DIGITAL HERITAGE (2025)
S. Campana, D. Ferdani, H. Graf, G. Guidi, Z. Hegarty, S. Pescarin, and F. Remondino (Editors)

Broadening Data and Digital Skills within Communities to Access Digital Research Infrastructures

Karina Rodriguez Echavarria† and Myrsini Samaroudi
University of Brighton, United Kingdom

Abstract

This paper advances understanding on how computational and data-driven research methods can be expanded within the Arts and Humanities community, including researchers and practitioners in the Cultural Heritage sector. In particular, the research investigates the catalysts and motivators to enhance the digital and data literacy of researchers and practitioners and recommendations for longer term adoption of these skills. The research is contextualised within current efforts to expand access to national and European Digital Research Infrastructures. It is conducted through a series of scaffolded learning interventions implemented through a training initiative in the United Kingdom (UK). The paper describes this training initiative and the evaluation of the effectiveness of these interventions employing a mixed data approach. The research concludes with a set of recommendations for designing training programmes amongst learning communities, including using skills curricula based on domain-specific data processes and infrastructures as well as active learning approaches.

CCS Concepts

• ***Applied computing*** → ***Arts and humanities;*** • ***Social and professional topics*** → ***Computing education;***

1. Introduction

This research is contextualised within current efforts to expand the use of computational and data-driven research methods, including Artificial Intelligence (AI) methods. These methods, also known as *data driven* or *data science* approaches are gaining popularity across disciplines. The Arts and Humanities (A&H) and Cultural Heritage domain are not exceptions. Such approaches rely on components, including services, facilities and support, within Digital Research Infrastructure (DRI) ecosystems. Thus, they bring together powerful computers, tools, data, techniques, people and skills to underpin cutting-edge research.

Increasingly, digital skill-related initiatives are being implemented. Their aim is to increase researchers' digital literacy. Nonetheless, such initiatives and related activities are often fragmented or even developed in isolation within projects. While they contribute to building digital skills' capacity, they do not provide a holistic approach to removing barriers which prevent communities from fully taking advantage of the opportunities provided by DRI ecosystems. Hence, the research in this paper addresses a gap in understanding the *catalysts and motivators*, as well as the *learning scaffoldings* that influence the long-term adoption of specialised data science approaches by A&H learning communities.

The deployed methodology is based on the study of a series of socialised formal learning interventions. These were implemented through a training initiative in the UK, which was evaluated using a mixed data approach. The contribution of the paper is a set of recommendations for designing data science training programmes amongst A&H learning communities.

The paper is organised as follows: Section 2 contextualises the research, presenting related work in this area. Section 3 defines the research questions and the methodology deployed for their investigation. Section 4 introduces the *Digital Skills in Visual and Material Culture* training initiative, which underpinned the implementation of the training interventions. Section 5 presents the data collected, and Section 6 presents discussions, as well as a set of recommendations for effectively supporting *learning communities*.

2. Related Work

Digital Research Infrastructures (DRIs) is a term increasingly used to refer to digital ecosystems supporting research across disciplines. At its core, they connect communities to data repositories, computing facilities, software codes, as well as specialised skills and cross-disciplinary collaboration to facilitate *data-driven and data science approaches*. Defined as "the science of (collaboratively) generating, acquiring, managing, analysing, performing inference and reporting on data" [Sto20], they offer an interdisciplinary approach to support creativity and innovation.

Common hurdles for users to fully take advantage of these infrastructures are the barriers to accessing DRI and the lack of associated specialised skills. This is caused by a wide range of fac-

† email:K.Rodriguez@brighton.ac.uk

DOI: 10.2312/dh.20253253

tors amongst user communities, including the lack of i) domain-specific digital skills and training resources; ii) engagement with computing services and software to promote learning and experimentation; and iii) easy-to-use workflows, codes/tools and user interfaces beyond command line prompts. The importance of upskilling and empowering heritage researchers and practitioners is well-understood amongst the community. Yet, there is a lack of knowledge on which are the driving motivators for people to engage with initiatives or training programmes, and how to ensure that these programmes deliver efficiently the skills to address the needs of communities across disciplinary, institutional and national boundaries; career stages and paths; while promoting diversity and inclusive approaches.

Meanwhile, there is a broader range of literature on research to convey skills and knowledge to novice communities. Although this subject is often studied within pedagogic and organisational disciplines, such research offers useful insights into how novices interact with experts to enter skilled domains. Whereas apprenticeship models and occupational socialisation (e.g. formal instruction or training programmes) often depict entry as a difficult journey overseen by experts, alternative models outside traditional pathways are being documented by scholars. For example, Dioun et al. [DPG24] show that experts conveyed skills and knowledge by (a) scaffolding to facilitate creation, encouraging skills and accompanying knowledge to be externalised, (b) relaxing hierarchy to reduce distance between themselves and novices to further facilitate their engagement, and (c) cultivating fun through enjoyment.

Furthermore, learning scaffolding structures is researched by *learning communities* in software development communities. A suitable example is boot camps, which use a flipped classroom approach. In this, students are given tutorials to watch, readings and exercises to complete on their own. This format emphasises the primacy of self-learning over learning from an expert teacher. Kaynak [Kay23] observes how these groups, composed of peers and near-peers, learn collaboratively despite limited access to proximate experts through three scaffoldings, including:

- **Peer team structures**, in the form of pair programming and group projects.
- **Near-peer role structures**, in the form of near-peer instructor, teaching assistant and mentor roles, engaging recent graduates in teaching relationships with novices, thus allowing novices to access help quickly and easily.
- **Encouragement to self-learn** by reaching out to the expertise of the broader occupational community.

It should be noted that computing and software-related communities share a philosophy of open access to knowledge, which can partially support these scaffolding requirements. This is evidenced within training and upskilling initiatives for DRIs, as many of the materials are open access.

