

## ***Towards a National Collection: Opening UK Heritage to the World***

### **Foundational Projects**

<b>Grant reference</b>	<b>Grant Title</b>	<b>PI</b>	<b>Co-Is</b>	<b>Research Organisation</b>	<b>Partner research organisations</b>
AH/T011068/1	Heritage Connector	John Stack	Jane Winters, Jamie Unwin	Science Museum Group	University of London
<p>As with almost all data, museum collection catalogues are largely unstructured, variable in consistency and overwhelmingly composed of thin records. This is largely a legacy of the development of these catalogues from handwritten paper records. The depth and form of collection catalogues has been primarily guided by collection management needs (records of acquisition, administration of loans, provenance documentation, etc.) where unstructured data can fulfil the needs of the organisation and comply with collection management standards. When computer technology was adopted for collection management in the 1980s it was implemented to handle these same back-office tasks rather than to support public access. The resulting form of the catalogues means that the potential for new forms of digital research, access and scholarly enquiry remain dormant, and searching across collections is currently possible only through aggregation which is labour intensive to implement, or by third-party search engines where results are unreliable. In this project, we will apply a battery of digital techniques to connect similar, identical and related items within and across collections. Our primary research question is ""How can existing digital tools and methods be used to build relationships at scale between poorly and inconsistently catalogued digitised collection objects and other content sources?"" Since the turn of the twenty-first century enormous and growing volumes of material have been digitised, and catalogues have begun to address the needs of digital public access. However, this has been mainly at an institutional level or via a handful of content aggregators and thus the enhancement of catalogues for the purposes of public access has been driven by the needs of individual collection websites, with little or no interlinking to other collections or content sources. Where that linking (people, places, events, objects, etc.) does exist it has been undertaken by human intervention, and because of the number of records, it has been limited in scale and scope and rarely an ongoing endeavour despite the evolving nature of the catalogue. Alongside the digitisation of collections, recent years have seen a growth in the publication online of scholarly research related to heritage collections: open access journals, theses, and other online resources. However, beyond the host institution, references to this material is rarely, if ever, ingested into the underlying collections systems and made available via links from related collection websites. This project will therefore also use computer analysis to attempt to identify and build links to this material. Finally, structured data and rich linking are an increasingly urgent concern as new forms of discovery and access emerge - notably artificial intelligence powered discovery and new interfaces such as voice search - that rely on these for their functionality. This project will explore an alternative approach - a "Heritage Connection Engine" - that will analyse catalogues, published material and knowledge graphs, and build links at massive scale between these that can then be used for new forms of research. It will explore the opportunity for computer generated links with Wikidata to provide new levels of structure and machine-readable data that can form the foundation of new types of discovery and access. The "Heritage Connection Engine" will use a range of technologies including machine learning; named entity recognition; open data; and persistent IDs. These methods will create a large-scale data source of links, each with a</p>					

confidence ranking. Computational enquiry to generated links via an application programming interface (API) will enable the creation of a range of proof-of-concept research and discovery tools. All software will be documented released under an open source Licence. All datasets will be released under the Creative Commons Zero license.

AH/T011076/1	Provisional Semantics: Addressing the challenges of representing multiple perspectives within an evolving digitised national collection	Emily Pringle	Helen Mavin, Tate Greenhalgh, Anjalie Dalal-Clayton	Tate	Imperial War Museums; The National Trust; University of the Arts, London.
--------------	--	---------------	---	------	---

Provisional Semantics interrogates how museums and heritage organisations can develop ethical and transparent readings to represent stakeholders appropriately and thereby foster engagement between a more diverse public and the digitised national collection. Currently, many subject index terms, catalogue entries and captions of artworks and artefacts have been informed by colonial contexts, attitudes and modes of perception. These can be outdated and/or offensive to contemporary audiences, not least people of African and Asian descent, whose diasporic histories are intertwined with Britain’s colonial past. Currently the lack of research examining ethical methodologies and required practical and attitudinal shifts prevents sustainable change from taking place.

Provisional semantics addresses this challenge by (1) examining what methodological, ethical and practical changes heritage organisations need to make to accommodate the multiple shifting interpretations needed for the digitised national collection to genuinely represent UK Heritage. And (2) testing what decolonial methods heritage organisations can employ to produce interpretive frameworks and terminologies fit for an evolving digitised national collection. Over 18 months the project will undertake a literature/practice review and utilise three case studies examining collections at the National Trust, Imperial War Museums and Tate that address the histories, representations and artistic practices of people of African and Asian descent. Each case study tests an approach to collaborating with key stakeholders of African and Asian descent whose expertise can help provide more equitable, multi-perspectival interpretations. The research is supported by the project’s academic partner, Dr Anjalie Dalal-Clayton, from the Decolonising Arts Institute, University of the Arts, London

AH/T011084/1	Practical applications of IIF as a building block towards a digital National Collection	Joseph Padfield	Torsten Reimer, Charlotte Bolland, Melissa Terras	The National Gallery	British Library; National Portrait Gallery; University of Edinburgh
--------------	---	-----------------	---	----------------------	---

How can we best build a common infrastructure for sharing high quality images from heritage collections and sites?

