Alessandra Esposito,[1] Pascal Flohr,[2] Shatha Mubaideen,[3] James Smithies,[4] Fadi Bala'awi,[5] Carol Palmer,[6] Sahar Idwan[7]

Digital Assets for the Study of Jordanian Heritage in the Nabataean and Roman Periods: The Data from the Madih (حيدم) Project

ABSTRACT: This study explores digital assets related to Jordan's Nabataean and Roman heritage, focusing on datasets from the MaDiH (Mapping Digital Cultural Heritage in Jordan) project. The MaDiH initiative aimed to assess Jordan's digital cultural heritage (DCH) landscape in order to inform research infrastructure policies and improve access to archaeological and historical data. Jordan has emerged as a regional leader in DCH development, with collaborations between the Department of Antiquities and international institutions yielding significant databases such as MEGA-Jordan and JADIS.

The research evaluates 133 datasets specific to the Nabataean and Roman periods, assessing their accessibility, ownership, and usability. These datasets primarily cover tangible heritage, including archaeological sites and objects, with limited representation of intangible cultural aspects. The majority exist in digital formats, yet accessibility challenges persist due to restrictive licenses, offline storage, and limited public availability. Ownership is primarily concentrated in universities, museums, and heritage institutions, but a significant portion remains in private hands.

The study identifies disparities in dataset accessibility, whether online or offline, depending on location, and highlights concerns regarding local access to national heritage. Additionally, language barriers exist, with most datasets in English rather than Arabic.

1 Dr Alessandra Esposito, King's College London, Strand (London WC2R 2LS.) Email: alessandra.g.esposito@kcl.ac.uk; ORCID: 0000-0001-8959-4519.

2 Dr Pascal Flohr, Affiliate Research Fellow, University of Oxford. http://orcid.org/0000-0003-3203-913X

3 Shatha Mubaideen, Senior Research and Partnerships Officer, Council for British Research in the Levant. ORCID: 0000-0002-0749-3851.

4 Prof James Smithies, Australian National University. ORCID: 0000-0003-4801-0366.

5 Prof Fadi Bala'awi, Director General of the Department of Antiquities.

6 Dr Carol Palmer, Former Director, Council for British Research in the Levant. ORCID: 0000-0003-4612-5426.

7 Dr Sahar Idwan, Professor of Computer Science, Hashemite University.

Findings emphasize the need for improved data standardisation, interoperability, and expanded public access to digital heritage. The MaDiH catalogue serves as a foundational tool for researchers and is a good starting point for a Jordanian heritage catalogue, while providing an example for similar efforts in other regions, however it requires further development to enhance engagement and integration with World Heritage platforms. Future efforts should focus on federating existing datasets, fostering multilingual accessibility, and strengthening Jordan's digital infrastructure to maximize the academic and economic benefits of its rich cultural heritage.

Keywords: Jordan; Digital Cultural Heritage; Digital Archaeology; Roman Archaeology; Nabataean Archaeology

1 Introduction

This paper[8] discusses trends that emerged from the analysis of datasets related to Nabataean and Roman heritage from Jordan, the result of a wider data collection carried out by the MaDiH (حيدم): Mapping Digital Cultural Heritage in Jordan project[9]. The project aimed to scope the cultural heritage dataset landscape of Jordan across all periods to inform research infrastructure policy (Smithies et al. 2021a) and to suggest ways research software engineering (RSE) and technical applications could be deployed to maximise the academic and economic usage of that infrastructure (Smithies et al. 2021b)[10].

Jordan is at the forefront of the MENA (Middle East and North Africa) region in terms of digital cultural heritage (DCH) (Smithies 2021a, 9–14). Both institutional and societal interest in the development of DCH assets (such as digital repositories and archives, websites, and associated infrastructure) are high, which has resulted in high-profile collaborations between the Department of Antiquity of Jordan (DoA) and international heritage organisations, leading to a regionally significant DCH infrastructural footprint.

8 I would like to thank Anja Hellmuth Kramberger and Kaja Stemberger Flegar for their invitation to present at the Erasmus Summer school 2024 – Heritage and Beyond and consequently to contribute to this volume. I extend my thanks to the anonymous reviewers and their valuable suggestions. Any mistakes remain my own.
9 MaDiH (حيدم) was a joint project funded by the Arts and Humanities Research Council and the Newton Khalidi fund. It involved the King's Digital Lab (KDL), the Hashemite University, the Council for British Research in the Levant (CBRL), the Department of Antiquities of Jordan, the Jordanian Open Source Association (JOSA), and the Endangered Archaeology in the Middle East and North Africa (EAMENA) project (ref. AH/S011722/1). It ran from January 2019 to April 2021 and is dedicated to the memory of Dr Andrea De Silva Zerbini (1984–2019), one of its driving forces.
10 Other topical research papers and reports focused on international project management (Bala'awi et al. 2021; Smithies et al. 2023) and on the reception and usability of the project catalogue (Mubaideen et al. 2022).

Most notably, these collaborations resulted in the creation of the Jordanian Archaeological Database and Information System (JADIS) in 1991, in partnership with the American Centre for Oriental Research (Palumbo 1994). It was later replaced by MEGA-Jordan (The Getty Conservation Institute and World Monuments Fund 2010), the current system of indexing Jordanian archaeological sites (Myers and Dalgity 2012; The Getty Conservation Institute and World Monuments Fund 2014). More recently, the 'Documentation of the Objects in Jordanian Archaeological Museums (DOJAM)' project, promoted by the German Protestant Institute of Archaeology (GPIA), has led to the creation of a prototype national database of archaeological holdings to improve museum-management and enable scientific research (GPIA 2025).