In the context of the A&H, several skills initiatives have taken place. At a European level, the Digital Research Infrastructure for the Arts and Humanities (DARIAH) offers the DARIAH-Campus which is a discovery framework and hosting platform for learning resources [Diga], as well as the SSH Open Marketplace, which provides access to tools, services, training materials, workflows and datasets for Social Sciences and Humanities research communities [Digb]. The European Research Infrastructure for Heritage Science offers resources including lectures, webinars and training camps [Eur]. Related training includes resources produced by projects such as ARIADNE [ARI], PARTHENOS [PAR] and CARARE [CAR]. Within national EU infrastructures, the NFDI4Culture aggregates educational and training offers on research data management, as well as code and data literacy [NFD].

Meanwhile, international initiatives such as the Programming Historian [Pro] publish digital tools, techniques, and workflow tutorials for non-specialist readers. Beyond the humanities, The Carpentries [The] offer resources for trainers and skills needed to do computational, data-intensive research in a variety of disciplines.

With a focus on Galleries, Libraries, Archives and Museums (GLAM) professionals, the scoping skills programme at the British Library [HMM*22] highlighted the need for appropriate training, promoting research collaboration, communication and open scholarship, while acknowledging the diversity of research activities within GLAMs. Developing specific curricula for the sector has been identified as key to addressing open access, efficient data management practices and use of data infrastructure.

Within these activities, there is evidence of best practices and lessons learned [Wri17, BP22, Wut19]. These refer to:

- Early guidance is essential for training, as learners might not always easily identify what they do not know;
- Incorporating collaborative and hands-on activities in active learning works best to practice skills;
- Balancing online and on-site offerings to cater for different training needs;
- Being flexible to adapt and mitigate unexpected circumstances (e.g. COVID-19 disruptions);
- Designing online training, such as webinars, is useful as they allow distant learning and decrease learning costs;
- Designing targeted on-site or online training events with fewer attendees and specific practice-based topics better suit active learning;
- Supporting individualised project work makes training relevant to learners;
- Offering access to equipment, software and resources lowers the barrier to access training;
- Promoting networking and peer support allows learning to continue beyond training;
- Making available the learning materials ensures that learners can access them.

Embedding curricula in Higher Education settings has also proven to be valuable to provide opportunities for formative and summative assessment, including assignments and evaluation in real-world settings. Practical project work and internships also offer highly rated opportunities to engage more deeply in learning.

As demonstrated in previous initiatives, the integration of concepts, coding skills, tools, and real-world use cases is key to supporting the research ecosystem [EGGS19]. The following sections describe how these insights have informed a series of interventions to better understand motivators, enabling factors and supportive scaffoldings for learning communities.

3. Methodology

This research aims to improve the understanding of how to effectively enhance digital and data skills within *learning communities* in the Arts and Humanities through training initiatives within the context of DRIs. Thus, the research questions are as follows:

- What are the *catalysts and motivators* for learning communities to engage with digital and data skills training?
- How do learning communities identify priorities for digital and data skills training?
- Which *learning scaffoldings* can effectively enhance digital and data skills to underpin research?
- What *enabling factors* are key to the training's short and long-term impact for changing research practices?

These questions were examined through a series of training interventions piloted in the UK. These interventions are described in more detail in Section 4. The activities aimed to transform novices or relatively novice learners into competent creators and users of multidimensional digital media assets.

To answer the research questions, the training was designed to include training activities within formalised University educational settings; and non-formalised settings, including workshops, webinars and internships. Before the training, a survey and round table discussions were used to collect the insights and requirements from a wider number of stakeholders.

Thereafter, pre-training and post-training surveys were deployed with learners attending training activities to examine the overall experience, quality, and effectiveness of training. This included semi-structured interviews with learners who participated in internships within the training programme. Feedback from the hands-on workshop learners was also collected a year after the completion of the training. This allowed us to gain insights into the impact of the training in terms of changing professional practices and passing on knowledge to others.

The research opted for non-probability convenience sampling. Feedback was gathered from those available, involved with the research either at its initial stages or later through the scaffolded training activities. The following section introduces the training initiative, its curriculum and the activities which were piloted.

4. Digital Skills for Visual and Material Culture

The research piloted a training initiative within the UK-DRI ecosystem branded as the *Digital Skills for Visual and Material Culture*, focusing on supporting practitioners and researchers in Galleries, Libraries, Museums, Archives, Universities and the creative industries. The training activities focused on skills related to the digitisation, processing, use/reuse, data management, preservation, and linkage of multidimensional (2D/3D) digital media assets.

As stated in section 3, the conception of the training programme was informed by quantitative and qualitative data collection, including a survey and interviews, to gather requirements from the Arts and Humanities communities. The survey was advertised amongst the Digital Humanities departments and research centres, as well as with researchers and practitioners in GLAMs. The results informed the design of the training initiative.

The data collected included participants' background, working practices with multidimensional media assets, domain-specific data processes and perceived training needs. The research also sought to investigate communities' requirements and future aspirations in terms of infrastructures which will support effective and FAIR -Findable, Accessible, Interoperable, Reusable- management of multidimensional media.

The responses highlighted that almost 70% of researchers use and manage 2D/3D media in research projects without or with minimal training. In addition, at most 30% of researchers have received formal and informal training, mostly in digitisation, but still lack data literacy skills to manage, analyse, visualise and provide access to the data. Further evidence corroborates the lack of confidence within the community with respect to knowledge of digital methods and domain-specific data processes, as well as broader enabling knowledge. The former includes technical knowledge such as linking data, documentation with metadata and paradata, data analysis, visualisation, and management of digital assets. The latter includes best practices for data preservation, intellectual property rights, copyright, ethical considerations, and organisational methods/tools for collaboration in ways which boost knowledge exchange and sharing. An emerging need is access to opportunities and development of digital literacy, while taking into account diversity, inclusion and principles around digital data ethics [Gre19].

The proposed training programme was designed across four key components : 1) *skills curriculum*, 2) *learner communities*, 3) *scaffolded training*, and 4) *infrastructures for training*. These are described in the following subsections.

