining search and discovery ontologies); delivering enhancement of sparse data across core datasets arising from WP1, developing UNPATH search tools (text, image, sound, space, time) for application to enhanced collections, enabling 'free-association' search beyond core datasets, and testing machine learning to assist searches and linkages.

WP3.1 will deliver four strands. (i) Dissolve boundaries between museum-based Mary Rose and Holland 1 and the still-submerged Holland 5, Mary Rose and The Needles Protected Wrecks and adjacent non-protected wrecks, to explore how conflict or trade vessels connect individuals and coastal communities; relating to construction, provisions, ordnance, crews, health, rescue and display; (ii) develop digital connectors linking disparate wreck artefact and archival collections; (iii) explore potential applications of Holland 1/5 real-time immersive digital experiences through XR; (iv) map the cultural and social attributes of the wrecks with contemporary communities, values and debates. These will be used to develop innovative co-designed, co-created wreck audience engagement and learning tools and packages and prototype connectors, enabling a range of consumers (Protected Wrecks Association, museum partners, wreck divers,

volunteers, researchers, school/university students and the public) to interact with and potentially enhance connected collections.

WP3.2 has defined a sample of unidentified wreck sites and will identify maritime collections and data sets required to support wreck identification. A documentary and data tagging initiative will be developed by creating a search criteria classification system designed to link key relevant historical collections and data. In addition, a site characterisation and classification approach will also be constructed, linking marine collections and data sets and principal geological and hydrodynamical processes and associated biological conditions to create a mechanism to evaluate likely condition and survival trajectories. The two protocols will be assessed in the context of the creation of appropriate adaptive management strategies, and on the applicability and relevance of this approach at a wider UK and international scale.

WP3.3 is developing a computer simulation to combine a variety of data sources to facilitate public access to Doggerland. Sources are disparate in spatial scale, storage size and computational demand (e.g. sonar bathymetric maps, seismic profiles, publicly available archaeological collections such as CITiZAN, local Historic Environment Records, ADS online library, and palaeoenvironmental seabed cores). The European Research Council Europe's Lost Frontier's output (the modelled landscape at the beginning of the Holocene) will form the base. New data collections will be integrated to test and advance the simulation. The simulation's front end will allow the user to 'fast forward' from the last Ice Age, melting glaciers and changing sea level to any selected time. Through a small band of Doggerland hunter-gatherers inhabiting the landscape at the time specified, the user will be able to either observe or take significant decisions for the simulated community, thus accessing the data patterns of the submerged landscape via the processes which formed them. The integrated collections will not tell the story of Doggerland but allow users to find their own stories within it.

WP4 is setting about engaging with existing audiences, finding new audiences and communities and co-designing the interface underpinning new ways of exploring and generating novel and unexpected narratives from data which may use multiple channels, e.g. embodied and sound-based. It will reveal user requirements and preferences for engaging with interoperable collections. It will deploy a Rapid Application Development methodology to transform requirements into new and sustainable interfaces for three case-studies, addressing specific audiences, audience requirements and three key research themes. (i) Co-design and User Generated Pathways data layer for non-coastal communities, with a focus on trading/commodities and migrant communities over time (cf WP3.1 datasets). This case study will develop new audiences, engage via new interaction modes and is intended to surface unexpected, community specified, connections between the interoperable data. (ii) The visually impaired public, using spatialised audio content (e.g. derived from WP3.2 sonar datasets) to both represent and allow interrogation of interoperable datasets, and address barriers to access presented by traditional complex database interfaces. Using creative response this case study will be extended to generate novel means of engagement with wider audiences whose interest goes beyond the historic and scientific to the aesthetic, social or spiritual (qv Burra Charter); and (iii) Cross-disciplinary research, using natural and cultural heritage collections derived from WP3.3 simulation, allowing researchers from both natural and cultural disciplines to explore and discover connectivity between these domains. The interoperable UNPATH datasets will be made amenable to multiple modes of access through an UNPATH 'Explorer' which will be demonstrated and promoted as a web-based resource, a mobile application, and an immersive VR interface (e.g. HMD) and installations evaluated and deployed via WP5.

