Understanding Data Reuse and Barriers to Reuse of Archaeological Data. A quality-in-use methodological approach

Kristy-Lee Seaton, Rimvydas Laužikas, Peter McKeague, Vera Moitinho de Almeida, Keith May and Holly Wright

Over the last decade, innovation has centred on making archaeological data more interoperable, increasing the discoverability of data through integrated cross-search and facilitating knowledge creation by combining data in new ways. An emerging research challenge for the next decade is optimising archaeological data for reuse and defining what constitutes good practice around reuse. Critical to this research is understanding the current state-of-the-art regarding both existing best practices and barriers to using and reusing archaeological data. This research aimed to understand how to optimise archives and interfaces to maximise the discovery, use and reuse of archaeological data and explore how archaeological archives can better respond to user needs.

The study was bound by (i) the reuse of digital archaeological archives; (ii) orientation to content usability and reusability; (iii) maintaining a user-orientated approach; (iv) collecting data from professionals in archaeology and heritage. The research group members adopted the quality-in-use conceptual approach for this study. Quality in use is 'the degree to which a product or system can be used by specific users to meet their needs to achieve specific goals with effectiveness, efficiency, satisfaction, and freedom from risk in specific contexts of use'. The research methodology is based on the SQuaRE (System and Software Quality Requirements and Evaluation) model, represented in the ISO/IEC 25000 standards series. In addition, the quality-in-use metric for investigation of reuse and barriers to reuse of archaeological data were adopted from the standardised measurement functions and methods of ISO/IEC 25022:2016. The result was a methodological model composed of 5 characteristics (Effectiveness, Efficiency, Satisfaction, Context coverage and Usability) with 14 measures (Task completeness, Objectives achievement, Task time, Cost-effectiveness, Overall satisfaction, Satisfaction with features, User trust in the system, data and paradata, User pleasure, Physical comfort, Context completeness, Flexible context of use and User guidance completeness). The methodology was tested with specific Contexts of use (use

cases), orientated to a distinct user with the specific professional goal of data reuse. Three use cases relating to 3D Pottery, radiocarbon, and GIS data were created. The pilot study has proven that the methodology works and could be applied in future research. This article discusses the application of the quality-in-use approach for evaluating the quality of digital archaeological archives, as well as presenting the methodology and the results of the pilot study.

1. Introduction

Over the past decade, innovation has helped address the interoperability of archaeological data, increasing the discoverability of data through integrated cross-searches and facilitating knowledge creation by combining data in new ways. An emerging research challenge for the next decade is understanding what constitutes good practice so that archaeological data can be optimised for reuse. Central to this research is understanding the current state-of-the-art, encompassing both existing best practices and barriers to using and reusing digital archaeological data. This topic is investigated within the framework of 'Saving European Archaeology from the Digital Dark Ages' (SEADDA), a European Cooperation in Science and Technology (COST) Action funded by the Horizon 2020 Framework Programme. Members of SEADDA represent 40 countries, 35 of which are COST members, and five are international partner countries (Fig. 1).
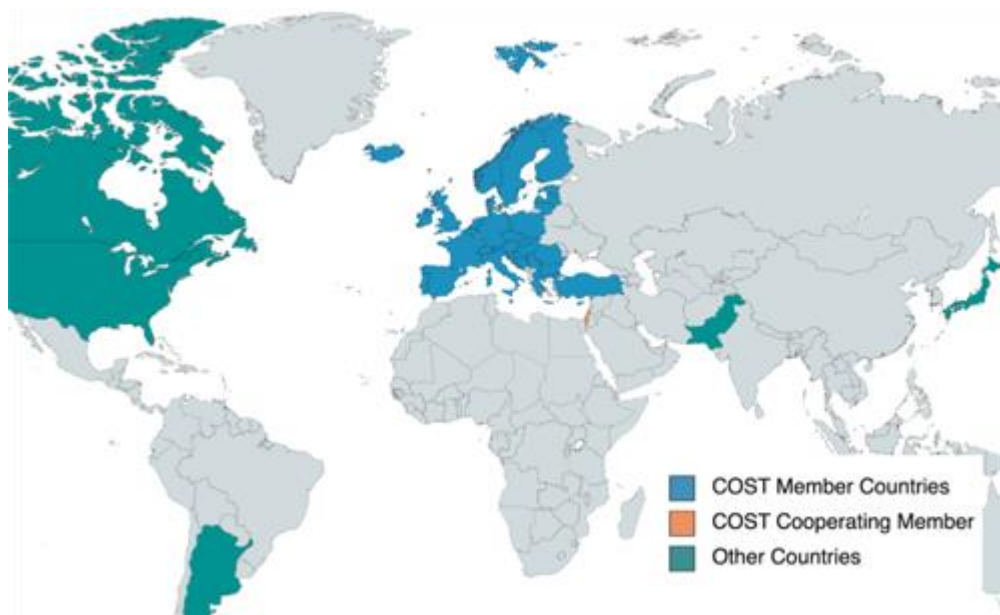

Figure 1: Map of countries participating in COST (European Cooperation in Science and Technology)

The SEADDA network comprises heritage specialists, archaeologists, academics, curators, and museum staff, who are all passionate about heritage data and have fostered a collaborative environment to progress this research. More specifically, the SEADDA research community is trying to understand the current state-of-the-art regarding the preservation, dissemination and reuse of archaeological data. Capacity

building has been addressed through holding data management and stewardship workshops in member countries. A special edition of *Internet Archaeology* explored the work of Working Group 1, presenting the current state-of-the-art regarding preservation for 24 member countries (Richards *et al.* 2021). This article relates more specifically to the research undertaken by Working Group 4 (WG4) that examined the use and reuse of archaeological data.

The ambitions of WG4 were to understand the qualitative and quantitative barriers to data reuse, crucial to understanding the impact of the broad adoption of FAIR principles in the research landscape (Richards et al. 2021). An exploratory workshop was held online between 31 March and 2 April 2020 to help define the current landscape pertaining to digital data use and reuse, providing a springboard to tackle further research (Wright 2020). WG4 developed the approach described in this article following the workshop. After considering a broad survey profiling the reuse sector, a targeted survey exploring quantitative and qualitative reuse was developed, and three research questions were defined:

- Research Question One: What are the ways to document and understand the quantitative and qualitative reuse of archaeological archives?
- Research Question Two: What are the barriers to the reuse of archaeological archives?
- Research Question Three: How can archaeological archives better respond to user needs?

The decision was made to limit their investigation to explore the underlying infrastructure of search interfaces used by data archives and repositories. Qualitative data were collected to evaluate the user experience. A conscious decision was made not to evaluate a particular archaeological archive but to understand the processes of discovery, use and reuse of archaeological data. This article outlines the processes of developing a methodology to collect and evaluate data on the user experience and the subsequent results. The survey had limited success owing to methodological issues. Notwithstanding, the results provided several interesting points of discussion. The methodology WG4 designed posed limitations on the research design that should be considered in future iterations of research, along with narrowing the scope of tasks for participants.