4.1. Skills Curriculum

Informed by Stodden [Sto20], the curriculum identifies skills surrounding data ecosystems within A&H research. Thus, the research considers the distinctive nature of A&H research. Whereas knowledge in other disciplines is more cumulative and rooted in discovering universally applicable insights; A&H researchers seek to understand human experience, agency, identity, and expression, as constructed through language, literature, material culture and performance amongst others. This means that very few branches of the A&H exhibit properties of a scientific mode of knowledge, which is predictive and universalisable; hence, easily reproducible. Instead, A&H research is particularistic and interpretive. Its interpretative nature often requires the close reading and critical analysis of a number of *texts*, forming an inheritance that can be reevaluated from new perspectives and new contexts continuously.

Therefore, digital data ecosystems in the A&H are often related to human constructs, evidenced, for instance, by literature, material culture and increasingly digital media. In this domain, digital methods are used to digitally collect, analyse and interpret these constructs. In doing so, data is created by researchers and others who continue to bring new perspectives and contexts.

This research deploys a multi-level approach to conceptualise skills for data-driven research life cycles within the A&H [Sto20]. This framework, illustrated in Figure 1, makes explicit the different steps in the discovery pipeline, including meta-level considerations for research, which is domain agnostic, such as ethics and IPR.

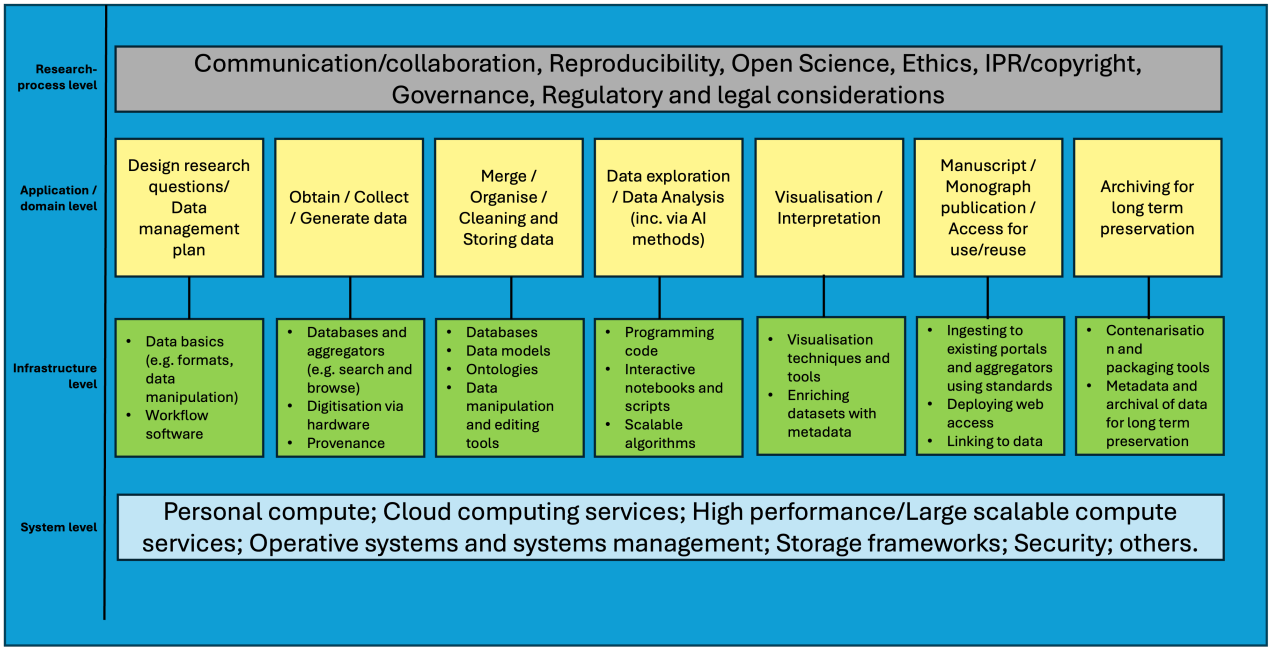

Figure 1: *Data-driven research life cycle adapted from [Sto20]. Skills are identified across four levels, including: 1) Research level - broader considerations for conducting research; 2) Domain level - example of research process in a digital humanities project; 3) Infrastructure level - software code, tools, workflows; 3) System level - computer and services on which the infrastructure runs.*

Others are domain- or even sub-domain-specific. Figure 1 lists an example of steps which a research project in the A&H or CH domain might require. Through these steps, a researcher will collect and/or generate data. The data will then be organised, analysed, enriched and interpreted until the research findings are published and the data is made accessible and/or archived for others to use/reuse and for preservation. The infrastructure and system levels identify the software code, tools, workflows, as well as computer systems and services on which the infrastructure runs.

In large projects, skills are spread across different roles and people, whereas smaller projects require skills to be concentrated in one or two people. This creates additional challenges for domains such as the A&H where lone researchers or smaller teams are prevalent. Thus, it creates a more challenging environment for the individual researcher as the barriers to entry are much higher.

4.1.1. Syllabus

The proposed training initiative concentrated on skills related to research processes underpinned by data related to collections and archives of the GLAM sector, including research related to their interpretation, preservation, as well as creative research practices within disciplines such as the visual arts, design, media studies and performance. Table 1 presents examples of Syllabus topics for the training programme.

With regards to the data, the focus was on multidimensional (2D/3D) media datasets, which are a common data type for documenting material culture. Datasets are often created as large-scale, multi-part and multi-format assets , making their management complex. Such datasets include born-digital data (i.e. produced via software tools) and data produced by digitisation processes. Examples include 2D photographic images (within and beyond the visible spectrum), 360 photography and video, film, 3D digital surrogate models, objects/environments produced through 3D modelling, as well as interactive digital environments for Mixed Reality (i.e. Virtual and Augmented Reality).

4.1.2. Levels of Competence

The targeted level of competence by the workshop was informed by national and international frameworks [VKP22, PKM*21]. The different levels were also influenced by frameworks in education and pedagogy, such as Bloom's taxonomy [Blo56] and Bigg's SOLO taxonomy [BC82]. These classify competencies and cognitive functions based on increased complexity. The levels used are:

- *Foundation*: Is aware of the skill and has an understanding of how it could be applied in a practical situation.
- *Intermediate*: Is aware and remembers the skill and how it could be applied in a practical solution.
- *Advanced*: Applies knowledge and has working experience of the skill, including tools and techniques. Adopts the most appropriate for the situation.
- *Highly specialised*: Holds specialist knowledge/expertise and critical/analytical skills to assess methods, practices and tools and guide on appropriate solutions.