WP5 is leading the co-design of evaluation methodologies for audiences of the products and will apply those methods to assessment of the Incubator, Innovator and Impact components of UNPATH. Following the

Digital Culture Charter and informed by the digital equity methodology of D'Ignazio & Klein (2020), WP5 began by establishing a values framework and associated metrics for the project and will audit progress against the framework on an annual basis. Via audience mapping, WP5 is currently identifying target audiences for recruitment into UNPATH in Year 1, and work between them and across WPs1-4 to ensure that project outputs accommodate user needs, that the project's exhibition programme is curated for accessibility, and that inclusive evaluation methods are developed and applied across the project to accommodate neurodiverse participants. WP5 will then coordinate user evaluation activities in the Innovator and Impact phases of the project, utilising (accessible) survey, interview and automated data acquisition tools. This includes an assessment of the efficacy and inclusiveness of the co-design approach itself at the close of the Innovator Lab, and analysis and publication (with WP3 and 4 leads) of learning, social and emotional outcomes of audience workshops, exhibitions and other dissemination activities. WP5 will ensure delivery of a final exhibition programme, with the National Maritime Museum coordinating museum partners and industrial and third sector collaborating organisations to extend the reach of the project further. The nature of the exhibition programme will be established in consultation with our key audience groups (including virtual and pop-up events), integrating (where possible) with existing programmes such as anniversary and thematic events. WP5 will also track the application of the research tools enhanced in the Incubator Lab across the life of the project, and map the novel interpretations, questions and research findings that emerge.

In summary, the research approach adopted by UNPATH is addressing the key Towards a National Collection Impact Areas in the following ways.

(a) **The exploration of how multiple collections can be explored in the future**. UNPATH does this by forming core metadata model and standards, linking disparate collections into a single schema, by testing those links through three real-world research queries, and by then freeing the results to be accessed by carefully defined invited research and public audiences using new tools co-designed by them themselves.

(b) **The dissolution of barriers and opening of public access**. The collections we are looking at range from born-digital to physical objects and from summary inventory data to point-cloud datasets. There are significant challenges (both technical and social) that we expect to come up against and either to resolve or to identify as currently beyond our ability to resolve. Some examples of these challenges include:

- *The analogue data and metadata of written sources, charts and records*. Considerable numbers of archive, museum and library holdings remain analogue and while digital datasets connect to them, there is no capability to link. We are aiming to explore a Digital/Analogue connector (WP3.1) which will at least begin to chart these connections.
- *The complexities of some of the collections/datasets*. For example: national inventories are not internally or comparably consistent; a 'wreck' can often include multiple sites (through drift of material); sonar survey technology over time has changed leading to differences within apparently coherent collections; some collections have a temporal dimension (seabed surveys, wreck surveys) complicating access.

By working with audiences to co-design pathways into these collections, we will develop a bottom-up approach to solutions which should resolve technical issues but create desirable access. The solutions are being designed to be transferrable.

(c) **The diversification of audiences and innovation of access**. The character and method of establishment of our three key target audiences (WP4 and 5) will ensure that UNPATH reaches a diverse group of potential

users and can accommodate ranges of (dis)ability and neurodiversity. The 'Unpath Explorer' (WP5) will extend access (via web, immersives, mobile app and museum installations) to the interoperable datasets which we anticipate will provide a model for future access. We have selected the three audiences to test each of the three kinds of motivation identified by Bailey-Ross (2021) (i.e. Casual Use, Personal Interest, Scholarly & Professional Research). In the case of the researchers, we anticipate cross-disciplinary interest in the interplay between natural and cultural heritage collections (especially as a result of the work in WP3.2 on the integration of ocean and seabed sciences and historical and archaeological sciences).

(d) **Setting a global standard for others building their own collections**. Work on integrating and enhancing marine inventory collections (national datasets) interoperability will allow international aggregation (such as the completed Perio.Do definitions for all of Ireland). The protocols established in WP3.2 likewise will have an international value. UNPATH has already led to a Getty Art and Architecture Thesaurus for maritime craft which represents an early impact on international standards and will deliver an ADS Open Access world-renowned Guide to Good Practice.

(e) **Enhancing collaboration between UK and national collections worldwide**. UNPATH has yet to develop significant international collaboration, but we expect this to arise through the issue of the outputs above (d), and through international events where UNPATH will be promoted (European Archaeological Association, Computer Applications in Archaeology conferences). We know that involvement with our Partner Manx Heritage will enable them to capitalise on UNPATH benefits in developing their own inventory of marine heritage, and we have colleagues from Ireland and the Netherlands on our Advisory Board.

Research results

The UNPATH programme is progressing well. We have completed the primary dataset collation (WP1) (c 94 different datasets) and established the UNPATH ontology (UO-CAT). The ingest of national level inventory datasets for Scotland is complete; work on Wales, Northern Ireland, England, ADS archives and OASIS grey literature reports is ongoing, with Isle of Man to follow. We have established Perio.Do (ie chronological) definitions for Scottish and Irish data (including the whole of Ireland), and we have agreed Getty AAT thesaurus mapping for maritime craft. The project triplestore is operational, and we have already been able to install a proof of concept instance of the European ARIADNE portal badged and hard-wired for UK maritime data for UNPATH.