2. Methodology

2.1. Methodological considerations

The objective of WG4 was to understand who was using and reusing digital archaeological data and what challenges they experienced. There have been few studies that have explored the attitudes and opinions of a broad audience. An exception to this is Faniel *et al.* (2013, 297), who interviewed 22 archaeologists about their experiences reusing data. Several articles explore personal experiences reusing archaeological data (see Huggett 2018; Marwick and Pilaar Birch 2018; Kansa *et al.* 2014 and Sobotkova 2018 for examples). Geser *et al.* (2022) collected quantitative data on the policies and practices of digital archives, aimed more at an institutional level. The challenge for WG4 then was determining how qualitative data

should be collected. The research members approached this challenge by adopting the 'Quality-in-Use' conceptual approach. Quality-in-use is 'the degree to which a product or system can be used by specific users to meet their needs to achieve specific goals with effectiveness, efficiency, satisfaction, and freedom from risk in specific contexts of use' (International Organisation for Standardisation 2016, 41). The research methodology utilised two existing standards to create a method and a uniform set of metrics.

The research methodology developed was based on existing approaches. A user-based approach to measuring product quality recognises that the values placed on product characteristics will vary between users (Garvin 1984, 27). In response to this approach, two existing ISO standards were adopted. The SQuaRE (System and Software Quality Requirements and Evaluation) model, represented in the ISO/IEC 25000 standards series, is a tool for evaluating software quality (International Organisation for Standardisation 2017), while ISO/IEC 25022:2016 provides a set of standardised measurement functions that measure software performance from the user's perspective (International Organisation for Standardisation 2016). The methodological model created consisted of five characteristics (effectiveness, efficiency, satisfaction, context coverage and usability) with 14 measures (task completeness, objectives achieved, task time, cost-effectiveness, overall satisfaction, satisfaction with features, user trust in the system, the data and metadata, user pleasure, physical comfort, context completeness, flexible context of use and user guidance completeness). In addition, the study was bound by (i) the reuse of digital archaeological archives; (ii) orientation to content usability and reusability; (iii) maintaining a user-orientated approach; (iv) collecting data from professionals in archaeology and heritage. Testing the methodology required a context that described the specific user, environment and goals for reusing data. Three user types were nominated to maintain comparable results, and user scenarios were developed that reflected real-life examples for reusing data. The scenarios were developed from existing research questions that research group members had.

Geographic Information Systems (GIS)

Three GIS user scenarios were developed that tested the different needs of GIS users. The first scenario tested the users' ability to locate data by means of the functionality of archive catalogues to find and access GIS data, not whether a particular dataset has been archived or interoperability across data discovered. The second scenario tested functionality and the ability to download datasets within web map browsers. The third scenario was more specialised and explored the awareness and use of web services to access information. While the scenarios tested archive discovery and accessibility, they did not explore issues of interoperability or formats across the datasets discovered.

3D Pottery data

Two scenarios were designed to examine 3D Pottery data. Both scenarios required the user to locate high-resolution 3D models of pottery. These two scenarios were similar to the GIS scenarios and examined the users' ability to locate data and their

awareness of web-based visualisation services. The limitation of the scenarios was a general lack of openly available 3D Pottery data.

Radiocarbon data

As with the other two research areas, the single radiocarbon scenario tested the ability of users to locate data. While the scenario developed was specific to an existing research question, the research question was somewhat niche, and there was an awareness that this may impact the overall result.

Each scenario was broken down into objectives and tasks to achieve those objectives. These tasks provided a series of templates that could be used to test the methodology. The number of tasks varied between four to five for each scenario. A workflow and online survey form, designed to emulate the ISO25022 metrics, was developed using Qualtrics software to test the Quality-in-Use methodology. The research group decided that a survey provided the best opportunity for testing the methodology as it could be sent to other members of SEADDA to increase the testing rate. A series of discussions occurred discussing the survey format and the best way to test the Quality-in-Use methodology. Survey development was iterative, allowing all members to contribute to the final product. The scenarios required some redevelopment to ensure consistency. When changes to one scenario were suggested, the other scenarios also needed to be changed to ensure consistency.

2.2. Tool description

The Quality-in-Use survey was divided into three sections:

1. Demographic information about the respondent
2. A series of tasks for the respondent to complete emulating the previously designed use-case scenarios
3. Qualitative questions addressed user satisfaction with certain features, asking the respondent about flexibility, satisfaction and trust

The data collected in the second section were primarily quantitative, with space for respondents to add comments (see Appendix A [PDF] for a full list of survey questions). The questions in the third section related to the 14 measures in ISO/IEC 25022:2016. These were qualitative questions centred on user satisfaction, context and usability concerning the tasks completed in the second section of the survey. There was concern that respondents may abandon the survey if they had too many text boxes to fill out. Therefore, questions relating to user satisfaction and trust were set out in a matrix, with free-text boxes to provide further comments. The matrix used a Likert scale, as suggested by the ISO standard. However, cultural implications exist regarding the Likert scale, ranging from how scores are assigned to a general sense of discomfort with this type of measurement (Lee *et al.* 2002, 296). Nevertheless, Likert scales are common for surveys and easily recognisable by respondents, so alternative scoring methods were not explored.

3. Results

The survey was distributed through SEADDA channels, and personal invitations to participate were sent by WG4 members to colleagues or promoted during presentations. The survey was open from 30 March 2022 to 30 April 2022. Preliminary survey results were presented to a joint Working Group 1 (WG1) and the WG4 SEADDA workshop in Braga, Portugal, in May 2022. Discussion of the results at the workshop helped identify bias within the results to improve the methodology and identify other issues, such as confusion with terminology or challenges in completing the survey. In total, 73 received responses were completed to 100%, and only the completed responses were used for this analysis.

3.1. Repositories tested

Survey respondents were asked to nominate a digital repository for testing. The total number of repositories tested was 41, with 30 being tested only once (Table 1). Initially, WG4 had discussed the possibility of limiting the number of platforms tested. Restricting testing to a small selection of repositories had the advantage of comparing their strengths and weaknesses. However, this also had the potential to single out specific repositories where the results highlighted weaknesses, thus causing reputational damage. Instead, the scope of the tasks was limited to ensure a consistent evaluation across multiple platforms. With hindsight, scores within the same repository could have verified the results. On the other hand, the survey has created a list of repositories with archaeological data, several of which are in languages other than English, which is a helpful resource. Future iterations of the methodology could focus on a finite list of repositories to capture data not recorded with the first survey.