The developed skills' curriculum targeted the novice or relatively novice learners, belonging to the foundation or intermediate level as described in the chosen set of competencies.

Data Life Cycle Step	Topics
Research Process Level	
	- **Copyright and licensing** - **Ethics** of multidimensional data - **Open-science** approaches
Application / Domain Level	
Design and data management	- **Research methods**: Digital methods for research in A&H - Introduction to **FAIR multidimensional media** - Planning for **data management**
Data creation and processing	- **Retrieving and curating datasets** - **Digitisation** of objects, environments via photogrammetry, RTI, 360 photography, etc.
Organising and storing data	- **Structured data**: metadata and paradata for CH collections
Data exploration and analysis	- **Programmatically manipulating data**, inc. various scales - **Artificial Intelligence driven methods** for analysis, classification of multidimensional media
Access for use/reuse of data	- **Digitally-enabled access**, inc. museum exhibition, research outputs - **Supporting organisational strategies**, including repatriation, decolonisation and environmental efforts
Archiving	- **Archival and long-term preservation**

Table 1: *Example topics for the Data and Digital Skills Curriculum in 2D/3D media for Visual and Material Culture research.*

4.2. Learner Communities

Known in the literature as *Learning Collectives* [Kay23], the training initiatives targeted researchers, postgraduate students, academics, and practitioners within Higher Education Institutions (HEI), Galleries, Libraries, Archives, Museums (GLAMs), and the creative industries sector. The availability of training was widely advertised through mailing lists and networks to reach learners' communities.

The training did not have an associated cost, beyond that of time and travel. In most cases, a selected number of bursaries were made available to those who might face difficulty in affording travel costs.

4.3. Scaffolded Training

Scaffoldings refer to the temporary support structures that learners experience to get to the next stage or level of their learning. Common scaffolding structures within socialised learning in both formal and informal contexts [Abb14, BB10] include *near-peers'* structures [Kay23] where learners within the same learning community socialise, share their experiences, provide guidance and support and overall act as role models for others.

Informed by these approaches, the following types of learning activities were implemented across one year during 2023-2024.

4.3.1. Webinars

Webinars focused on improving virtual attendees' awareness about topics such as AI, publishing and sharing complex visual data, digital storytelling, the development of digital outputs within the context of research assessment frameworks and more. Domain specialists with a wealth of experience in the skills-related topics delivered the one-hour webinars. Webinars were interactive, allowing time for questions and answers. Digital recordings were made available through the website for further access.

Webinars had the highest level of engagement, reaching out to national and international organisations. In total, four hundred seventeen (417) people registered for the webinars. The most popular webinars (with the highest number of registrations) covered topics around *Publishing and Re-use of Cultural Heritage 3D for Narrative Purposes* and *Artificial Intelligence Applications for Visual Archives*. While a high number of registrations were noted, the actual number of people who attended the webinars was around half the number of registrations (49% actually took part). The webinars proved valuable in terms of engaging broader audiences.

4.3.2. Socialised formal and informal learning through hands-on workshops

This included 1) half-day workshops for early career museum professionals enrolled in the MA module in Curating Collections and Heritage, as well as ii) 2 and 3-day workshops open to a wider number of learners, covering advanced knowledge on the data science life cycle of multidimensional media. Scaffoldings used in these workshops included: i) open access training resources in *The Carpentries* format; ii) prepared datasets, software tools and workflows for attendees to use; iii) peer-to-peer support implemented through hands-on experimentation with digitisation devices, tools, workflows and hardware.

Workshops were delivered in the following topics: Accessing and Interacting with Multidimensional Media, Advanced 3D Digital Documentation of Material Culture, Capturing Heritage Craft in 2.5D and 3D, Digital Skills in Visual and Material Culture (see Figure 2). The workshops included approximately 4-5 *sessions*, scheduled over 2 or 3 days. Each session lasted between 1.5 and 3 hours, including a theoretical presentation delivered in a classroom style and a hands-on exercise. Learners were encouraged to bring material relevant to their own research for the hands-on exercises. In case this was not possible, sample datasets were provided.

Various instructors with relevant expertise were involved in the

Figure 2: *Workshops' posters covering various training topics.*

delivery of each session, with most sessions being supported by at least two instructors. The majority of the instructors had a post-graduate level in relevant topics and had significant experience in deploying the methods within Cultural Heritage contexts.

Each session was complemented by an online training resource in the Carpentries style, adopting the Skills4EOSC - Skills for the European Open Science Commons approach [FGM*23]. The learning objectives were defined first, and then the content was generated to comply with such goals. Open and FAIR principles were also considered at each stage of the development. Therefore, training materials and datasets were made available under an open-access license. Additionally, the resources were available to learners both as a website and software code that learners could access and adapt to their own needs before, during, and after engaging with the training activities. These resources cover the topics of the curriculum including digitisation [RESS24] and FAIR data [RE23].

Furthermore, the data and code related to the training resources and datasets were curated with metadata to enable the discoverability of the resources. Metadata schemas [HBB*22, FGM*23, ORL*23, BLS*22] and related controlled vocabularies to outline the structure of information were extensively studied and mapped to ensure appropriateness for the scaffolded training resources. Hence, the deployed metadata schema complies with the RDA Minimal Metadata for Learning Resources [HBB*22]. Metadata included information such as title, abstract, learning outcomes, keywords, date of creation, license, URL, prerequisite knowledge level, and targeted expertise level.

The resources were made available via the project website and through DARIAH campus (https://campus.dariah.eu). The campus is a discovery framework and a hosting platform for DARIAH and DARIAH-affiliated offerings in Arts and Humanities training.

4.3.3. Individual learning through internships

Internships are aimed at providing learners with the opportunity to become involved in work, developing their digital skills within a professional setting.