In addition, other datasets that do not focus exclusively on Jordan are widely used, like the Aerial Photographic Archive for Archaeology in the Middle East (APAAME 2009), or the EAMENA project database (EAMENA 2023). The importance of these digital resources for conducting archaeological research in Jordan is confirmed by national and international researchers, however the number of these resources actually used was observed to be limited (Drzewiecki and Arinat 2017). They are often used in the initial phase of archaeological assessment for the preliminary surveying of the areas to be excavated (Zerbini and Banks 2015; Drzewiecki and Arinat 2017). In particular, MEGA-Jordan often informs heritage management in urban and extra-urban environments (Haroun 2016; Drzewiecki and Arinat 2017).

Set in this landscape, the MaDiH (حيدم) project intended to build a proof-of-concept digital catalogue to index all the above and other existing digital assets. The goal was to increase their use thus facilitating national and international engagement with Jordan's cultural heritage, either from an academic research perspective or to inform the development of tools enhancing public and touristic involvement with that heritage. Crucial to the spirit of the project was to help connect existing digital resources by standardizing the dataset metadata, while potential uses of the catalogue include the development of a shared space for the virtual federation of Jordanian heritage assets held around the world.

This paper illustrates the results of a test of the MaDiH (حيدم) catalogue by assessing whether it provides a useful go-to resource for researchers approaching specific aspects of Jordanian cultural heritage. The case study

for this test is a subset of Nabatean and Roman period related datasets[11]. To do so, it will consider aspects of data ownership and their accessibility, to evaluate how findable and accessible datasets related to these periods are, reflecting in particular on three aspects as a means of exploring how these datasets benefit the heritage community at large: who creates these digital assets, how they are shared with the heritage community, and how easy it is to access the information they hold.

The catalogue contains 326 datasets related to all periods of Jordanian history, from the Palaeolithic to the Hashemite period (Flohr forthcoming), while datasets related to the Nabataean period, the Roman period, or both, correspond to a total of 133 datasets. After briefly presenting the general project methodology used for the data collection (2), the paper will offer a summary of the Nabataean and Roman periods datasets sample in relation to the geographical areas of Jordan they cover and the types of evidence they are concerned with (3). It will then focus on specific aspects associated with datasets ownerships and their degrees of accessibility (4). Discussion of the findings (5) will be followed by concluding remarks (6).

2 Notes on the methodology of dataset collection in the MaDiH (مديح) project

The MaDiH (مديح) catalogue[12] was developed using CKAN, an open-source data management system based on Python, which facilitates publication and sharing of metadata and data (CKAN 2025). No raw data was collected in the catalogue, but only information *about* the datasets. In its conception, it was meant to function as both catalogue and *prototype* repository, but it is not an actual repository. Its goal is to provide a high-level 'landscape mapping' of Jordan's DCH infrastructure as an aid to analysis, research, and policy development, rather than, at this stage, a fully federated portal to the datasets or the digital objects they contain.

11 Two reasons have informed this decision: one has to do with the personal research interest of the first author, and the second with the good number of datasets making up this sample.

12 From now on referred to only as 'the catalogue'. After the completion of the dataset collection phase, the catalogue, originally hosted at the King's Digital Lab (KDL) in the United Kingdom, has since been transferred to the Hashemite University in Jordan, where it is maintained.

The data collection template used for the catalogue was based on a previous template developed by KDL to document their legacy projects (Ciula 2020) and customised by the MaDiH (مديح) research team following the Metadata Schema for the Description of Research Data Repositories (Rücknagel et al. 2015) and the DataCite Metadata Schema Documentation for the Publication and Citation of Research Data (DataCite Metadata Working Group 2016; Esposito et al. 2020).

The general datasets collection benefitted from the following sources:

- Online repositories
- Online library catalogues
- Online museum catalogues
- On-site visits to museums/institutions in Jordan
- Wide-spectrum internet search
- Word of mouth, based on the team's network
- Publications, e.g. journal papers
- Contacts established during MaDiH (مديح) outreach events

Highest priority was given to the collection of digital online datasets, followed by digital datasets stored offline, while the analogue datasets initially had a low priority (Esposito et al. 2020). In order to create a representative dataset, that also includes less visible (i.e. online) and accessible (i.e. online, digital) datasets, a deliberate effort was made to seek out more analogue-only datasets, but a (conscious) bias likely still exists in favour of the more accessible datasets. Moreover, fewer on-site visits than anticipated could be made due to the COVID-19 pandemic, which also likely reduced the number of offline and analogue datasets in the catalogue.