Related to the work on linking UNPATH and the English inventory (Mariner), we are testing concepts of connectivity between collections and marine inventories, including a 'Spatial Slider' capability: linking historical maps to modern charts (qv the British Library Georeferencer project) and a 'Temporal Slider' capability to test access to time series (e.g. journeys of vessels) in map form, to see whether these capabilities can be linked to national inventories.

For the access and search tools (WP2) we are now testing interoperability of data and reviewing the potential mix of NLP/zero shot/object detection to be deployed on samples of the datasets. Our initial focus will be on Welsh wreck inventory and sonar scan data and Lloyd's Register vessel loss data to test identifying wrecks.

Our *People and the Sea* (WP3.1) research lens has identified its target sample wreck data for the Digital/Analogue Connector. They are off the Needles, Isle of Wight, have a fairly wide date range and intriguing histories and include: *Campen*, Dutch East India vessel, lost 1627; *HMS Looe*, Fifth Rate, 1705; *HMS Incendiary*, fire ship, 1780; *Dream*, schooner, 1850; *Anglo Saxon*, Channel Island brig, 1879; *Serrana*, steamship, 1918; *SS Varvassi* cargo ship, 1947. To support immersives (qv WP4), a student hackathon at Portsmouth piloted an immersive competition. Two teams of five (post-graduate) students explored visualisation of genetic data from the Mary Rose including from the Archer Royal skeleton to create a visual immersive/interactive. The brief for this pilot hackathon was student-led. They also looked at other datasets (one team looked at ancestry.com's DNA service display data). The outputs were two prototype digital immersive, one targeted at families and one at researchers.

Our *Science and the Sea* (WP3.2) lens has completed sample selection of unidentified wrecks in an area west of the Isle of Man, and correctly formatted related multibeam datasets for analysis. We are now identifying relevant collections and archive sources likely to aid in their identification. For the environmental work, development of a scientific metadata catalogue is underway, as is evaluation of scientific data topics and outputs (e.g. hydrodynamical model outputs and relevance/value of specific parameters such as spring/neap variance, flow depths and tidal direction).

For *Lands Beneath the Sea* (WP3.3) the focus has been on establishing and integrating all the necessary datasets and collections to permit the development of the primary simulation infrastructure. This has progressed far enough to permit the first surface terrain model for the study area (southern North Sea).

Work on developing the immersives and visualisations has not yet started in earnest: the principal effort has been on ensuring alignment and integration between this work and WP2 to identify any impact on WP4 from AI data enhancement; to ensure links between this and immersive pilots in WP3.1; and to assess any impact of Unity based Doggerland modelling (WP3.3) and how this will articulate with the Unpath Explorer.

WP5 has published a [Values Framework](#) on the website, and prepared an associated plain-language description of the methodology, and has now completed the initial Audience Mapping. The team gathered audience information from the UNPATH consortium to understand current and previous audiences of their organisations, and how they overlap (if at all) with the three target audiences of UNPATH. Analysis is ongoing.

Project outputs (to date)

While it is still relatively early in the project, the following project outputs have been created as of end July 2022, which we consider provides evidence of good progress. Many are essential stepping stones within UNPATH, but some are (or will be) standalone products with enduring impact. The latter are emboldened below:

WP1

- UNPATH ontology (UO-CAT) v1.1 finalised
- Metadata licence drafted
- Provisional Perio.do definitions for Scotland and Ireland published
- Triplestore operational with Scottish national Canmore dataset loaded
- Getty Art and Architecture Thesaurus for Marine Craft established

WP2

- Parsing of .json file from Coflein for initial processing
- Integration of dictionary into Natural language tool kits

WP3.3

- Extracted 2D seismic layer for integration into simulation.
- Base GIS created.
- **Prototype surface created**.
- Corpus of reports relating to submerged landscapes created for AI investigation by other WPs

WP5

- 1st release of Co-designed Unpath'd Waters Living Values
- **Online update on methodology behind the creation of Living Values**

WP6

- Publication of UNPATH project website.

Cross Project Collaboration

In the set-up stages of the project, we have had little opportunity to begin to work with the four other TaNC projects in any great depth outside of the strong inter-project collaborations initiated by the Programme Director and her team. These TaNC sponsored events (all-projects meetings, workshops on technical aspects, project management approaches, decolonisation of research, and media interaction training) have provided fertile grounds for identifying some key cross-project collaboration strands. One example of strong cross-project collaboration which has emerged from UNPATH is the work undertaken as part of WP5 on project values, as mentioned above. At a cross-TaNC workshop, enthusiasm was shown from all the other projects on understanding what we were trying to accomplish with work on Values and how this might tangibly affect TaNC project engagement with different audiences.