Internships usually lasted 70 hours (or 2 weeks) with flexibility as to how/when these were organised. Internships were supported by an organisation, mostly a research team within a University, and a heritage institution or a company, which acted as a host. A supervisor within the organisation defined the overall aims of the internship, as well as its deliverables and supported the advertisement and the selection of the interns. Interns were not paid, but a stipend was provided to support the project. There was flexibility as to how this stipend was used, for instance, to pay for travel, digitisation, or other expenses related to the project. A hybrid setup allowed interns to spend some time at the organisation and work from home, enabling further flexibility. Interns were supervised by experts within the organisation, company or heritage institution. The internship often involved gaining new insights into the life cycle of key datasets for the organisations.

4.4. Infrastructures for Training

An infrastructure based on D4Science (https://services.d4science.org) was available to learners. This is a Virtual Research Environment that includes digital tools and shared data spaces. It provided access to the datasets, collaborative editing and JupyterLab to ensure learning communities could train together while supporting individual contexts.

5. Data Analysis

As outlined in Section 3, a mixed-methods approach was used for the analysis of data collected from the *learning communities* who engaged with the training activities [RL14]. The following subsections will present the data for the scaffolded learning activities.

5.1. Evaluation of socialised formal and informal learning through hands-on workshops

Quantitative and qualitative data were collected through a survey before and after the training activities. The provision of feedback was not mandatory. Twenty (20) learners provided responses before the training. All of them were working in A&H-related projects, and amongst them, the vast majority were women (15). Thirteen (13) of them were between 25 and 39 years old, followed equally (2 participants per age group) by "mature" groups, including those

between 40-49 and 50-59 years old. Two (2) more participants belonged to younger or older age groups.

All of them, apart from one learner, had completed formal education and were holders of degrees or postgraduate degrees. Fourteen (1) learners had research roles either as undergraduate, postgraduate or academic staff. Four (4) learners were practitioners, and two (2) identified both as practitioners and researchers with responsibilities including digital content creation, exhibition design, educational programmes and digital curation.

Most learners (11) reported their interest in expanding their own digital skills as a reason to attend the training, while others also mentioned an interest in teaching others. In the majority of the cases, learners were motivated to deploy those skills in their research and working practices. Few of them were motivated by the idea of improving their employability or by personal interest in the topic.

The findings demonstrate that learners' expectations, once they made the decision to attend the training, included the ability to access practical hands-on knowledge relevant to their context; while being equipped with transferable skills coupled with fundamental knowledge about methods and best practices. Learners also expected that the training will boost their professional development and capacity to adapt in a rapidly changing technological environment. The fact that training was offered free, as well as the provision of all necessary equipment and software, was highly valued. The clear way the training activities were structured and communicated was also very well received by the learners.

To better understand learners' digital literacy, people were asked if they had received training in the past on the given topic. Most learners stated that they had not received relevant training (16 people). Thus, these learners reported their competence level as basic/foundation, while one reported it as intermediate, two as advanced and one as being highly specialised. Common obstacles that learners reported for not undertaking training before included not knowing where to reach for training, not knowing what kind of training they needed, and encountering too many offerings with no clear structure.

After the completion of the training, twenty-two (22) learners provided quantitative and qualitative feedback. When asked about understanding the purpose of the training before this took place, all learners stated that this was communicated to them in a clear way by saying that*"All the information provided during the application process as well as the content previously sent to us participants were very precise about the purpose of the workshop"* or *"The workshop information on the website was very clear"*. Most of them (14 people) stated that they had enough time to prepare, while eight (8) people stated that they could not prepare as they wished, mainly due to dealing with other responsibilities or deciding to join last minute.

When examining how well the experience of the training matched the information and intent of it, as provided to learners before the training took place, seventeen (17) people rated it as "extremely" or "very well", while five (5) thought that this matched "quite well". Qualitative comments reveal points of improvement with regard to providing more information to learners before the training, referring to more detailed descriptions of hands-on exercises, better pre-training preparation of individual project materials, clearer software requirements and provision of more pre-training knowledge materials and optional tasks.

Moreover, learners provided positive feedback on the training being delivered via a hands-on workshop. Positive points highlighted by the learners include the provision of both "theoretical" and practical parts (*"workshop is for both practical and theory lessons"*) and the collaborative nature of learning (*"The delivery allowed for a collaborative approach where we came together and worked"*). The face-to-face format was positively assessed. A plethora of comments from the learners corroborate the importance of hands-on in-person learning: *"It might be able to deliver it also online, but face-to-face is much better. Two full days and it was not difficult to follow, something that might be difficult if it was online"; "The face-to-face training was very appropriate for this kind of workshop because it allowed me to have direct supervision and prompt feedback from the facilitators"; "Face-to-face is the best way to teach this sort of research, it's hands-on and physical"*.

A majority of learners (14) expressed they would have liked to see more time dedicated to the workshop to allow for training at a slower pace. Learner's comments reveal that they needed more time to go through the materials, deepen understanding based on enhanced explanations, improve work with their data, practise software and receive guidance from colleagues. They also highlighted that they would need more support to execute processes themselves instead of having these demonstrated by the instructors.

Another aspect of the training assessment was the learning resources and infrastructure. Most learners found them easy to access, easy to understand in terms of structure and language, interesting, practical and helpful. Learners commented positively on the availability of clear, comprehensive and accessible materials (*"Very helpful, all the information are available in one place"*), as well as the provision of the software code (*"The github is very helpful! Thank you so much for building the page. It allows me to easily share it with my peers and teammates"*). When it comes to further important aspects or suggestions for improvement, the learners pointed out the need to sustain materials and associated training resources online, as well as including more information on the "basics" to support learners with more fundamental knowledge before proceeding to more complex processes.

Overall, learners were very satisfied with the practical, interactive and applicable character of the training; the quality of the resources; the opportunities to network and collaborate; the supportive and friendly environment; and the opening of new perspectives with regard to their own research and ongoing learning. On the other hand, suggestions to improve the training resonate with findings about the provision of information prior to training, the pace of the training/s and the access, structure, content and delivery of the materials.