To analyse aspects related to the themes of dataset ownership and accessibility, we relied on the following tags:

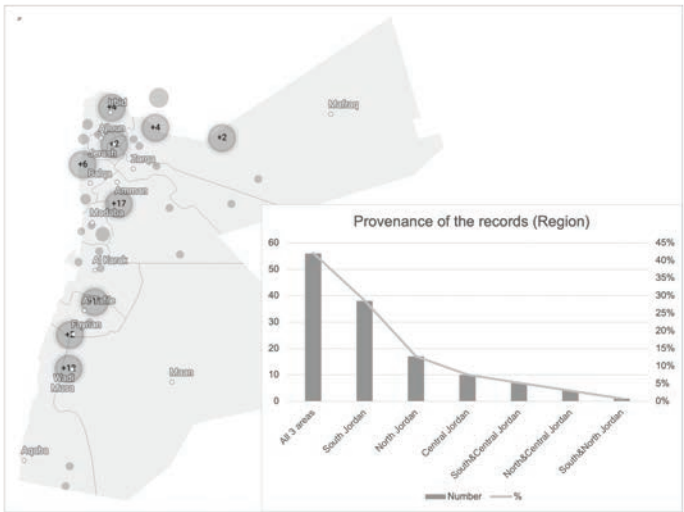
- Whether the dataset is available in digital, analogue form, or both.
- Whether the dataset is available online, offline, or both.
- Whether the dataset is publicly accessible, private, or partially accessible (tagged as 'partially public').
- Whether the dataset is held inside or outside Jordan and in which country.
- The dataset language(s).

3 The Nabataean and Roman periods datasets: An introduction

The Roman and Nabataean periods overlap chronologically, with 'Nabataean' covering the period from 100 BCE to CE 400 and 'Roman' referring to 50 BCE to CE 400 (MaDiH Team 2020a, 2020b). However, it is worth noting that Roman evidence is present in Jordan at least since the landing of Pompey in 63 BCE and that, although Jordan became part of the Roman province of Arabia Petrea in CE 106, this did not signal the end of Nabataean culture (Al-Otaibi 2011; Cimadomo 2018), as relevant material evidence continued well after this date (Amr 2004). The issue of periodisation in digital archaeology is a rich one, with many intellectual and technical issues, but is outside the scope of this paper (see for example Rabinowitz et al. 2016, 2025). The approach chosen by MaDiH (حيدم) is aligned to accepted scholarly practice employed in MEGA-Jo. What does this mean for our datasets collection? Depending on where an archaeological site is located, or the characteristics of a group of materials included in a certain dataset, that dataset might have been considered as pertaining to the Nabataean or Roman period, or both.

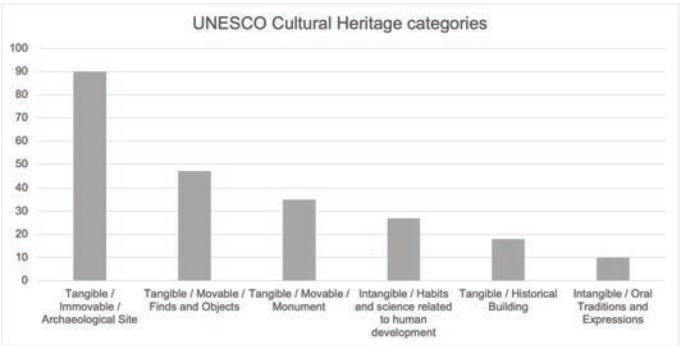
The Nabataean and Roman periods datasets mostly cover all regions of Jordan in the catalogue (42 %) (Figure 1). This group includes large national datasets, such as the Department of Antiquities Online Publications Archive (DOA 2025), or international ones (ICOMOS 2025; Ifpo 2025). In terms of individual regions, perhaps unsurprisingly, the most represented region is the South of Jordan, where Petra is located, (29 %), followed by the North (13 %) and Central (8 %) regions (Figure 1, inset).

Figure 1: Distribution of the records provenance by site (main) and percentage of the provenance of the records by region (inset).



The majority of the sites belong, again unsurprisingly given the periods, to the UNESCO Heritage categories[13] of Tangible Heritage, and more than a third (47) to the Movable Heritage category of 'Finds and Objects', showing a preponderance of datasets related to sites and monuments over finds, studies and catalogues (Figure 2).

Figure 2: Number of the Nabataean and Roman datasets by UNESCO Heritage category.



13 UNESCO 1992–2005, 2003.

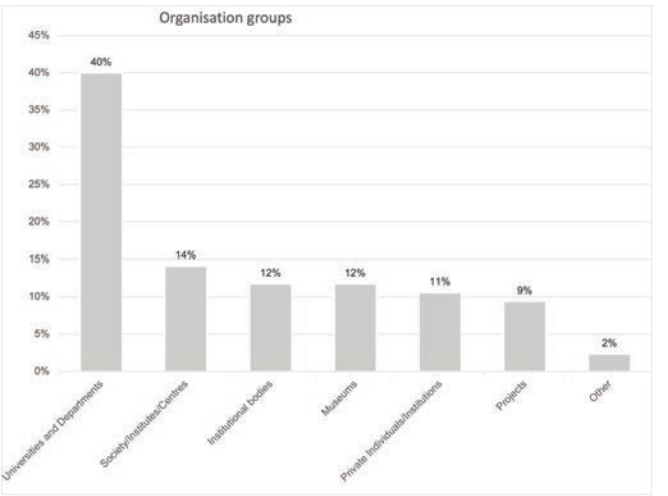
4 Finds from the Catalogue: Datasets ownership and accessibility

To analyse how the Nabataean and Roman periods datasets are created and shared with the heritage community, and how easy it is to access the information they hold, we will focus on two sets of characteristics recorded in the catalogue: one related to ownership, and one to accessibility. The first covers questions on who creates and owns the datasets and the formats in which they are available, while the second considers whether these datasets are publicly accessible or are kept private and to what extent.