View Table 1 List of all repositories tested by respondents

3.2. Demographics

Demographic data were collected to determine whether the user experience varied between user locations or job roles. Primarily, the research area of respondents was in archaeology or data management (Fig. 2).
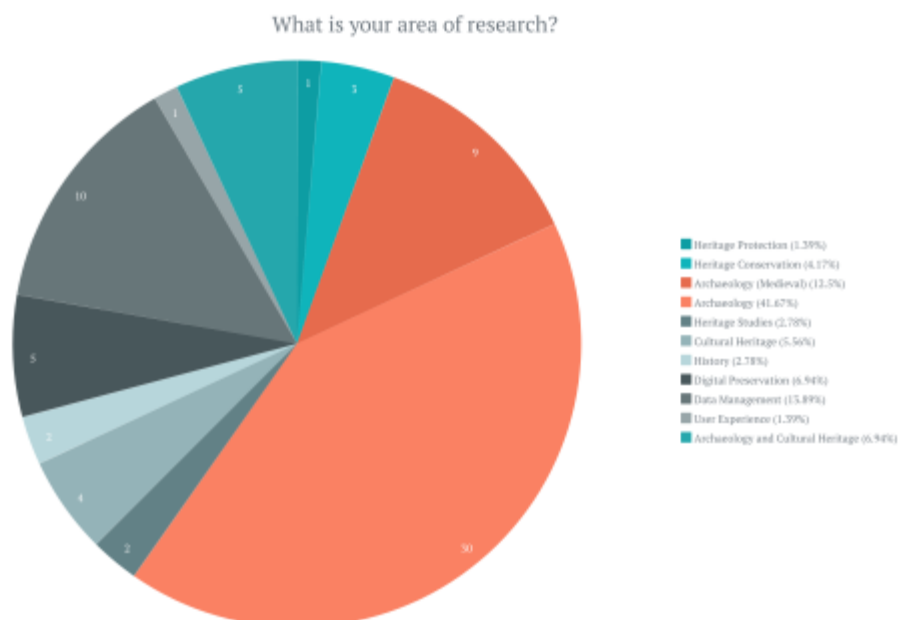
Figure 2: 'What is your area of research?', a list of responses regarding participants' occupations

Responses were received from 21 countries, with a significant response from the United Kingdom (UK) (Table 2). This result may be biased as staff from the Archaeology Data Service (ADS) were asked to complete the survey. Multiple responses from individual respondents further skewed analysis of the survey results. However, making comparative interpretations based on geographic location is problematic owing to the sample size. Therefore regional variations were not considered for this analysis.

View Table 2: : List of countries where respondents were located

3.3. Effectiveness and efficiency

Two indicators of user performance are effectiveness and efficiency. Effectiveness evaluates whether the respondent could complete the tasks on their chosen platform (refer to The British Standards Institution 2016, 11, for further detail). This quantitative measurement is not concerned with how a goal might be performed. The interpretation of these measures is subjective based on the context to which it is applied (The British Standards Institution 2016, 35). Effectiveness is measured by the number of tasks completed and the number of objectives achieved. Task completion is measured as the total number of tasks the user was asked to do and compared to the number of tasks the user could complete. Objectives achieved examine the tasks' complexity and are measured through the weighted values allocated to each task. In this survey, weighting was used to assess the importance of the task rather than its complexity. Unlike measuring the tasks, objectives achieved take partially completed tasks into consideration.

The results indicate that for each scenario, at least half of the tasks could be completed (Figs 3 and 4).

Effectiveness measures - Completed tasks (5 tasks attempted)
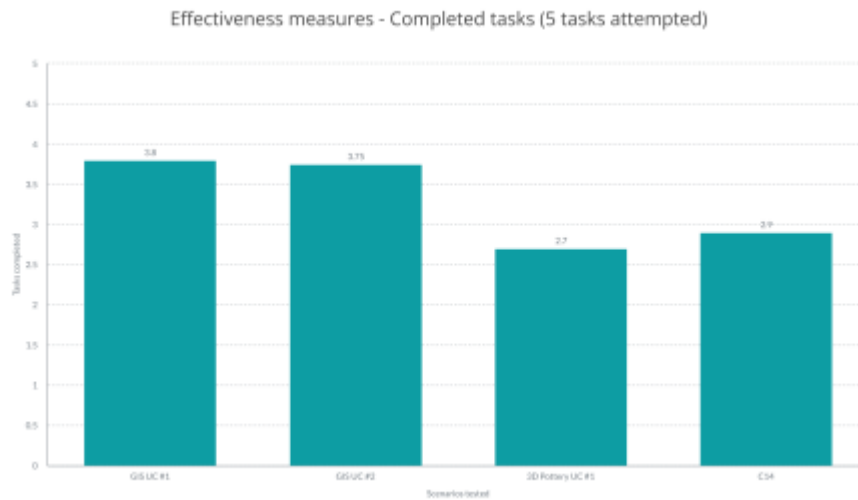
Figure 3: Effectiveness evaluates the respondents' ability to complete the tasks in a given scenario. The measurement is a ratio between the tasks supplied and the number of tasks respondents could complete. This graph displays the effectiveness of scenarios with five tasks

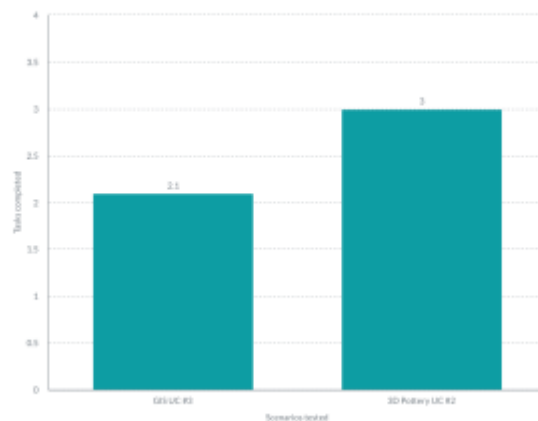Effectiveness measures - completed tasks (4 tasks attempted)

Figure 4: Effectiveness evaluates the respondents' ability to complete the tasks in a given scenario. The measurement is a ratio between the tasks supplied and the number of tasks respondents could complete. This graph displays the effectiveness of scenarios with four tasks

For example, for the 3D Pottery #1 scenario, 15 responses were received, and seven of these respondents could not locate any 3D Pottery models on the platform they were testing. However, these respondents could complete other unrelated tasks, such as filtering search results by material or chronology or locating generic licensing information. Another user commented that despite being unable to find a 3D Pottery model, they completed the tasks using other data from the platform they were testing. Statistically, should responses where the appropriate data not be

located be included in the results? In the case of 3D Pottery data, this would decrease the average task completion score from 2.7 to 1.7. However, the GIS scenario scores would remain unchanged, and the radiocarbon scores would decrease from 2.9 to 2.7 tasks completed. Given the difficulty locating 3D Pottery models, a score of 1.7 better reflects the case in trying to work with this type of data.

Efficiency is measured to determine the cost of resources required to perform the tasks. Cost can be measured in time, opportunities lost or monetary costs for resources consumed (refer to The British Standards Institution 2016 11-12, for further detail). In this instance, the cost is measured in time. Firstly, efficiency is measured by the average time taken to complete all tasks (Fig. 5).

