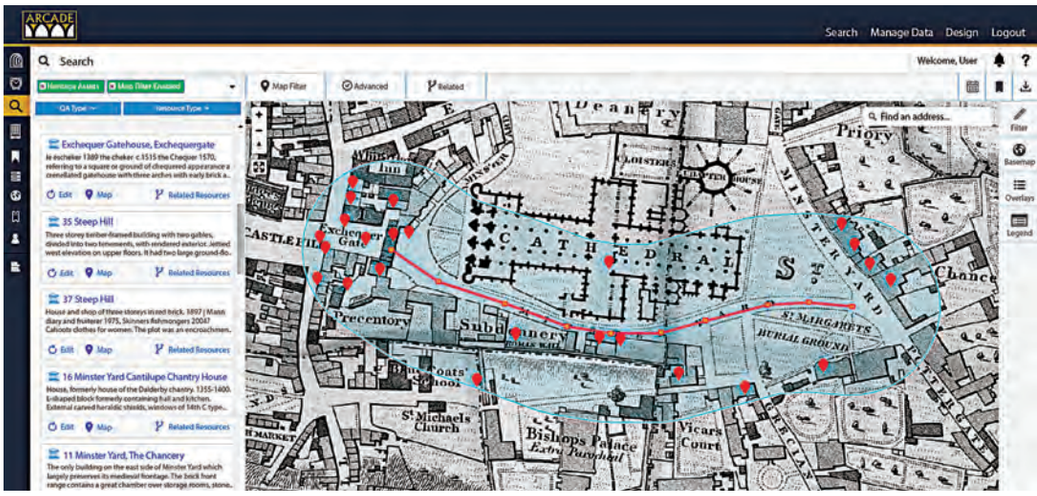
THE ARCHES PLATFORM

Bridging Heritage Pasts and Data-Rich Futures

BY ALISON DALGITY, DAVID MYERS, AND CATHERINE SCHMIDT PATTERSON



Using the Location Filter in Arches, users can spatially query data by drawing a line or an area and specifying a buffer size, viewed here over an underlying 1842 base map. This search functionality may be used to identify heritage resources that would be impacted by proposed development projects in this example from the City of Lincoln Historic Environment Record (HER) deployment of Arches. Graphic: Getty Conservation Institute.

CULTURAL HERITAGE DATA BY ITS VERY NATURE IS largely subjective, frequently incomplete or imprecise, and typically changes over time. These and other characteristics make it challenging to manage, even with the most modern and sophisticated digital systems. Moreover, public heritage organizations charged with the protection of our cultural heritage—and, of course, knowledge *about* that heritage—are notoriously underfunded worldwide, often lacking the resources and technical expertise needed to develop effective data management systems. This often results in the use of proprietary software never intended for the heritage field. Because of the commercial origin of such software, organizations may find themselves and their data locked into ongoing licensing fees and maintenance contracts. Organizations that create their own systems tend to end up reinventing the wheel by developing custom systems that share features with systems already developed. In addition, heritage organizations are sometimes forced to continue using outdated software, which not only is incompatible with other systems, but eventually renders the managed data inaccessible and therefore unusable.

After years working jointly in this domain, the Getty Conservation Institute (GCI) and World Monuments Fund (WMF) recognized the need to address these and related challenges. Inventories of heritage places such as archaeological sites, buildings and structures, and cultural landscapes must be kept up to date to ensure that management decisions are informed by accurate and current information. As such, an inventory system must facilitate the seamless incorporation of new data and the ability to readily share information. Failure to properly manage heritage data can have serious consequences, such as delays protecting heritage places in the wake of natural disasters or armed conflicts, and it can create uncertainty about whether heritage resources would be impacted by proposed development.

CREATING ARCHES TO SUPPORT THE FIELD

In 2012, following consultations with several heritage organizations with experience in managing inventories and developing standards, including Historic England and the Flanders Heritage Agency, the GCI and WMF jointly invested in the development of Arches, a generic software platform to help address the challenges described above. Using the latest technologies and an open-source approach, version 1 of the Arches platform was released in 2013, freely available worldwide to independently deploy and customize as needed.