4.1 Dataset ownership

Data ownership was tracked by analysing the types of organisations that produce datasets and the formats in which they are kept, reflecting their commitment to sharing their findings and/or collections and broader levels of accessibility. The majority of datasets are held by universities (40 %), while societies/institutions/centres host 14 % of them (Figure 3). Institutional bodies are responsible for 12 % of them, however more than half of the datasets in this category of organisations (15, corresponding to the 56 % of the datasets in this category) are 'owned' by the DOA, while 'Museums' hold *only* 12 % of the datasets related to the periods considered here. These numbers depend on the fact that many Jordanian museums are 'owned' by the DOA 'organisation', and thus do not appear in the catalogue as individual museums.

Figure 3: Types of organisations holding Nabataean/Roman periods datasets.



Private institutions or individuals own 11 % of the datasets, which heavily influences the modalities of their accessibility, a point that will be addressed in section 4.2. Furthermore, 9 % of the datasets are held by 'project' organisations, which refers to collaborative arrangements across different universities or institutions (e.g., the EAMENA project).

81 % of the datasets from the Nabataean-Roman period are tagged as 'digital', 14 % are double tagged as 'digital' and 'analogue', and 5 %, corresponding to 6 datasets, are tagged exclusively as 'analogue'. Before focussing on the digital datasets, it is worth describing the datasets that are tagged as 'analogue'. Five out of the six analogue datasets are research projects. It is likely that these datasets are available in some digital form, completely or at least in part, however this information was not available from the public sources consulted to record these databases. Only one dataset can be considered 'purely' analogue, i.e. the paper catalogue of the Karak Archaeological Museum (Jarajreh pers. comm.).

The datasets from the Nabataean-Roman period double tagged as digital and analogue (14 %) are primarily held in Jordanian museums. These collections are often recorded both in analogue (e.g. paper records) and digital files. Some archaeological projects, such as the iterations of the Madaba Plains Project (2025) or the Building Archaeology in Jordan Project (Parenti 2012), also fall into this hybrid category, as they compile and analyse both analogue and digital data from old and current archaeological excavation seasons. This 'digital/analogue' double tagging is useful for finding current digitisation projects, where only parts of the records have currently been digitised, (e.g. The French Institute of the Near East Library Catalog (Ifpo 2025)).

There are 108 exclusively 'digital' datasets, covering the Nabataean and Roman periods in part or exclusively. We have already addressed how this value is likely the result of the dataset collection methodology employed by the MaDiH (مديح) research team, which prioritised datasets available online over digital-offline or purely analogue ones, as these were both more accessible to the team and increased the speed of collection, thus supporting the project's scoping aim[14].

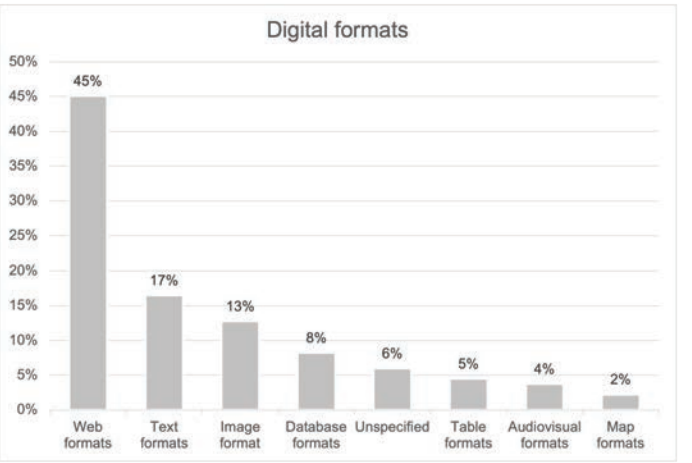
14 See section 2.

The 'digital' tag umbrella covers a variety of digital formats:

- Audiovisual (MP4, WMA)
- Database (ACCDB, ADLIB, ARCHES, ARK, FMP, JSON, SQL)
- Image (JPEG, TIFF, U3D)
- Map (QGIS, FMI, KML)
- Tabular (CSV, XLS)
- Text (DOC, PDF, TXT)
- Web (HTML, PHP, XML)

Web-based formats (45 %) are the most common, as they are often used to present archaeological projects, but this category also includes online catalogues and collections produced by museums and libraries (Figure 4). 17 % of datasets are in text formats (typically PDF files) and correspond to datasets published as journal articles, books, and reports. These text files often contain data in embedded tables or map images used to summarise the findings. Database (8 %) and tabular (5 %) formats are much less represented than expected, considering that many research projects, especially archaeological ones, almost always produce lists and/or tables of raw data.

Figure 4: Percentage of Nabataean and Roman periods datasets digital formats.



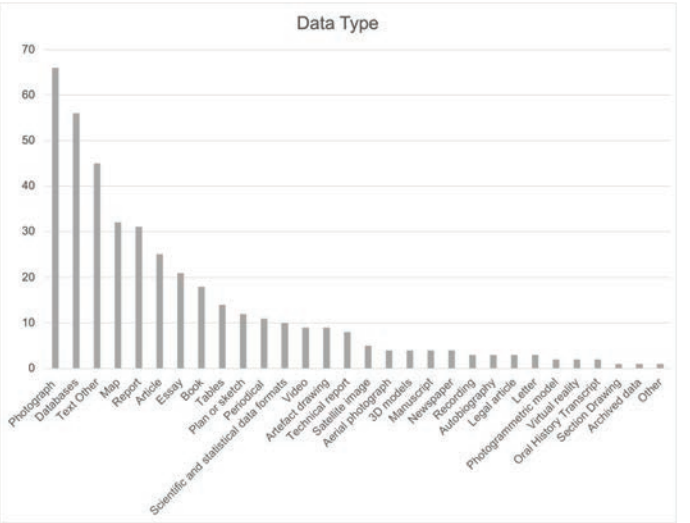
Audio-visual formats, although relatively uncommon (4 %), highlight the existence of datasets of video or audio recordings related in part to the preservation of intangible cultural heritage and in part to documenting the process of archaeological research of tangible heritage. Notable in this

category were videos of excavations at archaeological sites, often made via drones. Examples of audio-visual datasets include audio recordings of interviews from the Udhruh Oral History Project (Hageraats 2013), or videos and recordings of the excavations at the Petra area by the Association for the Understanding of Ancient Cultures (The Association for the Understanding of Ancient Cultures 2025).