Figure 5: Efficiency determines the resources required to complete a given scenario. The first efficiency measure was the average time taken to complete a scenario

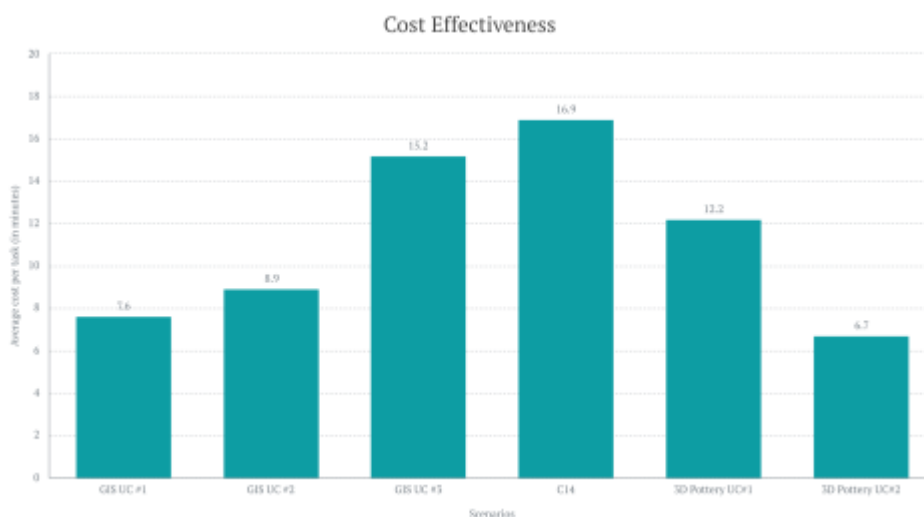The second measure calculates the cost, in time, per task (Fig. 6).

Figure 6: Efficiency determines the resources required to complete a given scenario. The second measure of efficiency is the cost-effectiveness of the user, where the cost of completing each task is considered in relation to the time taken to complete individual tasks

The trend between the two measures is similar for effectiveness. The radiocarbon scenario proved to be the most challenging in terms of efficiency. The radiocarbon respondents took the longest to attempt all tasks, with the longest scenario taking over 80 minutes to complete. The average time taken was 49 minutes, and the cost to complete each task was almost 17 minutes.

3.4. User Satisfaction

The results for user satisfaction were a surprise. There was an expectation that the user community was unsatisfied. The response rate did not reflect this, with responses ranging from neutral to positive (see Table 3 for a comparison of scenarios). Note that the 3D Pottery #2 scenario has been excluded from the comparative analysis as more responses were needed in order to discuss the results in any meaningful way.

Table 3: User satisfaction responses, compared across use case scenarios. Was the platform easy to use?

Was the platform easy to use?

The majority of respondents, 75% (n=55), scored somewhat satisfied or extremely satisfied with the ease of platform use (Fig. 7). This trend was seen across most scenarios, with slightly higher satisfaction with the 3D Pottery scenario. These respondents made comments on design and presentation features. Respondents who gave a low score commented on the functionality of platforms or an inability to locate an appropriate dataset.
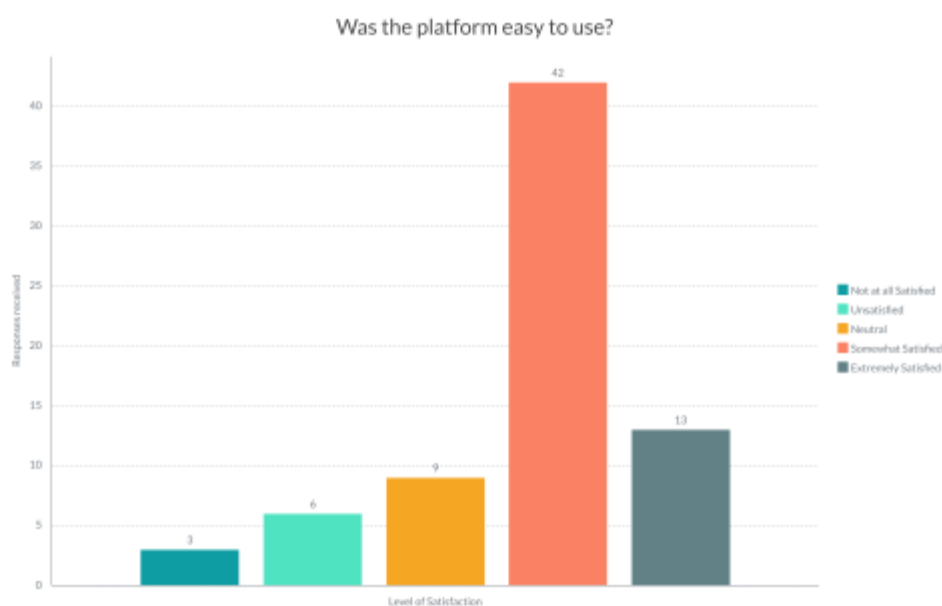

Figure 7: 'Was the platform easy to use?' assessed the respondents' satisfaction when using their chosen platform. Satisfaction was used to determine the degree to which users' expectations were met when using their chosen platform

Were you satisfied with the features offered?

When asked about satisfaction with the features offered, 30% (n=21) of respondents scored a three, which increased to 73% (n=52) if somewhat satisfied scores were included (Fig. 8). Respondents for the GIS web services scenario commented that the survey did not apply to the tested service, with 37% of respondents scoring as not at all satisfied. Other comments stated that the data were available but not reusable, the data could not be exported, or there needed to be more resources for archaeologists. This sentiment was echoed elsewhere with respondents of other scenarios who scored more neutrally across the board.
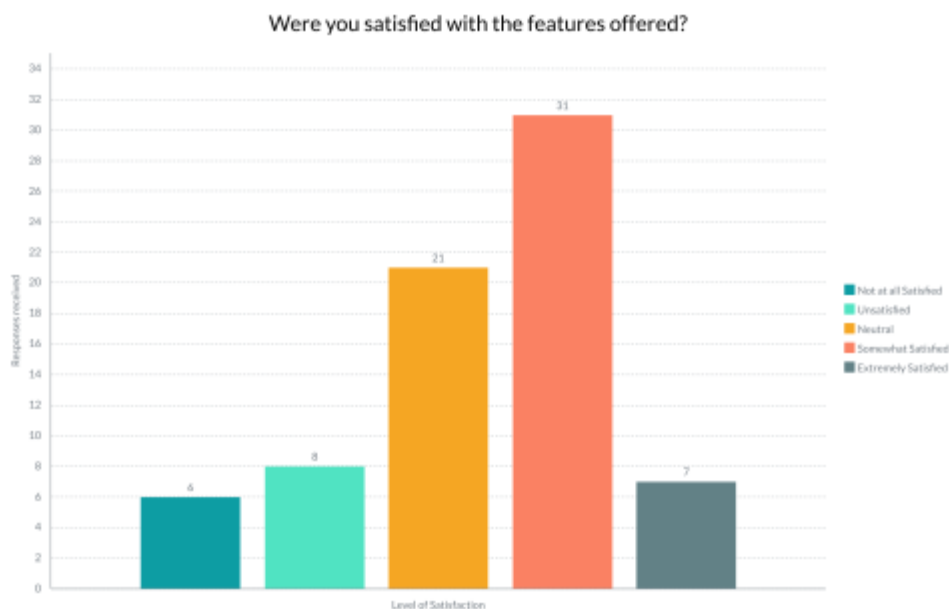

Figure 8: 'Were you satisfied with the features offered?' assessed the respondents' satisfaction with the features offered on their chosen platform. Satisfaction was used to determine the degree to which users' expectations were met when using their chosen platform

Did you enjoy using the platform?