This project will explore and demonstrate the possibilities of the International Interoperability Framework (IIIF) to support the dissemination of born digital and digitised heritage images for research and engagement. IIIF, along with a wide range of freely available IIIF compliant software, represents a flexible, standard approach to providing reliable access to such images. However, though well described, setting up and re-using such IIIF resources can still be complex, particularly for smaller institutions or individual researchers. Also, a better understanding of how to combine IIIF resources across multiple institutions and present a National Collection, to diverse audiences, is needed.

This project aims to demonstrate the opportunities and benefits that IIIF offers, to a wide audience of users and help to define more robust use cases of IIIF, for institutions, but also for individual researchers who want to re-use and exploit IIIF resources to carry out new research, create new opportunities and tell alternative stories. This work will highlight existing software and resources and identify what new tools, services or training might be required to maximise the potential of IIIF within the heritage community.

This will be achieved through a series of targeted workshops and surveys, along with the creation of pilot demonstrators, to provide tangible examples of what can be achieved. This process will involve both technical and non-technical researchers in collaboratively building resources, to increase mutual understanding of the possibilities and requirements.

AH/T011092/1	Persistent Identifiers as IRO Infrastructure	Rachael Kotarski	Lorna Mitchell, Roderic Page, Joseph Padfield	British Library	Royal Botanic Garden, Edinburgh; University of Glasgow; The National Gallery
--------------	--	------------------	---	-----------------	--

Museums, heritage collections and sites in the UK house at least 200 million physical and digital objects. Being able to identify these objects supports their use and curation - you cannot provide persistent or even consistent access to an item if you don't know what it is. Accession numbers are therefore a key component in all collection and library management systems but these only cover selected objects within an individual collection. To fully realise the potential of our national collections, we need identifiers that can bring together collections across institutional boundaries.

Persistent Identifiers (PIDs) provide a long-lasting click-able link to a digital object. They are recognised by UKRI as a tool for enabling data discovery, access and citation. Supporting wider use of PIDs for collection objects, environments, specimens and related items will allow long-term, unambiguous linking of collections that will create a digital UK National Collection. However, the challenges, utility and wider benefits of PIDs are less well understood across the heritage sector.

This project will bring together best practices in the use of PIDs, building on existing work and projects. We will share expertise and provide recommendations on the approach to PIDs for colleagues in institutions across the UK heritage sector. Through a mixture of workshops, desk research and case studies, the project will answer questions such as 'What are the gaps in the existing PID landscape for heritage

collections, buildings and environments?' and 'What should a PID infrastructure, strategy and governance framework look like for a unified national collection?'					
AH/T011114/1	Deep Discoveries	Lora Angelova	Pip Willcox, John Collomosse, Joanna Norman, Robert Cubey	The National Archives	University of Surrey; Victoria and Albert Museum; Royal Botanic Garden, Edinburgh
<p>The way we access information in the virtual space is changing. Discovery and exploration are no longer constrained by a keyword entered into a blank search bar. Instead, museums, libraries, archives, and galleries worldwide are welcoming a shift to 'generous interfaces' – presenting their collections online in browsable and linkable networks of information that allow users to explore and discover new ideas through meaningful and contextualised relationships.</p> <p>A key component in this emerging virtual browsing landscape is 'visual search', an AI-based method for matching similar images based on their visual characteristics (colour, pattern, shape), rather than a keyword description. The Deep Discoveries project aims to create a computer vision search platform that can identify and match images across digitised collections on a national scale. The research will focus on botanically themed content, allowing us to test how far we can stretch the recognition capabilities of the technology (can it recognise a rose in a textile pattern and the same flower in a herbarium specimen? How about on a ceramic vase?).</p> <p>Deep Discoveries is a collaboration between The National Archives, the University of Surrey, Royal Botanic Gardens Edinburgh and the V&amp;A. We will start with ~30,000 botanical images from the institutions and our project partners Gainsborough Weaving Company, the Sanderson Design Archive, and the Museum of Domestic Design and Architecture. The wide range of partners will also allow us to explore the necessary criteria for our nation's image collections to be linked, and to survey the searching needs of diverse audiences.</p>					
AH/T011122/1	Engaging crowds: citizen research and heritage data at scale	Pip Willcox	Christopher Lintott, Elspeth Haston, Martin Salmon	The National Archives	University of Oxford; Royal Botanic Gardens, Edinburgh; National Maritime Museum
<p>Engaging Crowds: citizen research and heritage data at scale harnesses the capabilities of people-powered research to enrich understanding of cultural heritage collections through digitally enabled participation. Citizen research enables volunteers, including people previously not reached by cultural heritage organizations, to participate in research projects, helping classify, annotate or transcribe collections.</p> <p>Engaging Crowds explores the current and potential practice of engaging diverse audiences with the creation, use and reuse of heritage data. It investigates best practice in the citizen research landscape, by considering how institutions assess, present and value the work of</p>					