Most datasets are image-related (specifically photographs), which are used widely both for datasets recording sites and/or monuments and finds catalogues. There is also a strong presence of archaeology-focused datasets, i.e. reports (31 datasets), plans or sketches (12), artefact drawings (9), technical reports (8), and section drawings (1). Interpretative visual formats, such as 3D models, photogrammetric models, and virtual reality formats only account for a small number of the datasets recorded (respectively 4, 2, and 2).

Purely map-based resources were the least common format (2 %), corresponding to just three datasets. The limited number of map formats might seem unexpected given the time periods considered here, as maps are crucial for visualizing the locations of archaeological sites and finding spots for objects. However, the low number only shows the limited availability of datasets published using sophisticated GIS-based technology. As mentioned earlier, maps are often included as text or image formats (if, for example, a map collection has been digitised as JPEG files). Looking at findings in the 'Data Type' field, we see that the Map category counts for 32 datasets (Figure 5). Similarly, another 'Data Type' field allows us to appreciate the number of 'Scientific and statistical data formats' (56 datasets), compared to the 'Database formats', which accounted for only 11 datasets. This is because many databases, such as museum collection catalogues, have been recorded as web formats rather than database formats (i.e. they are published in a web format, but there is an underlying database (Smithies et al. 2021b)). This 'hidden' information is an important aspect of the catalogue; essential information was gained but the project could only provide a tantalising snapshot of the extensive data relating to Jordanian archaeological heritage.

Figure 5: Numbers of the Nabataean and Roman periods datasets by 'Data Type'.



4.2 Dataset accessibility

Assessing how accessible these datasets are to different audiences was one of the main goals of the data analysis. 'Accessibility' is one of the four principles of FAIR[15] data (Wilkinson et al. 2016), but this does not necessary imply that the data collected is to be available in open access, only that the modalities of its accessibility are clearly stated. We considered physical accessibility for both Jordanian and international audiences in the catalogue, indicating whether a dataset is available online or offline, and whether it is hosted outside or inside Jordan. The language in which the dataset was produced was also recorded to highlight the presence of potential language barriers.

74 % of the datasets are accessible online. This percentage includes databases and institutional collections that allow access to their actual data, often through an online search tool, as well as archaeological and heritage projects, which usually have a website to showcase their activities, share bibliographic lists, and include contact details of researchers involved in the projects. The latter, however, do not provide access to their data. These websites have been included in the catalogue to ensure that the us-

15 Findability, accessibility, interoperability, and reusability.

ers would have information on how to access those datasets by contacting their owners, thus promoting research networking and boosting data accessibility. In this sense, the findings for the Nabataean and Roman subset show that although most datasets recorded are available online, the actual amount of data (i.e. the information contained in the datasets) that is actually accessible online is limited.

The smaller yet significant percentage of offline datasets (26 %, equalling 34 datasets) further confirms the challenges related to data accessibility. The reasons for this vary from restrictive licenses to concerns regarding data disclosure, lack of funding, and/or ability to provide online access to sometimes large datasets.

Knowing the 'privacy' status of a dataset is important for planning any engagement activity. Just over half of the datasets recorded in our subset were tagged as 'public' (53 %, 71 datasets). A 'public' dataset is one whose data is readily available to be consulted, either online or by visiting its holding institution. Surprisingly, given the importance of publicly funded research, 'private' datasets account for 27 % of the total of our subset, while 20 % are 'partially public', highlighting the hybrid accessibility of those datasets that can be consulted and used upon receiving authorisation from the owners (for example, the EAMENA database, which requires user registration).

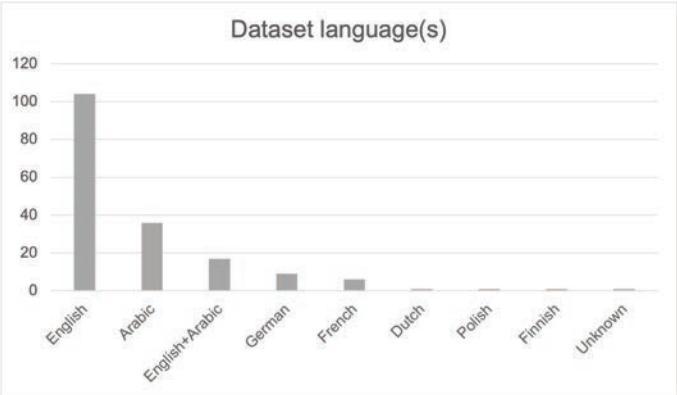
The geographical location of a dataset also influences its accessibility, especially if it happens not to be digital and publicly available. Most of datasets are located outside of Jordan (98 datasets, 74 % of the total). Of these, the most represented country is by far the USA, followed at a distance by the United Kingdom (Table 1). Other European countries are represented with much smaller numbers, although Canada and South Africa also appear. The only country from the MENA region is Lebanon. Of these datasets, the majority are digital. This highlights the efforts, interests and capacities of these countries to digitize. In fact, these might be the criteria by which we can best judge these datasets, rather than which countries actually hold more Jordanian heritage datasets. A second aspect worth considering is what level of public access is possible for these datasets. While Jordan itself has a low number of public datasets compared to the total, all the other countries have at least half their datasets publicly available.