In general, a low response rate was given for this question, particularly for the GIS scenarios, with 25% (n=4) of respondents selecting unsatisfied or not at all satisfied (Fig. 9). The comments noted a need for more available datasets and a need for more functionality on the platform being tested. However, typically across the other satisfaction measures, the GIS scenarios score more highly, with higher task completion rates. This suggests that functionality does not necessarily equate to user satisfaction with the tools provided.
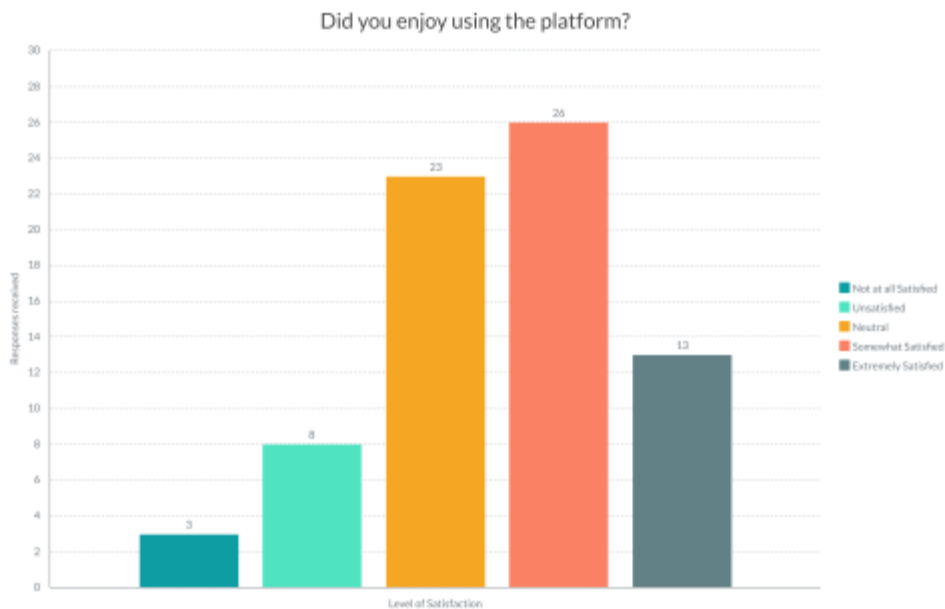
Figure 9: 'Did you enjoy using the platform?' assessed the respondents' enjoyment when using their chosen platform. Satisfaction was used to determine the degree to which users' expectations were met when using their chosen platform

Did the platform have the data you were looking for?

The response to this question was more positive than expected, with 53% (n=39) of respondents scoring highly (Fig.10). However, this question also had the highest dissatisfaction response, with 33% (n=24) of respondents scoring a one or two. Again, there was some variation, with the GIS #2 scenario having a positive response of 87% (n=14) and the 3D Pottery scenario having a negative response of 67% (n=10). The comments from respondents referred to the limited data quality for 3D models or, more generally, about a low quantity of available data. For users with higher satisfaction levels, comments related to search functionality that they found counterintuitive.
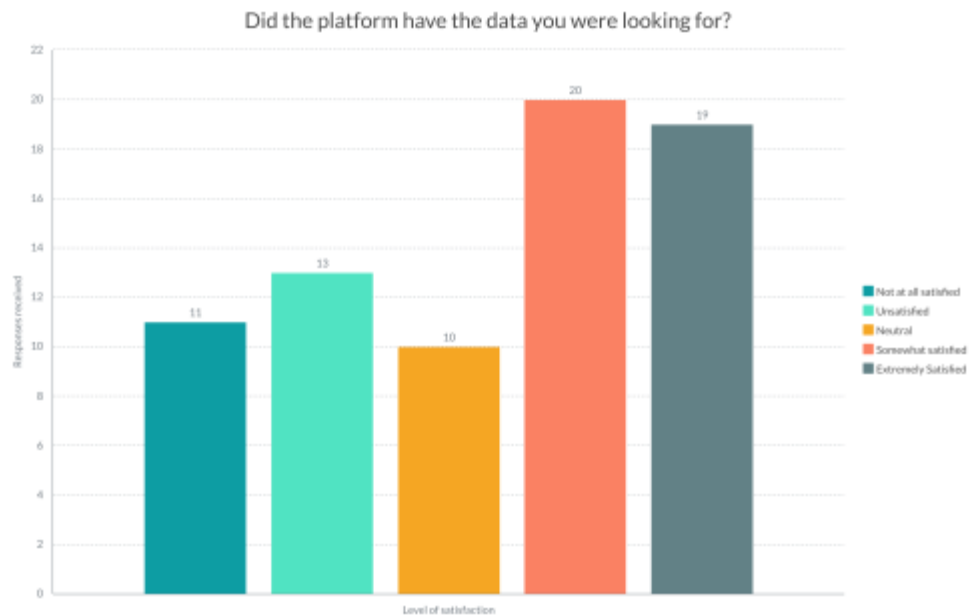
Figure 10: 'Did the platform have the data you were looking for?' assessed the respondents' ability to find relevant data when using their chosen platform. Satisfaction was used to determine the degree to which the user's expectations were met when using their chosen platform

Did the platform help you better engage with the data?