With the release of version 7 in 2022, Arches[1] has expanded far beyond the original requirements and now addresses data management demands in areas of the cultural heritage field as diverse as conservation science, publishing archival cultural materials, and heritage provenance. This expansion was made possible by the original Arches software design principles and priorities, which included:

- purpose built for the cultural heritage field;
- standards based (technical and heritage open standards), with ability to override standards if necessary;
- economical (freely available with no licensing fees or vendor lock-in);
- independently deployable;
- built with flexible, customizable, and publicly accessible software code;
- able to control access to data at any level (from fully restricted to openly shared);
- user friendly;
- committed to establishing a broad community for input and promoting collaboration.

One of the first important decisions in developing Arches was to make it open-source software (OSS). This facet offers numerous benefits. Unlike proprietary software, OSS code is freely available

1. https://www.archesproject.org/

and open for further improvement and customization, eliminating dependence on an individual vendor and the consequent risk of being locked into long-term licensing and maintenance costs. While the code is open, data in Arches is independently controlled and can be made open or private. The overriding ethos of an open-source environment is one that promotes community support, collaboration, and resource pooling. Additionally, the Arches open-source license stipulates that all improvements made to the code are to be freely shared with the broader community.

ARCHES CAPABILITIES

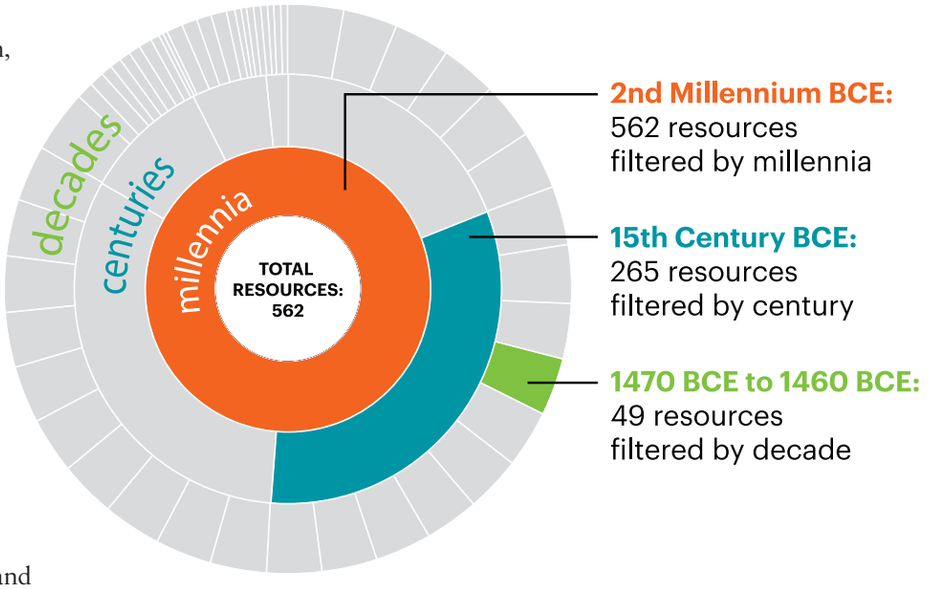
The capabilities of Arches fall into three broad categories: data management; data visualization and discovery; and workflows for task management.

Data management tools include the ability to create, edit, and share data, and to publish selected information online, while maintaining granular control over data access. Arches uses common nonproprietary file formats, and once data is structured within Arches it will outlive the software. Therefore, data entered or migrated into Arches today will be available to support a vast array of conservation goals far into the future.

There are a variety of tools that facilitate the *visualization and discovery* of data within Arches, including a map interface that can integrate historical maps and satellite imagery, allowing heritage resources to be discovered geographically. Relationships can be established and visualized on an interactive graph depicting an expanding web of relationships where previously unknown information can be discovered. For example, the graph can visualize relationships between a heritage site or object and historic events, conservation activities, scientific or other scholarly reports, and people and organizations associated with it. Linking this information to dates or other temporal attributes allows visual time-based searches using the so-called Timewheel. The ability to build these relationships and search by concept as well as by keyword is made possible by the underlying semantic data structure[2] and the use of localized controlled vocabularies.[3]

To help organizations automate their business processes, *task management workflows* can be customized to create data editing sessions that mirror those processes. For example, a workflow can be created to track the status and outcomes of heritage impact assessments. Another might record the process of taking a sample from a heritage object or place, describing its location, the purpose of the sample, and the instruments used to analyze it. Moreover, the process of sophisticated data structuring is automatically folded into the intuitive data-entry forms in a workflow, thereby shielding the user from the complexity that makes this possible.