Table 1: Roman and Nabataean datasets numbers by location and according to the 'digital', 'online', and 'public' tags.

	Dataset locations			
	Total	**Digital**	**Online**	**Public**
United States of America	42	39	32	23
Jordan	35	33	15	9
United Kingdom	19	19	18	13
Germany	10	8	8	6
France	8	8	8	7
Netherlands	6	6	6	3
Canada	2	2	2	2
Denmark	2	2	2	2
Italy	2	2	2	1
Poland	2	2	2	2
South Africa	2	2	1	2
Finland	1	1	1	1
Lebanon	1	1	1	1
Switzerland	1	1	1	1
Unknown	2	2	2	1

Even if a dataset is publicly accessible and online, a further barrier to its usability may be the language used to record the data (Figure 6). English is the most commonly used language, while Arabic is a distant second. Only a small number of datasets (17) are available both in English and Arabic, mostly Jordanian institutional datasets (7). Concerns over the accessibility of English-produced content for Arabic speakers are demonstrated by multilingual databases such as the EAMENA project database, which has an Arabic version. However, the number of datasets in other languages poses a further obstacle to the accessibility of Jordanian data.

Figure 6: Numbers of Nabataean and Roman datasets by language.



5 The relevance of the MaDiH (مَديح) catalogue for the study of the Nabataean and Roman periods in Jordan

The MaDiH (مَديح) catalogue was intended from its conception the intention of being a scoping exercise in understanding the landscape of digital datasets related to the Jordanian cultural heritage of any period. By looking at a subset period, the goal was mainly to test whether such catalogue could also be considered a useful resource for researchers interested in a specific period of Jordanian heritage, for example the Nabataean and Roman periods.

The total number of datasets in our subset is 133, which certainly does not represent all relevant datasets. It is also biased towards those datasets that were digital and available online at the time of data collection. Nevertheless, in its current form and for the periods concerned, the catalogue *is* a useful resource. For students and researchers approaching the study of Jordanian heritage, it showcases the necessary datasets to start off research on several aspects of Nabataean and/or Roman period heritage, covering archaeological sites and finds. Crucially, the data collection process involved meetings and encounters with professionals from Jordanian institutional and museum organisations, who provided precious insights about their institutions' holdings that are not available online or might be exclusively in Arabic.

The datasets currently in the catalogue might be already familiar to more expert researchers working on the Nabataean and Roman periods.

However, the value of this catalogue in this instance might not be in knowing that a dataset exists, but whether that dataset is accessible physically (where the dataset is stored), legally (who owns the dataset), in which format its content was recorded, and whether/to what degreecan be reused

For future dataset creators, the catalogue offers an overview of the work that has already been done. This limits the perils of creating digital 'islands', i.e. research efforts that are isolated from each other, which can lead to duplication of digital projects and consequent waste of funding. It is, in fact, becoming challenging for researchers to keep track of an increasingly complex landscape of digital outputs produced internationally and in different languages (Bala'awi et al. 2021). Growing concerns on this topic have funded interest in providing infrastructure to digital archaeology, as shown by the creation of large repositories of archaeological data like The Digital Archaeological Record (tDAR 2025) or the Archaeological Data Service (ADS 2025). The multiplication of digital repositories created by museums and institutions to digitise and share their collections has created a variety of outcomes both in terms of quality and sustainability. For this reason, initiatives such as the Archaeological Archives Forum (AAF 2017) and the Digital Archives in Archaeology project (DigVentures 2025) are aiming to provide guidance on how to create and manage digital cultural heritage data. Certainly, the creation of collections of such repositories in containers like re3data (2025) or Europeana (2025), which aim to connect repositories of archaeological and heritage data, provides important short-term resources that have not escaped the interest of tech giants such as Google (2025). But ideally, long-term solution would create a federated network of archaeological repositories underpinned by long-term funding and automatically aggregated by national heritage portals. On a more focused subset than the one considered here, the creation of similar repositories for other provinces of the Roman Empire could be facilitated by the adoption of similar ontologies that would enable the employment of linked data. In this respect, the Ariadne Plus Portal (ARIADNE Plus 2025), which aims to implement interoperability, linking resources, and enabling data mining (Geser 2016; Wright 2016) provide a successful example in this direction.

6 Conclusions

The data in the MaDiH (حيدم) catalogue highlight how rich the landscape of digital resources is for the study of Jordanian heritage related to the Nabataean and Roman periods, while at the same time also emphasising how these assets are currently underused. Rather than envisioning a vast new archive for these resources, the analyses conducted during the project and in this paper suggest how it would be possible to strengthen the already existing digital resources to build an integrated system of data preservation and research. This approach can similarly inform research on other periods of Jordanian history, as well as other areas of the Roman world: in many ways, it provides methodological and conceptual direction for the aggregation and analysis of archaeological data, as much as rigorous empirical evidence.

Bibliography

AAF. 2017. *Archaeology Archive Forum*. Access: https://archives.archaeologyuk.org/#proj (24 January 2025).