Overall, the response to this question proved neutral (Fig. 11). The exception was the GIS #3 scenario, with 50% (n=5) of its respondents scoring extremely satisfied. Only eight surveys were undertaken for this scenario, so it is not a significant representation. Two respondents scored two for all of the answers related to user satisfaction. In comparison, the remaining six respondents scored highly for the other categories, particularly for the previous data availability question. The same trend is not seen across the other scenarios.
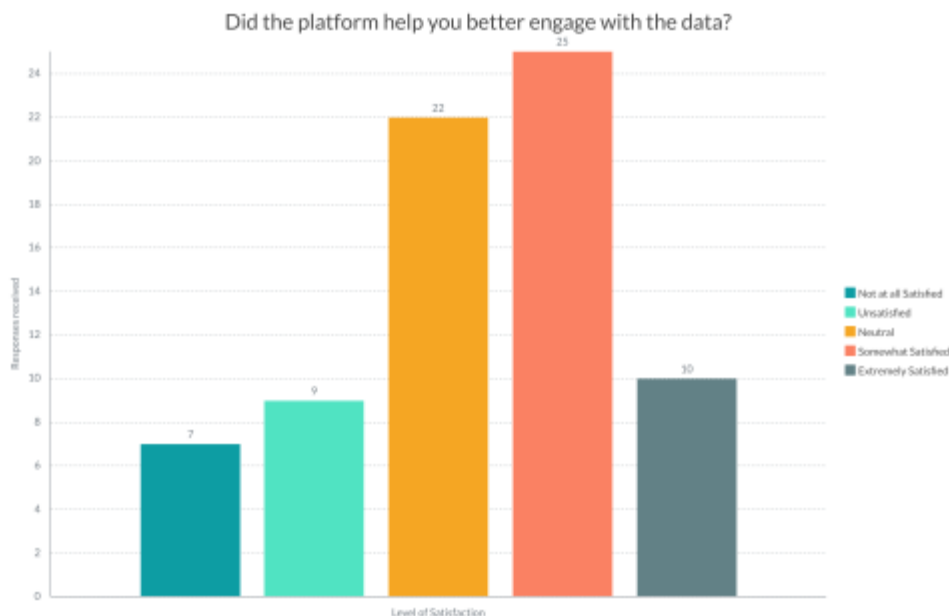
Figure 11: 'Did the platform help you better engage with the data?' assessed whether the respondents' chosen platform improved their engagement with the data offered. Satisfaction was used to determine the degree to which users' expectations were met when using their chosen platform

What was your overall satisfaction with using your chosen platform?

The GIS scenarios scored highly for the final satisfaction question, while the other scenarios were more neutral (Fig. 12). These scores suggest that some data reusers are satisfied with the tools available and can find the relevant data, while others are unsatisfied.
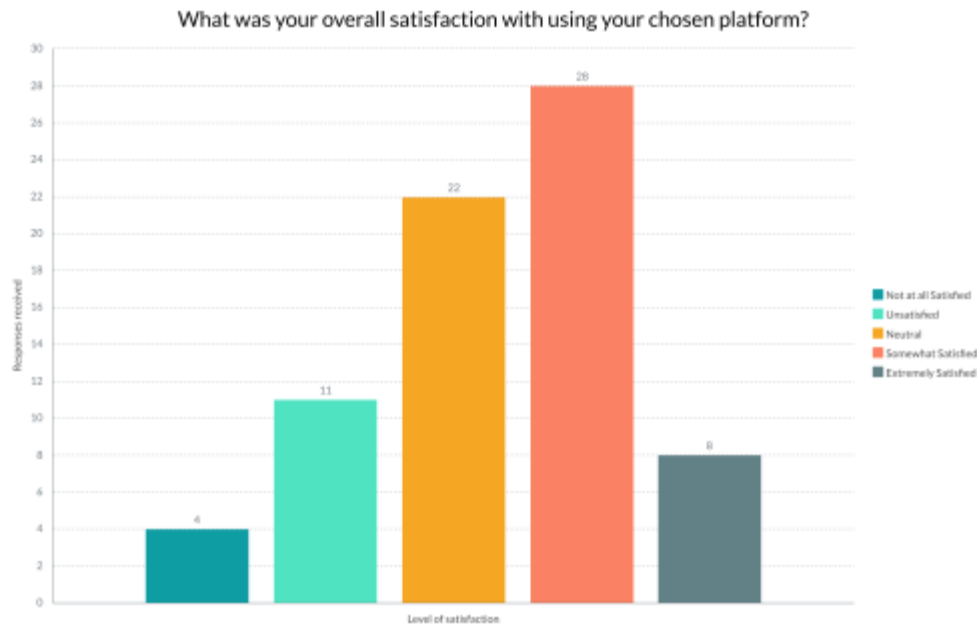
Figure 12: 'What was your overall satisfaction with using your chosen platform?' assessed the respondents' overall satisfaction when using their chosen platform. Satisfaction was used to determine the degree to which users' expectations were met when using their chosen platform

We expected that our user community would be unhappy with the tools available. Instead, the comments suggest that the user community is more confused than unsatisfied. However, the confusion may be a by-product of the survey design. Feedback received at the 2022 workshop pointed out the confusion expressed in the survey and the inability to relate to the scenarios chosen to test the data found on the platforms tested.

3.5. Trust

Like satisfaction, respondents were asked to answer questions on trust using a Likert scale where one is 'Not at all Confident' and five is 'Extremely Confident'. The questions on trust were compulsory, but the comments were optional and provided as a single text box. The original questions were divided into three categories: trust in the system, trust in the data and trust in the metadata. The questions were worded and reworded several times to ensure they addressed what we needed to know. The expectation was that trust scores would be low. However, trust was generally rated more positively compared with the previous satisfaction scores.

Was there enough metadata for you to understand the dataset?

For the first question on trust, 60% (n=44) of respondents were somewhat confident or extremely confident (Fig. 13). The scores varied, with 38% (n=3) of GIS #3 responses scored not at all confident, while 91% (n=10) of Radiocarbon respondents

scored somewhat or extremely confident (see Table 4 for a comparison of scenarios). Respondents who gave a low score commented that they could not locate any metadata or that familiarity with the data was required to make sense of it. Respondents who scored low for this question also had a similar score for the other questions, with eight respondents scoring a one across all three questions.

Table 4: User Trust scores compared across use case scenarios.
Was there enough metadata for you to understand the dataset?
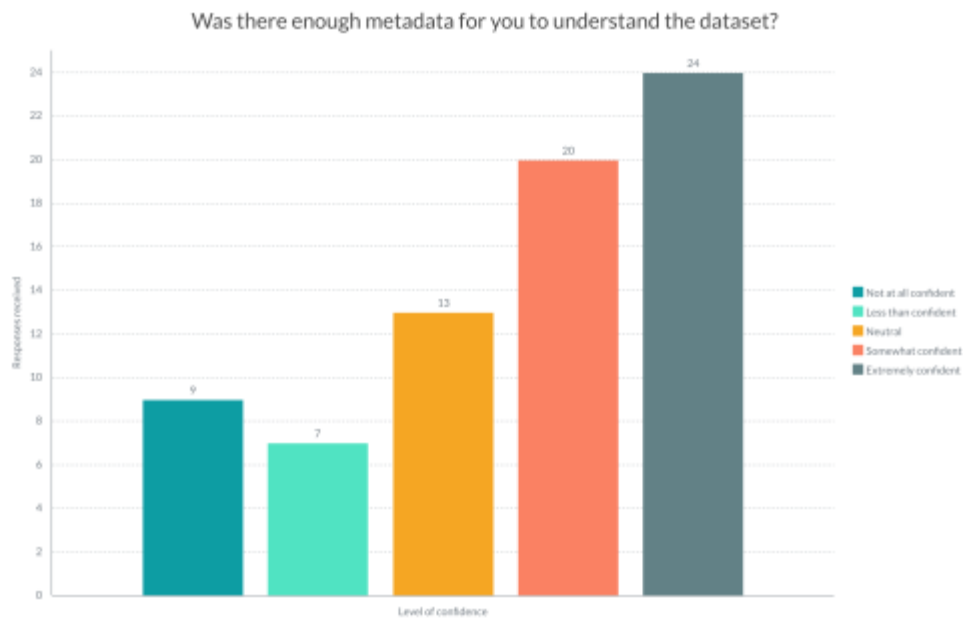
Figure 13: 'Was there enough metadata for you to understand the dataset?' assessed the extent to which the respondents' trusted their chosen platform. Trust was a sub-characteristic of satisfaction that determined the degree to which the user's expectations were met when using their chosen platform

Was the metadata sufficient for your research needs?