After the completion of the training, when learners were asked to self-assess literacy in the relevant digital skills, they clearly acknowledged that their skills had improved. In addition, learners' future aspirations communicated their desire to continue improving skills; applying the knowledge to their work/research; pursuing professional goals, training others and creating educational re-

sources; while also advocating for the use of technology; discussing and co-creating with the communities they work with.

A year after the completion of the training, learners were contacted for further input. The main areas for feedback were related to a. how research/work practices might have changed by using new methods, workflows, outputs; and b. sharing of knowledge with others. In total, nine (9) people responded to the request for feedback. Learners stated that the training increased their awareness with regard to the potential of technologies (*"The training was an eye opener on the potential of technology for documenting material culture. It made me think differently about how I could enhance my research, especially in terms of dissemination and reaching non-academic audiences"*). Others mentioned that their work/research was transformed, since they were able to put in practice new skills (*"The training has significantly benefited my research by enabling me to create 3D images of the containers I am documenting"*).

Furthermore, some people reported continuous progression in terms of advancing their skills by exploring new workflows and integrations (*"Since the workshop, I also started working on research into the integration of archaeological photogrammetry of material culture into videogames"; "Tools like model annotation and visual storyboarding have become integral parts of our final outputs, enhancing our work beyond just photos and videos"*), advancing their career (*"I am now going to be pursuing a master's in digital archaeology next September to continue developing my skills"*) or becoming "reference points" within their organisation for using the acquired digital skills and sharing knowledge with others (*"Nowadays, I can be a reference in sharing knowledge about 3d digitalisation at the university, and also at meetings with indigenous teachers"*). Such developments seem to have opened novel or further research opportunities (*"The best way of showing the traces of the tool used to create a pot is by 3D imaging. It is now something we also encourage other pottery researchers to use"*), while enabling researchers to provide better access and engage audiences with their work (*"We incorporated 360-degree photography extensively, which allowed us to create a more immersive experience for our audience"*). Apart from the positive developments, difficulties in applying the acquired skills were mostly related to time constraints, lack of access to suitable infrastructure in their own environments, and lack of resources.

Evidence to investigate impact with regard to sharing the knowledge with others reveals that learners have been mostly sharing these skills with colleagues in informal or near-peer structures (*"I have primarily shared these skills with colleagues working on my container documentation project"; "I have helped several of the 3rd year archaeology undergraduate students with the integration of photogrammetry into their dissertations. This has been done mostly during informal study sessions"*). Such mechanisms of knowledge transfer can be potent and direct in their ability to empower researchers and practitioners in the use of digital skills, while also promoting research collaboration and innovation. As another learner pointed out *"The lessons learned during the workshop have significantly influenced our organisation's approach and practices... This shared learning experience has not only enhanced our individual skills but has also contributed to a more cohesive and effective work environment for everyone involved."*

Fewer learners, have organised or plan to organise formal training events for their networks or communities (*"I'm currently involved in training indigenous teachers and the topic of 3D digitalization was part of one of our meetings"; "We want to do a formal event at university about these new learned techniques"*).

5.2. Evaluation of individual learning through internships

After the completion of the internships, two interns were interviewed to provide further insights about their experience. An academic who provided the internship opportunity within their project and supported the intern with mentorship was also interviewed.

All three (3) interviewees had foundation skills within their research context and professional practice to use and manage 2D digital media, including searching digital archives and image repositories. Their motivation to engage with training revolved around benefiting their own research and work context and generally broadening interests and potential, such as enabling new engagements with heritage. All interviewees had also had previous training, which touched upon data management, but often with a different focus. Previous learning had happened via formal education, individual study, and professional networks and other training activities, such as seminars. Some of the difficulties that the interviewees had encountered when trying to attend training before include lack of time, not knowing what to look for or not finding specific training opportunities (*"it was just me who couldn't find them or maybe they weren't. There wasn't any training available"*) or the cost of training itself (*"It would be having to pay an amount that I'm just not able to do or I have paid which is just, you know, unsustainable to do regularly"*).

The interns provided positive feedback about their experience, which contributed to improving their digital skills, focusing mainly on data planning/management, including data archiving and preservation (*"Metadata and controlled vocabularies... I'm aware of those things, but I've never really consciously understood what I've been doing. So this is definitely the first time that I've had a sense of how important this is"; "It has raised my awareness of the complexities around copyright"*). Moreover, the "exposure" of interns within creative industries brought additional benefits, such as providing new insights on other heritage related processes within the sector (*"Relevant is that I became also more aware of how films should be conserved"*). The internships' supervisors also reported benefits, noting that the internships enhanced their own data management skills, as well as an understanding of technologies for viewing and interacting with visual data (*"it was a catalyst for me, dedicating some time to it"*, *"this was an opportunity for me to put that into practice and design a kind of catalogue of my own"*).

When examining the overall experience, all involved parties appreciated the support and training provided through mentorship and near-peer structures. Such support, in the form of training meetings, project-based discussions, and provision of documents and resources to self-study, was highly appreciated. The flexibility of training through internships was also valued, as the hybrid format enabled interns to spend time at the host organisation, their workplace or at home, working in a flexible manner (*"And also I'm able to kind of work at different times... in the evening and stuff like*

that"; *"I found it very efficient, like the first weeks.. dealing with the photographs and speaking with the designers at the company. And (then) working on my own to try to help them reach their goal"*). Yet, some difficulties were also encountered. From the mentor's perspective, these were related to the time allocation to support the intern and the balance between training and internship outputs. From the interns' perspective, issues were encountered regarding various support systems (e.g. university payment platforms), platforms to access training content (e.g. permission issues) and time allocation.

6. Discussion and Conclusions

This research provides a holistic understanding of the catalysts and motivators to enhance digital and data literacy, as well as the enabling factors for longer-term adoption of digital methods and tools. In turn, such understanding can support designing, developing and implementing effective training programmes.