ADS. 2025. *Archaeology Data Service - ADS*. Access: https://archaeologydataservice.ac.uk/ (24 January 2025).

Al-Otaibi, Fahad Mutlaq. 2011. *From Nabataea to Roman Arabia: Acquisition or Conquest*. Oxford: Archaeopress.

ʼAmr, Khairieh. 2004. Beyond the Roman annexation: the continuity of the Nabataean pottery tradition. *Studies in the History and Archaeology of Jordan* (7): 237–245.

APAAME. 2009. *Aerial Photographic Archive for Archaeology in the Middle East - APAAME*. Access: https://www.flickr.com/photos/apaame/collections (24 January 2025).

ARIADNE Plus. 2025 *Portal page*. Access: https://ariadne-infrastructure.eu/portal/ (24 January 2025).

Bala'awi, Fadi, Shatha Mubaideen, James Smithies, Pascal Flohr, Alessandra Esposito, Carol Palmer, and Sahar Idwan. 2021. The MaDiH: Mapping Digital Cultural Heritage in Jordan, Opportunities and Limitations. *Culture and Computing: Interactive Cultural Heritage and Arts: 9th International Conference C&C 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part I*, ed. Matthias Rauterberg, 15–26. Cham: Springer

Cimadomo, Paolo. 2018. The controversial annexation of the Nabataean Kingdom. *Levant* 50: 258–266. Access: https://doi.org/10.1080/00758914.2019.1614769 (24 January 2025).

Ciula, Arianna. 2020. *Exposing legacy project datasets in Digital Humanities*. Access: https://kdl.kcl.ac.uk/blog/legacy-project-datasets/ (24 January 2025).

CKAN. 2025. *CKAN - Homepage*. Access: https://ckan.org/ (24 January 2025).

DataCite Metadata Working Group. 2016. *DataCite Metadata Schema: Documentation for the Publication and Citation of Research Data: Version 4.0*. Access: http://doi.org/10.5438/0012 (24 January 2025).

DOA. 2025. *Department of Antiquites of Jordan: Publication Archive*. Access: http://publication.doa.gov.jo/ (24 January 2025).

DigVentures. 2025. *Digital Archives in Archaeology*. Access: https://digventures.com/projects/digital-archives/ (24 January 2025).

Drzewiecki, Mariusz, and Mahmoud Arinat. 2017. The impact of online archaeological databases on research and heritage protection in Jordan. *Levant* 49(1): 64–77. Access: https://doi.org/10.1080/00758914.2017.1308117 (24 January 2025).

EAMENA. 2023. *EAMENA – Endangered Archaeology in the Middle East and North Africa*. Access: https://eamena.org/ (24 January 2025).

Esposito, Alessandra, Shatha Mubaideen, Pascal Flohr, James Smithies, and Fadi Bala'awi. 2020. *MaDiH (حيدم) Mapping the Digital Cultural Heritage in Jordan Project: Datasets Identification and Publication Protocol*. Access: https://doi.org/10.5281/zenodo.4146756 (24 January 2025).

Europeana. 2025. *Homepage*. Access: https://www.europeana.eu/en (24 January 2025).

Flohr, Pascal. Forthcoming. Mapping Digital Heritage in Jordan (MaDiH): What are the data telling us?

Geser, Guntram. 2016. *ARIADNE WP15 Study: Towards a Web of Archaeological Linked Open Data*. Access: http://legacy.ariadne-infrastructure.eu/wp-content/uploads/2019/01/ARIADNE_archaeological_LOD_study_10-2016-1.pdf (24 January 2025).

Google. 2025. *Dataserch – About page*. Access: https://datasetsearch.research.google.com/help (24 January 2025).

GPIA. 2025. *DOJAM Project - Homepage*. Access: https://www.zitadelle-amman.de/ (24 January 2025).

Hageraats, C. 2013. *Thematic collection: Udhruh Oral History Project: DANS*. Access: https://doi.org/https://doi.org/10.17026/dans-275-pdb6 (24 January 2025).

Haroun, Jihad. 2016. *MEGA-Jordan as a Preventive Resource*. Access: https://podcasts.ox.ac.uk/mega-jordan-preventive-resource (24 January 2025).

ICOMOS. 2025. *Open Archive: EPrints on Cultural Heritage*. Access: https://openarchive.icomos.org/ (24 January 2025).

Ifpo. 2025. *Catalogue de la médiathèque de l'Institut français du Proche-Orient*. Access: https://mediatheque.ifporient.org/cgi-bin/koha/opac-main.pl (24 January 2025).

Madaba Plains Project. 2025. *The Madaba Plains Project Centered at Andrews University and LaSierra University*. Access: https://www.madabaplains.org/ (24 January 2025).

MaDiH Team. 2020a. *MaDiH Vocabulary (English and Arabic)*. Access: https://doi.org/10.6084/m9.figshare.12950879.v1 (24 January 2025).

MaDiH Team. 2020b. *The MaDiH data collection form: Zenodo*. Access: https://doi.org/10.5281/zenodo.4289918 (24 January 2025).

Mubaideen, Shatha, Pascal Flohr, James Smithies, Fadi Bala'awi, Carol Palmer, Sahar Idwan, and Alessandra Esposito. 2022. Assessing the MaDiH CKAN Catalogue as an Engagement Tool for the Jordanian Cultural Heritage Community. *Levant* 54(2): 277–284.