The second question scored somewhat evenly (Fig. 14). For each category between neutral and extremely confident, there were 17 responses recorded. Again, there was some variation between the scenarios, with the GIS scenarios scoring more positively and the 3D and Radiocarbon scenarios scoring neutrally. The 3D Pottery #1 scenario had a response rate of 27% (n=4) for not at all confident.
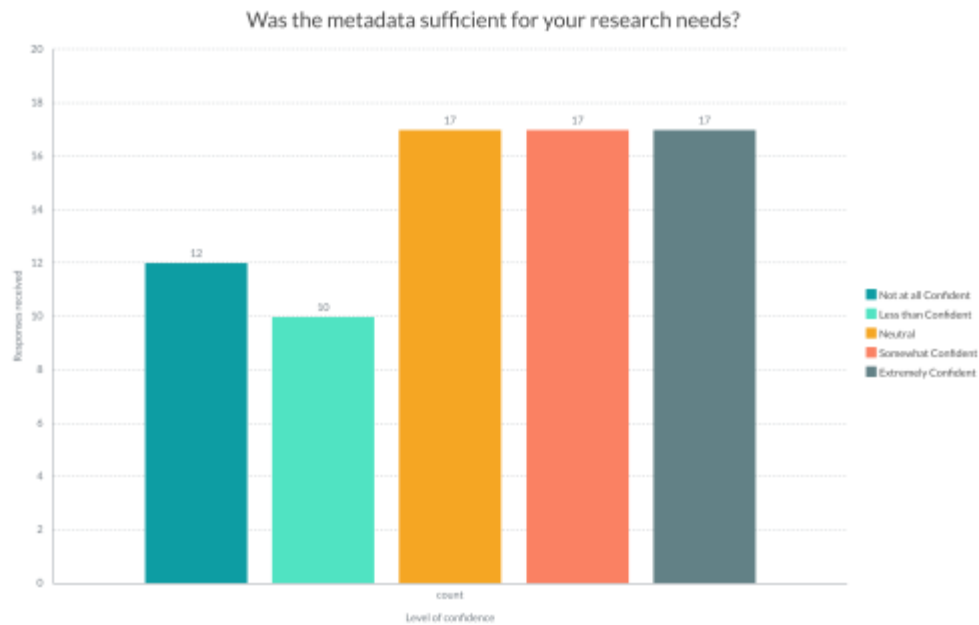
Figure 14: 'Was the metadata sufficient for your research needs?' assessed the extent to which the respondents' trusted the data used for the survey. Trust was a sub-characteristic of satisfaction that determined the degree to which the user's expectations were met when using their chosen platform

Are you confident that you could use the metadata in your own research?

The final measure for trust tended to have a more positive response than the previous two questions (Fig. 15). The response for the GIS scenario is interesting, with 50% (n=4) of responses scoring extremely confident and 38% (n=3) scoring not confident at all. Contrasted with the previous result, fewer GIS #3 respondents thought that there was not enough metadata for their needs, yet they would still be willing to use the dataset. The scores were evenly spread for the 3D Pottery scenario, although there appears to be a slight increase in confidence compared to the previous question.

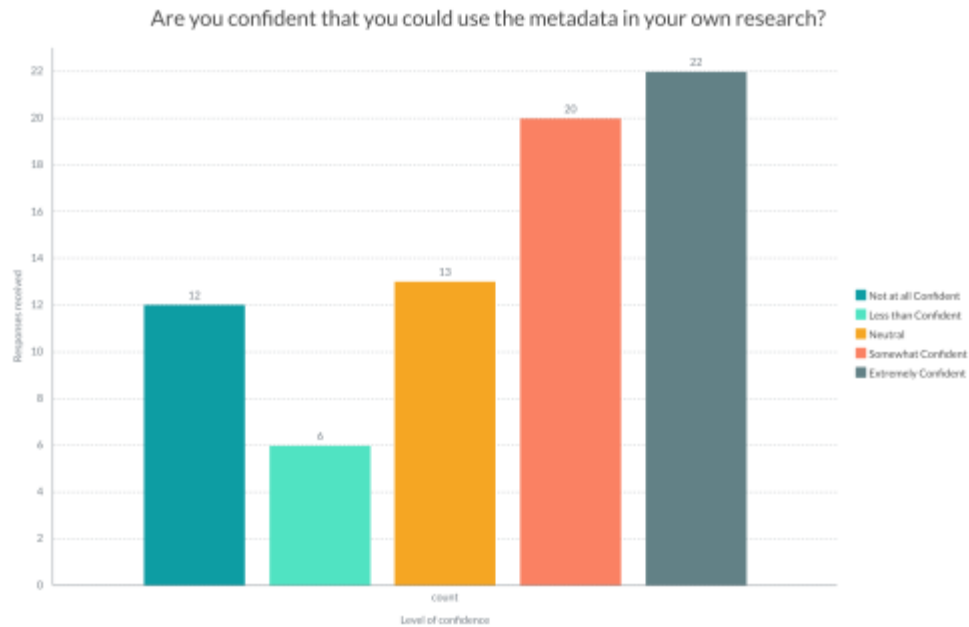Are you confident that you could use the metadata in your own research?

Figure 15: 'Are you confident that you could use the metadata in your own research?' assessed the extent to which the respondents' trusted the metadata or paradata found on their chosen platform. Trust was a sub-characteristic of satisfaction that determined the degree to which the user's expectations were met when using their chosen platform

3.6. Context coverage

Context coverage is the evaluation measure that assesses the overall usability of software in the ISO 25022 standard. For the survey, the data collected addressing context was qualitative, with text boxes asking questions about satisfaction and trust. The first question asked respondents about their operating systems and internet browsers. Primarily the operating system was Microsoft Windows, and the primary browser was Google Chrome. This question was asked to determine if the software used to access platforms impacted the user experience. However, these data were not used further in this analysis.

The following three questions asked the user about the flexibility of the data and their personal preferences. Half of the respondents said yes when asked if they could export the data for use on another platform (Fig. 16).

Could the data be exported and used with different software or a different platform?