volunteers and by working with three groups that are potential recipients of the data volunteers produce: collections-holding organisations, machine learning algorithms, and the citizens themselves. Engaging Crowds is developing a new indexing tool that gives volunteers the agency to choose their own pathway through a project and will evaluate it by running three citizen research projects. A report based on these findings together will recommend the best ways of encouraging and supporting meaningful public interaction with heritage collections.

The project is led by The National Archives, with the Zooniverse team at the University of Oxford, Royal Botanic Gardens Edinburgh, and Royal Museums Greenwich. All partners have significant experience of conducting citizen research projects, not least using Zooniverse, a free, open source platform with more than 1.9 million volunteers worldwide. Engaging Crowds aims to promote open discovery and research, interdisciplinary working across sectors and equality, diversity and inclusion.

AH/T01119X/1	Locating a National Collection (LANC)	Gethin Rees	Leif Isaksen, Alex Hunt, Anthony Musson	British Library	University of Exeter; The National Trust; Historic Royal Palaces
--------------	---------------------------------------	-------------	---	-----------------	--

This project will help cultural heritage organisations to use location data - such as where objects were made and used or the places they depict and describe - to connect cultural heritage collections and engage audiences. Through scoping, workshops and audience research, the project will establish best practice and provide technical recommendations for a national system that allows collection items from diverse organisations to be discovered through location. The Pelagios Network of researchers, scientists and curators has developed a methodology that uses gazetteer referencing to link research data with considerable success, building a community of partners and stimulating research. This project builds on their methodology, scoping improvements and exploring ways to render content accessible and meaningful to different public audiences. To test and scope this ambition, participating organisations will work on a set of thematic and technological case studies to understand the requirements of stakeholders, institutional, academic and public. The objective is to understand the technical components required, the options available and to make recommendations for potential solutions, all described in the project report. The report will constitute a strategy to inform developments in the next phase of Towards a National Collection and across the cultural heritage sector more broadly. It will encourage cultural heritage organisations to take up a common geospatial approach and provide a roadmap that enables diverse organisations to enrich their metadata and expose this in a consistent and joined-up way, ultimately spear-heading a movement beyond text-based searches in cultural heritage.

AH/T01122X/1	Preserving and sharing born-digital and hybrid objects from and across the National Collection	Natalie Kane	Joel McKim, Richard Palmer, Stephen McConnachie	Victoria and Albert Museum	Birkbeck College; British Film Institute
--------------	--	--------------	---	----------------------------	---

This project addresses the challenge of collecting, preserving, linking and sharing born-digital and hybrid collections with the public across the national collection. It aims to encourage knowledge exchange in order to build on collective experience and challenges and drive thought-leadership across the sector regarding digital collecting.

The project seeks to tackle the unique issues that face the collecting of born-digital and hybrid objects (which consists of both physical and digital elements) into museum and collecting organisations, for example, internet-enabled product design such as Apple iPhone, which requires preservation of the software alongside the hardware it is presented on. This includes other complex digital objects such as social media, open-sourced and user-generated design, videogames and design that relies on 'live' networked interactivity.

Led by the V&A, with the BFI and Birkbeck, University of London, the 18-month project will include three workshops, involving invited participants from across the national collection and sector outside UK, focusing on three key areas of digital collecting; collections management, conservation, and access, experience and meaning. Outputs will include a report that addresses the current state of preservation and access of born digital and hybrid objects, a set of co-authored recommendations for skills and resources to help develop digital collections, and guidelines on changes to governance and acquisition policies that would support the management of digital objects across the national collection. It will also prototype a data model based on linked open data to describe born-digital and hybrid objects and related "digital preservation" conservation documentation.