Arches continues to expand and add capabilities. For example, internationalization, which allows all languages and scripts to be used, is now available in version 7. The first iteration of Arches for Science



The Timewheel is a circular histogram allowing users to filter data based on a time period—i.e., a millennium, century, or decade. The size of each time period segment represents the frequency that period appears in the data. Graphic: Getty Conservation Institute.

(AfS) has also been completed and is currently being tested. AfS is an expansion of the platform that will help conservation scientists and others to secure, retrieve, visualize, compare, and share scientific data and to track technical examination projects of heritage objects and, potentially, heritage places. It also includes the ability to annotate images of cultural materials under study, such as museum objects. Development has begun on building a new Reference Data Manager (RDM), which allows organizations more control to integrate, combine, and manage their local vocabularies and thesauri, helping ensure consistency of data entry and greatly improving search results.

ARCHES IN USE

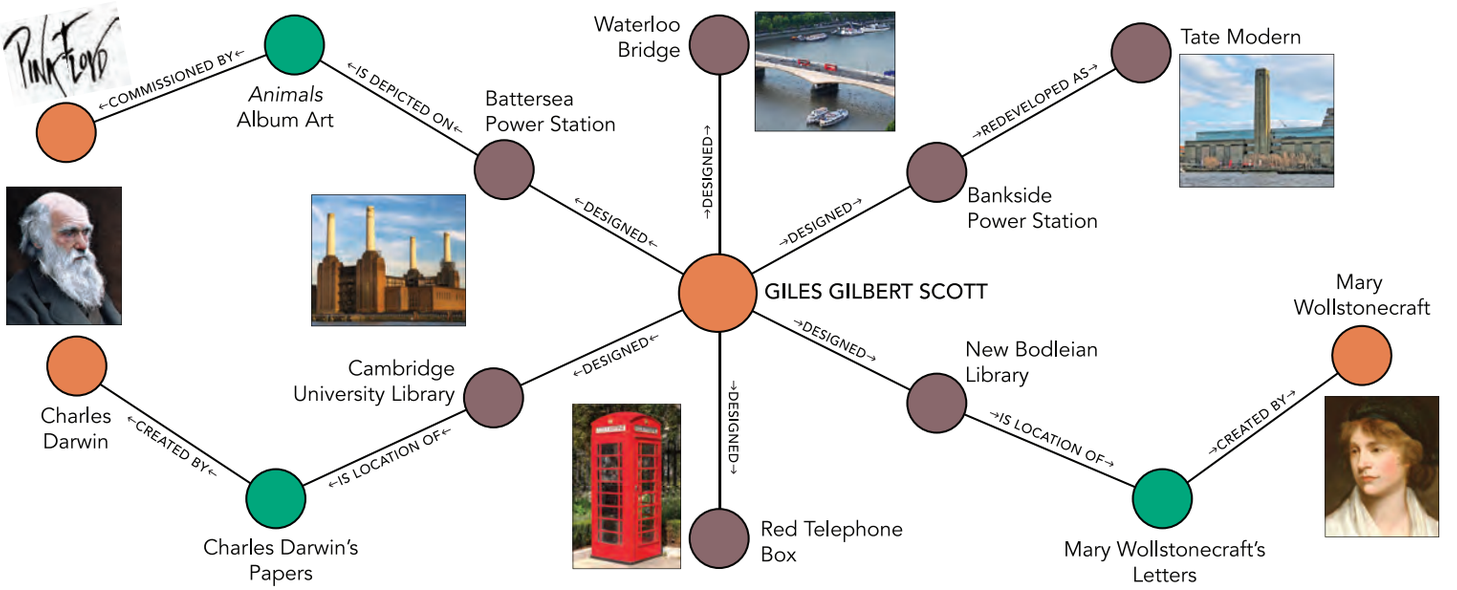
To date, the GCI knows of nearly one hundred implementations of Arches (already launched or in preparation), with many more in the planning stage; the open-source nature of Arches means there may be many more, including for uses beyond cultural heritage. The known implementations collectively record heritage spanning five continents and nearly sixty countries.