It seems very likely that another of UNPATH's key strengths – the strong spatial relationships embedded in most of the collections with which we are working – may provide a similar opportunity for cross-collaboration in the near future. We also believe a third opportunity for collaboration will arise as all five projects develop their approaches to AI.

Sustainability and Infrastructure

Sustainability of the UNPATH project products and outputs will be ensured by adhering to FAIR principles and using international data standards.

In the short term (i.e. during the duration of the project), university partners use networked computers which are regularly backed up to secure storage, using a mix of tape, remote hard drive and AWS or Google cloud provision. Project management information and shared documents will be held on shared team drives, or cloud services, securely managed and backed up by the respective IT services and to which only authorised staff have access. Non-HEI partners all follow industry-standard backup and data management procedures, and use networked drives and computers, and either designated data centres, or third-party servers, operated by Microsoft and Oracle and located in Western and Northern Europe, backed up daily. For new primary data transferred to ADS, core data in the form of AIPs and original (the slightly altered SIP) files are regularly synchronised from a local copy in the University of York to AWS.

Long-term preservation of new digital data generated or enhanced within UNPATH will be the responsibility of the archivists within the national heritage bodies, supported by ADS for data that would otherwise be orphaned, and specialist data types. All follow the OAIS model and ADS and RCAHMW are able to assign Datacite DOIs to all data sets and reports but will also observe recommendations of the TaNC Persistent Identifiers Foundation Project. In ADS spatial data is published using Geoserver and incorporated into web interfaces using Openlayers (v5). The Highslide (javascript) application is currently used to provide zoomable access to high resolution JPG images in the archive, but ADS & RCAHMW are investigating the deployment of IIIF and will build upon the recommendations of the TaNC IIIF Foundation Project.

ADS archives are preserved into perpetuity; additional data held by university partners and not transferred to ADS will be held for a minimum of 10 years by the institutional repositories. Beyond the project lifetime, data which forms the UK marine record will continue to be updated by the national bodies for their areas of responsibility as part of their national roles. For England, the National Marine Heritage Record is currently held within Warden, a bespoke database developed on the Getty Arches platform; in Scotland, Canmore is Oracle-based; in Wales Arches provides the front-end and Preservica is used to ensure long-term preservation and access; in Northern Ireland data is held on a SQL Server with public access provided by an ESRI web application.

ADS and the national bodies will maintain access to the distributed marine record and associated digital archives using their respective current and enhanced search interfaces: ArchSearch, Heritage Gateway, Mariner, Canmore, Coflein etc. For external portals we are reliant on their sustainability plans, but MEDIN's marine portal is a national service, and there is a commitment by the Swedish National Data Service and CNR in Italy to maintain the ARIADNE portal for two years beyond the end of the current funding period in 2023, with plans being developed for longer term support as part of an EU ERIC infrastructure and/or EOSC. All data held by ADS is available online free of charge, and under a CC-BY or other open licence, and metadata is CC-0.

The TANC Github site will become a repository for all source code generated in WP2-WP4 and made available CC-BY-NC. WP3.1 Holland submarine visualisations will be maintained on web platforms by the University of Portsmouth, and/or Nautical Archaeology Society for five years after the project. For WP3.3 the software simulation source code will also be freely available via the CoMSES simulation repository. The software will be available free as an installer for desktop PCs via Itch.io. While the installer is vulnerable to

future operating systems changes and may not function reliably in the long term, availability of the source code will allow the software to be altered and recompiled to ensure compatibility with any such changes. There will be no dependencies on commercial libraries to facilitate the recompilation. For WP4, the Explorer project files will be maintained on Github and the Glasgow School of Art but will be vulnerable to changes in dissemination technology. Versions created during the project for particular events e.g. standalone, web, mobile etc, will be archived for re-use by SimVis at GSA. Historic England will maintain a dedicated project website, with links to all outputs, for 10 years after the project end date.

Contacts

Barney Sloane: barney.sloane@historicengland.org.uk for general overview of project research

Julian Richards: julian.richards@york.ac.uk for technical data information

Fraser Sturt: F.Sturt@soton.ac.uk for AI information

Stuart Jeffery: S.Jeffrey@gsa.ac.uk for simulation and visualisation information

Sara Perry: sperry@mola.org.uk for audience research information