Acknowledging the distinctive, interpretative and evolving nature of A&H research, the research identifies a sense of *agency* and *self-reliance* as catalysts for long-term adoption of data and digital skills amongst learners. Whereas supporting research and working practices, professional development and career advancement motivates people to seek out training, these efforts are often hampered by not knowing what to look for and where to access training. Affordable, inclusive and effective *scaffolded learning*, including clearly structured and communicated *skills curricula*, as well as effective learning methods, training infrastructure and FAIR resources, are also key. Also, expanding and democratising *access to research infrastructures*, through which learners can continue to build capacity past the training and share their knowledge with others, are important to ensure the long-term impact of training efforts. The research also identifies a set of recommendations for the planning, design, implementation and aftermath of scaffolded training activities.

Recommendations to consider pre-training

- Design training based on **the targeted community's research processes** and priorities.
- Design **contextualised learning** activities where learners work toward producing an outcome that is relevant to their work.
- Incorporate **disciplinary or sector advancements**, examples and best practices into the learning resources and materials.
- Offer a **variety of content**, which covers both foundational, theoretical knowledge and practical learning.
- Consider offering **pre-training consultation** or **information sessions** offering learners opportunities to self-assess their needs, prepare for the upcoming activity/ies and address questions.
- **Advertise training** well in advance through relevant platforms.
- Honestly describe **clear details about training**, including purpose, learning objectives, prerequisite knowledge and structure to ensure learners form realistic expectations about the training.
- Suggest **optional pre-reading resources** to enable learners to spend more time in practising skills during the training.
- Design **learning resources following FAIR principles**, documenting them with suitable metadata (e.g. RDA Minimal Metadata for Learning Resources [HBB*22]) to increase findability, accessibility, interoperability and re-use of the materials.
- Integrate **progressively complex exercises** or challenges to favour learning-by-doing, while employing constructive alignment principles.
- Where possible, favour software with low or no cost, provide access to sample datasets, while providing flexible equipment setups to **adapt to different budgets and infrastructures**.
- Clearly communicate **hardware/software requirements** and ensure software **compatibility** before the commencement of the training activities.
- Establish an **evaluation plan** in advance to examine training details, delivery, effect and learners' aspirations.

Recommendations to consider during training

- Include **practical, hands-on training**, as it promotes active learning, while enhancing learners' skills and boosting their confidence. While webinars might not offer the same ground for interactive learning, it is possible to include interactive elements and provide examples for individual experimentation.
- Provide a **friendly and supportive environment**, where there are opportunities to engage with collaborative peer-to-peer and near-peer work, as well as active instructor feedback.
- Ensure instructors' **learning and teaching style is consistent** and works harmoniously across all training activities or sessions.
- Allocate **plenty of time** to engage with activities and prioritise answering questions and practical work.
- Gather **feedback** to assess the training activities.

6.1. Post-training

- Provide support via **communicative/collaborative services** and platforms to enable multi-directional feedback, exchange of ideas and skills' practice.
- Offer opportunities for **follow-on learning**, either focusing on specific topics or providing advanced skills' training.
- **Act on learners' feedback** to identify and address comments about training access, delivery, learning materials, resources and the overall quality of training.
- Regularly **update learning resources** with current, topical information, best practices and examples to maintain relevance and speak to the sector's needs.
- Support efforts and initiatives to **strengthen** policy formulation, strategies and best practices to increase the uptake of **data science approaches** and DRI.

Drawing on the experience of the training and the findings of the research, further work includes standardising curricula for data and digital skills by expanding the proposed framework (Figure 1) and vocabularies [BHK*21]. Enabling researchers and organisations to self-assess their digital skills and competences with the aim of easily locating further development areas, which in turn will aid in shaping organisational policies and strategies, is also a key focus.

7. Acknowledgments

This work was supported by the Arts and Humanities Research Council (AHRC) [AH/X007626/1]. We acknowledge all the attendees to training events who provided feedback.

References

[Abb14] ABBOTT A.: *The system of professions: An essay on the division of expert labor*. University of Chicago Press, 2014. 5

[ARI] ARIADNE: ARIADNE training resources. http://legacy.ariadne-infrastructure.eu/training-materials/. Accessed: 2025-03-17. 2

[BB10] BAILEY D. E., BARLEY S. R.: Teaching-learning ecologies: Mapping the environment to structure through action. *Organization Science 22* (2010), 262–285. doi:10.1287/orsc.1090.0511. 5

[BC82] BIGGS J., COLLIS K.: *Evaluating the quality of learning: The SOLO taxonomy*. Academic Press, 1982. 4

[BHK*21] BOREK L., HASTIK C., KHRAMOVA V., ILLMAYER K., GEIGER J. D.: Information organization and access in digital humanities: Tadirah revised, formalized and fair. In *Information between Data and Knowledge*, vol. 74 of *Schriften zur Informationswissenschaft*. Werner Hülsbusch, Glückstadt, 2021, pp. 321–332. Session 5: Knowledge Representation. URL: https://epub.uni-regensburg.de/44951/. 9

[Blo56] BLOOM B. S.: *Taxonomy of educational objectives, handbook I: The cognitive domain*. David McKay, 1956. 4

[BLS*22] BRAUKMANN R., LEENARTS E., SALDNER S., KECK V., WEHMEYER J., ALBUERNE A., ILLMAYER K., DURCO M., KALAITZI V., YANKELVICH T., FISER D., CVEK A., VLAD A.: *Deliverable 6.11 SSHOC Trainer Toolkit (final)*. Tech. rep., SSHOC Social Sciences & Humanities OPen Cloud, 2022. URL: https://doi.org/10.5281/zenodo.6564283, doi:10.5281/zenodo.6564283. 6

[BP22] BASSETT S., POTENZIANI M.: *D11.1 Quantity of Access Offered-TA4 Title and Organisation*. Tech. rep., ARIADNEplus, 2022. URL: https://doi.org/10.5281/zenodo.7651114, doi:10.5281/zenodo.7651114. 2

[CAR] CARARE: CARARE Training Hub. https://www.carare.eu/en/training/. Accessed: 2025-03-17. 2

[Diga] DIGITAL RESEARCH INFRASTRUCTURE FOR THE ARTS AND HUMANITIES: DARIAH Campus. https://campus.dariah.eu/. Accessed: 2025-03-17. 2

[Digb] DIGITAL RESEARCH INFRASTRUCTURE FOR THE ARTS AND HUMANITIES: Social Sciences & Humanities Open Marketplace. https://marketplace.sshopencloud.eu/. Accessed: 2025-03-17. 2