A few examples of Arches deployments convey the range of its uses in the field:[4]

- Jamaica, Barbados, Wales, Jersey, and the Isle of Man have deployed Arches for their national inventories of heritage places, and Canada is preparing to do the same. England is additionally preparing to implement Arches for its national maritime heritage inventory;
- on a regional scale, the Dunhuang Academy in China is deploying Arches to record and help monitor Buddhist grotto sites in Gansu Province, and the Florida Public Archaeology Network has implemented Arches as a tool for an ongoing citizen science program to monitor the condition of archaeological sites across the state;

2. Semantic data structuring organizes data in a logical way using an ontology, which is a specification that categorizes data elements and the relationships among them and ensures that data remain humanly readable regardless of the software platform (Arches uses the CIDOC Conceptual Reference Model).
3. The use of integrated controlled vocabularies (or thesauri) significantly improves the accuracy of data input and allows searching by broad concepts rather than by the exact term originally entered in a database.

4. For further information on these and other deployments of Arches, visit the "Who is using Arches?" webpage at: https://www.archesproject.org/implementations-of-arches/



In Arches, relationships between different data types can be established and visualized on an interactive graph. This example shows relationships between architect Giles Gilbert Scott and buildings and structures he designed, as well as further related archival materials, artworks, and other persons. Graphic: Getty Conservation Institute. Pink Floyd Logo: CMetalCore, Pink FloydSVG, 2016. Charles Darwin: © 2019 Julius Jääskeläinen. Battersea Power Station: © 2013 Tosh Marshall. Telephone box: 2012 Christoph Braun. Waterloo Bridge: © 2009 Tom Arthur. Tate Modern: © 2018 Acabashi. Mary Wollstonecraft: John Opie (1761–1807), *Mary Wollstonecraft*, c. 1797, oil on canvas. Artwork: © National Portrait Gallery, London.

- the cities of Los Angeles, Philadelphia, and Greater London and Lincoln in England have implemented Arches for their inventories of heritage places. San Francisco is also in the process of deploying Arches;
- the Arcadia Fund in London is supporting at least ten projects to record endangered heritage places in more than forty countries on three continents, each deploying Arches;
- the GCI is preparing to deploy Arches for Science to manage its conservation science data.

Other Arches deployments focus on heritage site management and on enabling online access to scholarly archival collections and historical maps, as well as 3D data of heritage places and collections.

THE ARCHES COMMUNITY

A defining goal of the Arches project has been to build a collaborative open-source community around the software. From its inception, the project has created infrastructures to attract new members, enable their collaboration, and amplify work of individual contributors. The Arches community has grown to include institutions and individuals representing the government sector, NGOs, philanthropy, academia, and commercial entities. Their involvement ranges from responding to questions from other community members on the forum and submitting bug fixes, to funding software enhancements, such as Arcadia's generous support for the development of internationalization. Through a partnership between the GCI, Historic England, and the City of Lincoln, a version of Arches has been customized to meet many of the requirements of the more than eighty UK Historic Environment Records, called Arches for HERs. This freely available, ready-made, and comprehensive inventory and consultations system will serve as a model of how purpose-built open-source software can provide exciting opportunities for entire heritage sectors. With the impending release of Arches for Science, the community will expand to include conservation scientists and others working with heritage-related scientific data. To date, community-driven Arches User Groups have been established in the United Kingdom and the United States, offering organizations and professionals further opportunities to collaborate and share resources.

WHAT'S NEXT?

The Arches project was established with the strategic aim of helping break the cycle of individual heritage organizations investing scarce resources in re-creating software. It has resulted in a freely available, state-of-the-art software platform requiring only marginal investments for customizations. Implementations around the world have demonstrated that collective investments in information infrastructures can allow heritage organizations to instead focus resources on documenting and protecting heritage and advancing their individual missions.

Even the most cutting-edge and well-engineered software will eventually become obsolete, but most importantly, data must outlive current software to be usable in as yet unknown future technologies. Just as technology advances, so should the structures that support it. As such, the long-term sustainability of Arches will be based not only on technology but also on people and institutions. The GCI is now advancing the Arches Governance Initiative to establish a framework for broader community participation in determining the future priorities and direction of Arches. This is a logical conclusion to a decade of work that has sought to more fully embed effective data management practices into the work of heritage organizations and professionals worldwide, helping the heritage field use the best tools available to protect and conserve our collective cultural legacy.

Alison Dalgity is a GCI senior project manager. David Myers is a GCI senior project specialist. Catherine Schmidt Patterson is a GCI scientist.