Myers, David, and Alison Dalgity. 2012. The Middle Easter Geodatabase for Antiquities (MEGA): An Open Source GIS–Based Heritage Site Inventory and Management System. *Change Over Time* 2(1): 32–57.

Palumbo, Gaetano. 1994. *JADIS: The Jordan antiquities database and information system: A summary of the data*. Amman: Department of Antiquities of Jordan: American Center of Oriental Research.

Parenti, Roberto. 2012. Building Archaeology in Jordan: Preliminary Report on the 2009 – 2011 Surveys at Umm As-Surab. *Annual of the Department of Antiquities of Jordan* (56): 187–195. Access: http://publication.doa.gov.jo/uploads/publications/44/ADAJ_2012_56-187-196.pdf (24 January 2025).

Rücknagel, Jessika, Paul Vierkant, Robert Ulrich, Gabriele Kloska, Edeltraud Schnepf, David Fichtmüller, Evelyn Reuter, Angelika Semrau, Maxi Kindling, Heinz Pampel, Michael Witt, Florian Fritze, Stephanie van de Sandt, Jens Klump, Hans-Jürgen Goebelbecker, Michael Skarupianski, Roland Bertelmann, Peter Schirmbacher, Frank Scholze, Claudia Kramer, Claudio Fuchs, Shaked Spier, and Agnes Kirchhoff. 2015. *Metadata Schema for the Description of Research Data Repositories: Version 3.0.* Access: https://doi.org/10.2312/re3.008 (24 January 2025).

Rabinowitz, Adam, Ryan Shaw, and Patrick Golden. 2025. *A gazetteer of periods for linking and visualizing data*. Access: https://perio.do/ (24 January 2025).

Rabinowitz, Adam, Ryan Shaw, Sarah Buchanan, Patrick Golden, and Eric Kansa. 2016. Making sense of the ways we make sense of the past. *History and Landscape* 59(2): 42–55.

re3data. 2025. *Registry of Research Data Repositories*. Access: https://www.re3data.org/ (24 January 2025).

Smithies, James, Fadi Bala'awi, Pascal Flohr, Shatha Mubaideen, Alessandra Esposito, Sahar Idwan, Carol Palmer, and Issa Mahasneh. 2021a. *MaDiH (حيدم): Mapping Digital CulturalHeritage in Jordan: Policy White Paper*. Access: https://doi.org/10.5281/zenodo.4867530 (24 January 2025).

Smithies, James, Fadi Bala'awi, Pascal Flohr, Shatha Mubaideen, Alessandra Esposito, Sahar Idwan, Carol Palmer, and Issa Mahasneh. 2021b. *MaDiH (حيدم): Mapping Digital Cultural-Heritage in Jordan: Technical White paper*. Access: https://doi.org/10.5281/zenodo.4867530 (24 January 2025).

Smithies, James, Pascal Flohr, Fadi Bala'awi, Sahar Idwan, Carol Palmer, Alessandra Esposito, Shatha Mubaideen, and Shaher Rababeh. 2023. MaDiH (حيدم): A Transnational Approach to Building Digital Cultural Heritage Capacity. *Journal on Computing and Cultural Heritage* 15(4): 1–14.

tDAR. 2025. *The Digital Archaeological Record – tDAR*. Access: https://www.tdar.org/ (24 January 2025).

The Association for the Understanding of Ancient Cultures. 2025. *Project Website*. Access: https://www.auac.ch/ (24 January 2025).

The Getty Conservation Institute and World Monuments Fund. 2010. *MEGA Jordan*. Access: http://www.megajordan.org/ (24 January 2025).

The Getty Conservation Institute and World Monuments Fund. 2014. *Middle Eastern Geodatabase for Antiquities, Jordan*. Access: https://www.getty.edu/projects/middle-eastern-geodatabase-antiquities-jordan/ (24 January 2025).

UNESCO. 1992–2005. *Convention Concerning the Protection of the World Cultural and Natural Heritage*. Access: https://whc.unesco.org/en/conventiontext/ (24 January 2025).

UNESCO. 2003. *Convention for the Safeguarding of the Intangible Cultural Heritage*. Access: https://ich.unesco.org/en/convention (24 January 2025).

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E. Bourne, Jildau Bouwman, Anthony J. Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T. Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J.G. Gray, Paul Groth, Carole Goble, Jeffrey S. Grethe, Jaap Heringa, Peter A.C 't Hoen, Rob Hooft, Tobias Kuhn, Ruben Kok, Joost Kok, Scott J. Lusher, Maryann E. Martone, Albert Mons, Abel L. Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene van Schaik, Susanna-Assunta Sansone, Erik Schultes, Thierry Sengstag, Ted Slater, George Strawn, Morris A. Swertz, Mark Thompson, Johan van der Lei, Erik van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao, and Barend Mons. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3(160018). Access: https://doi.org/10.1038/sdata.2016.18 (24 January 2025).

Wright, Holly. 2016. *Linked Open Data Approaches within the ARIADNE project*. Access: https://archaeologydataservice.ac.uk/app/uploads/2022/05/WrightEAA2016.pdf (24 January 2025).

Zerbini, Andrea, and Rebecca Banks. 2015. *The Madaba Ring Road, Jordan: Evidence of Cultural Heritage Assets from Remote Imagery in the Madaba Hinterland*. Access: https://eamena.org/article/madaba-ring-road-jordan-evidence-cultural-heritage-assets-remote-imagery-madaba-hinterland (24 January 2025).