National Heritage Science Forum

Digital Society 'Deep Dive' – Event Summary

On 3 May 2023, the NHSF Research Working Group coordinated an online member meeting to look at UK heritage science digital infrastructure, aiming to uncovering structural challenges, hear about potential solutions, and inform the development of the future digital landscape for heritage science research.

'Digital society' is one of five societal challenges identified by the NHSF Research Working Group that can inspire and encourage connections between heritage science research and issues of importance to society. You can find out more about the societal challenges here.

This societal challenge also intersects neatly with a major development in the heritage science sector: the upcoming roll-out of the Research Infrastructure for Conservation and Heritage Science (RICHeS) programme, as well as with several other existing UK and European digital infrastructure projects, such as Towards a National Collection (TaNC) and the Capability for Collections Fund (CapCo).

The morning began with a presentation by **Caroline Peach on behalf of Ben Edwards** from Manchester Metropolitan University. In 2021, Ben was heavily involved in conducting scoping research funded by the Arts and Humanities Research Council (AHRC) to determine the priorities, risks and values for a UK heritage science infrastructure. This presentation was a summary of his findings. You can read the full report here.

The report found three major priorities for infrastructure investment within the sector:

1. Sustainability of digital resources
2. Connections (between people, data, and technologies)
3. Digital technological infrastructure

It also identified some gaps in the existing infrastructure provision:

- Aggregation of the results of heritage science investigations
- Discoverability, accessibility, interoperability and reusability of these results, and of localised sample and reference collections beyond the GLAM sector
- Digital technology connecting curators in the GLAM sector with heritage science technology and expertise to meet GLAM sector priorities
- Digital technology fostering new and innovative forms of analysis of heritage science data

And summarised the primary challenges in addressing these priorities and filling these gaps:

- Existence of data repositories and physical archives of the results of scientific investigation may not be known to the wider community

- Content of physical and/or digital archives of heritage science data may not be accessible or available
- Resources (finance and staff expertise) lacking
- Lack of computer science expertise
- Creating and maintaining relevant digital collections
- Incomplete understanding of the audience(s) for digital data

The presentation concluded with the report's recommendations:

- Aggregation not centralisation of data
- Training and resources to mitigate against technological obsolescence
- Investment in computing resilience
- Ensuring compatibility with TaNC
- The infrastructure should survey and understand users and creators of digital heritage science data
- Developments could take place across 2 phases: landscape mapping and creating connections.

Aurélie Turmel spoke next about Historic Environment Scotland's (HES) ongoing digital infrastructure project, Heritage Hub. This will be an accessible, navigable website that pools all of HES' disparate heritage data, information and knowledge. Aurélie presented some challenges and learnings that HES has encountered during this project:

- Many stakeholders so communication and involvement are a high priority
- Many user groups across the existing websites so extensive and ongoing user research is needed
- Data is held in many locations so this needs to be partially consolidated, focusing on core technologies
- The same people and places may have multiple names across databases so it's necessary to identify connections between data elements and establish unified search
- There are many standards in use across the heritage and GLAM sectors. Heritage Hub has based theirs on Dublin Core.

Heritage Hub will launch in late 2023 and will continue to develop once in operation to keep up with data types and capabilities, and community needs. Before then, the sector is invited to contribute to the user-testing phase by contacting HeritageHub@HES.scot.

Joanna Dunster, Head of (Research) Infrastructure at AHRC, then provided an update on the Digital Research Services branch of RICHeS. This encompasses increasing data repository capacity, and creating an accessible directory of heritage science equipment, collections, expertise and reference data. This will benefit heritage science researchers and raise the profile of the UK heritage science community by establishing a 'front door' for the sector. Joanna acknowledged that there is an outstanding question surrounding the incorporation of legacy heritage science data into the repository but that this cannot be addressed by the current programme. However, the RICHeS Infrastructure Head Quarters will explore opportunities to keep developing the Digital Research Services, such as by funding new approaches to data analysis.

The fourth talk was presented by **Joseph Padfield** (The National Gallery, London). He discussed the European Research Infrastructure for Heritage Science (E-RIHS) DIGILAB platform and its adherence to

FAIR data principles. DIGILAB is a broad-ranging branch of E-RIHS, comprising several different components. For heritage science researchers, its key function is facilitating virtual access to heritage science research data, tools, services, expertise, and collaborative opportunities.

Joseph explained how DIGILAB is founded on FAIR data principles:

- **Findable** – having consistent identifiers
    - Standard PID registries for data, tools, services, equipment, and people
    - Open vocabulary servers
    - Catalogues
- **Accessible** – ensuring metadata is retrievable online
    - Central Core E-RIHS Data Repository is connected to appropriate standardised repositories
    - Agreed data management plans that are 'as open as possible, as closed as necessary'
- **Interoperable** – having common formats and standards so that different systems can talk to each other
    - Agreement on standard formats led by heritage science experts
    - Agreed, open and shared documentation using standardised models and schema
- **Reusable** – optimising data reuse
    - Having clear, appropriate licences so users know how they can use research data
    - Having persistent, citable digital resources for sustainability

The final speaker was **Rob Baxter** sharing learnings from DARE UK's national research infrastructure project for sensitive data. Currently in Phase 1 (Design & Dialogue), Rob described how they undertook an extensive process of landscape review, community engagement, public dialogue and guided research to determine a vision and blueprint for the project. DARE UK is now funding several driver projects to better formulate a UK-wide network of trusted research environments (TREs). Rob explained that one of the main challenges so far has been achieving interoperation between TREs and data providers, at an infrastructure, data and governance level.

The meeting concluded with a question-and-answer session. Discussion points included:

- Standardisation versus variability in digital infrastructures
- Balancing top-down systems and human processes in digital infrastructures
- Meeting community needs with digital infrastructure
- It is more feasible that the final outcome of RICHeS will be a hub of distributed research infrastructures rather than one amalgamated system.
- How the NHSF can support infrastructure development and user involvement. Suggestions included developing an FAQ and best practice guides with accessible use cases, and promoting the strategic value of the RICHeS Digital Research Services to the heritage science community.