[DPG24] DIOUN C., PAMPHILE V., GORBATAI A.: 'Making it Easy to Do Hard Things':How experts help novices perceive craft as accessible. *Organization Studies: Special Issue: Rediscovering and Theorizing Craft in Organization Studies* (2024), 1–25. doi:10.1177/01708406241282126. 2

[EGGS19] EDMOND J., GARNETT V., GOULIS H., SCHUSTER K.: *D7.5 Embedding PARTHENOS Training into Higher Education Curricula Report on activities of Task 7.4*. Tech. rep., PARTHENOS, 4 2019. URL: https://doi.org/10.5281/zenodo.3356752, doi:10.5281/zenodo.3356752. 2

[Eur] EUROPEAN RESEARCH INFRASTRUCTURE FOR HERITAGE SCIENCE: ERIHS Academy. https://www.e-rihs.eu/hs-academy/. Accessed: 2025-03-17. 2

[FGM*23] FILIPOSKA S., GREEN D., MISHEV A., KJORVEZIROSKI V., CORLETO A., NAPOLITANO E., PAOLINI G., GIORGIO S. D., JANIK J., SCHIRRU L., GINGOLD A., HADROSSEK C., SOUYOULTZOGLOU I., LEISTER C., LAZZERI E.: *D2.2 Methodology for FAIR-by-Design Training Materials*. Tech. rep., 2023. URL: https://doi.org/10.5281/zenodo.7875541, doi:10.5281/zenodo.7875541. 6

[Gre19] GREENHALL M.: Digital scholarship and the role of the research library: The results of the rluk digital scholarship survey. https://www.rluk.ac.uk/wp-content/uploads/2019/07/RLUK-Digital-Scholarship-report-July-2019.pdf, 2019. [Accessed 19-July-2022]. 3

[HBB*22] HOEBELHEINRICH N., BIERNACKA K., BRAZAS M., CASTRO L. J., FIORE N., HELLSTRÖM M., KUCHMA I., LAZZERI E., LEENARTS E., LAVANCHY P. M., NEWBOLD E., NURNBERGER A. L., PLOMP E., VAIRA L., GELDER C. V., WHYTE A.: *Recommendations for a minimal metadata set to aid harmonised discovery of learning resources*. Tech. rep., 2022. URL: https://doi.org/10.15497/RDA00073, doi:10.15497/RDA00073. 6, 9

[HMM*22] HOLT A. I., MILES S., MARPLES A., KAUR K., COPE J.: *Scoping Skills and Developing Training Programme for Managing Repository Services in Cultural Heritage Organisations*. Tech. rep., British Library, 2022. URL: https://doi.org/10.23636/nyee-xt63, doi:10.23636/nyee-xt63. 2

[Kay23] KAYNAK E.: Leveraging learning collectives: How novice outsiders break into an occupation. *Organization Science 35* (2023), 948–973. doi:10.1287/orsc.2020.14214. 2, 5

[NFD] NFDI4CULTURE: Educational Resource Finder of the Cultural Research Data Academy (CRDA) . https://nfdi4culture.de. Accessed: 2025-03-17. 2

[ORL*23] ORLANDI S., ROVERSI R., LONGO D., TURILLAZI B., FLOHR P., VAN HORIK R., CASSIDY C., FERNIE K., DRITSOU V.: *Deliverable 4.2 Report on service deployment and training*. Tech. rep., 4CCH4 Competence Centre for the Conservation of Cultural Heritage, 2023. URL: https://doi.org/10.5281/zenodo.7898350, doi:10.5281/zenodo.7898350. 6

[PAR] PARTHENOS: PARTHENOS training. https://training.parthenos-project.eu/. Accessed: 2025-03-17. 2

[PKM*21] PLUMMER D., KEARNEY S., MONAGLE A., COLLINS H., PERRY V., MOULDS A., MROCZKA M., ROBERTSON J., SMITH T., TREWAVAS S.: *Skills and Competency Framework Supporting the development and adoption of the Information Management Framework (IMF) and the National Digital Twin*. Tech. rep., CDBB, 2021. doi:https://doi.org/10.17863/CAM.65347. 4

[Pro] PROGRAMMING HISTORIAN: Programming Historian . https://programminghistorian.org/. Accessed: 2025-03-17. 2

[RE23] RODRIGUEZ ECHAVARRIA K.: Fair multidimensional media: Version 1.0.0, Nov. 2023. URL: https://doi.org/10.5281/zenodo.10212738, doi:10.5281/zenodo.10212738. 6

[RESS24] RODRIGUEZ ECHAVARRIA K., SAMAROUDI M., SCHIAVOTTIELLO N.: Universityofbrighton/2024-3d-digitisation-methods: Version 1.0.0., July 2024. URL: https://doi.org/10.5281/zenodo.12687073, doi:10.5281/zenodo.12687073. 6

[RL14] RITCHIE J., LEWIS J.: *Qualitative research practice : a guide for social science students and researchers*. Sage Publications, 2014. 6

[Sto20] STODDEN V.: The data science life cycle: a disciplined approach to advancing data science as a science. *Commun. ACM 63*, 7 (June 2020), 58–66. doi:10.1145/3360646. 1, 3, 4

[The] THE CARPENTRIES: The Carpentries . https://carpentries.org/. Accessed: 2025-03-17. 2

[VKP22] VUORIKARI R., KLUZER S., PUNIE Y.: *DigComp 2.2 - The Digital Competence Framework for Citizens*. Tech. rep., European Union, 2022. URL: https://ec.europa.eu/jrc, doi:10.2760/115376. 4

[Wri17] WRIGHT H.: *D5.1: Report on Transnational access activities and training activities*. Tech. rep., ARIADNE, 2017. URL: http://legacy.ariadne-infrastructure.eu. 2

[Wut19] WUTTKE U.: The "parthenos training webinar series": Webinars as a means of delivering successful research infrastructure training in ehumanities and eheritage. *LIBER Quarterly 29* (2019), 1–35. doi:10.18352/lq.10257. 2