Legend:
- Yes (51.39%)
- No (11.11%)
- Partially (8.33%)
- With Difficulty (6.94%)
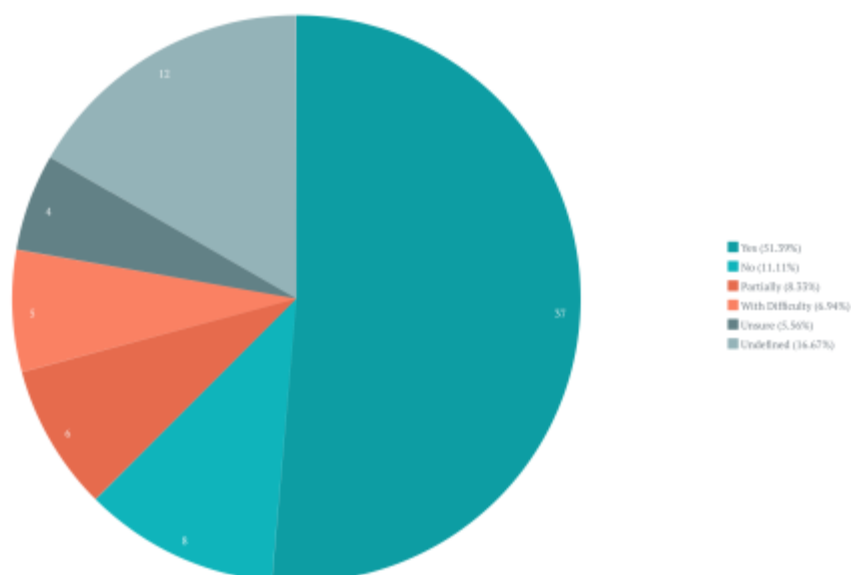- Unsure (5.56%)
- Undefined (16.67%)

Figure 16: Flexibility measured the extent to which the data could be used in additional contexts, in this instance, whether the data could be exported for use with different software. The question provided a text box, and the responses are summarised here. Respondents who failed to answer the question were marked as 'undefined'

Some respondents gave one or two-word responses, while others provided additional detail. For example, one respondent commented that the data could not be exported but could be copied and pasted in a tabular format. Comments also included an objection to registering to view data. The next question asked if the exportable format was their preferred option (Fig. 17).

What file types could be exported? Are they the ones you would prefer?

Legend:
- Yes (12.33%)
- No (5.48%)
- Prefer a different format (17.81%)
- Positive response, no elaboration (54.25%)
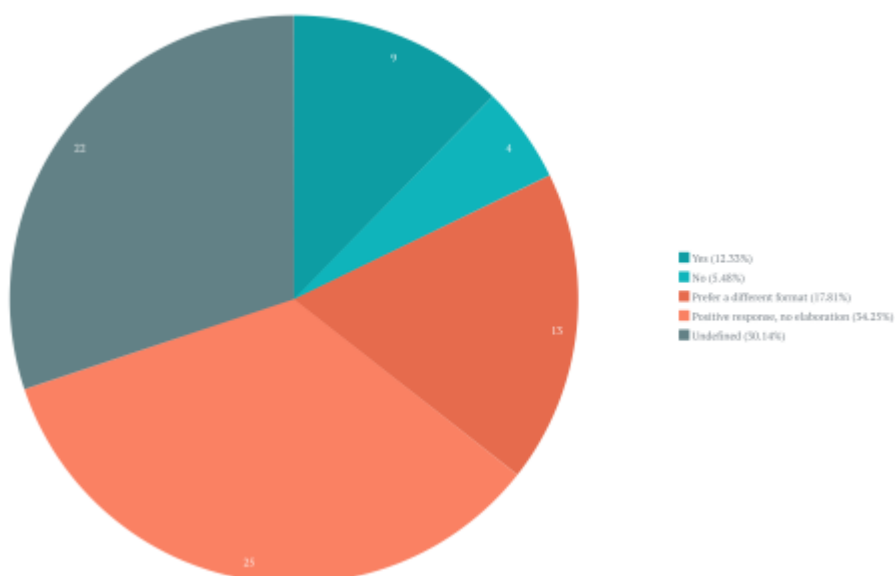- Undefined (30.14%)

Figure 17: Flexibility measured the extent to which the data could be used in additional contexts, in this instance, by considering whether the respondent was satisfied with the file

formats available for download. Respondents who failed to answer the question were marked as 'undefined'

For this question, 12% (n=9) of respondents said emphatically yes, while 30% (n=22) responded positively but did not confirm. For example, 'Most of them are in OBJ, which is a common solution'. The final question for this section asked how flexible the data were or whether further modification was required before it could be used. The responses to this question were mixed, with 36% (n=26) of respondents saying yes. The remaining responses either answered negatively or gave no response at all. As one respondent pointed out, the response to this question depended on the research question.

The following two questions focused on available documentation from the data platform. First, the respondent was asked to comment on the availability of licensing information (Fig. 18).

Figure 18: Usability determined whether sufficient documentation was provided to use their chosen platform. The question posed whether licensing information was available. Respondents who failed to answer the question were marked as 'undefined'

This question received a positive response of 64% (n=49). This result was not surprising. Some repositories have policies in place for applying licensing rules for deposited data. There is also an increasing awareness of the need for licensing with initiatives such as Creative Commons. The next question asked, 'Were most functions explained in sufficient detail in the user documentation and/or help facility to enable you to use them?' (Fig. 19).

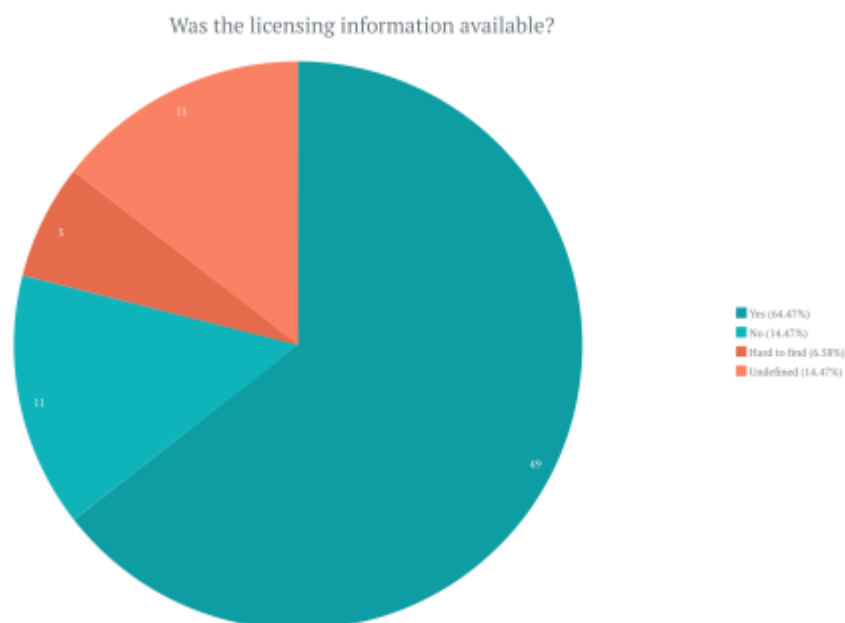Were most functions explained in sufficient detail in the user documentation and/or help ?
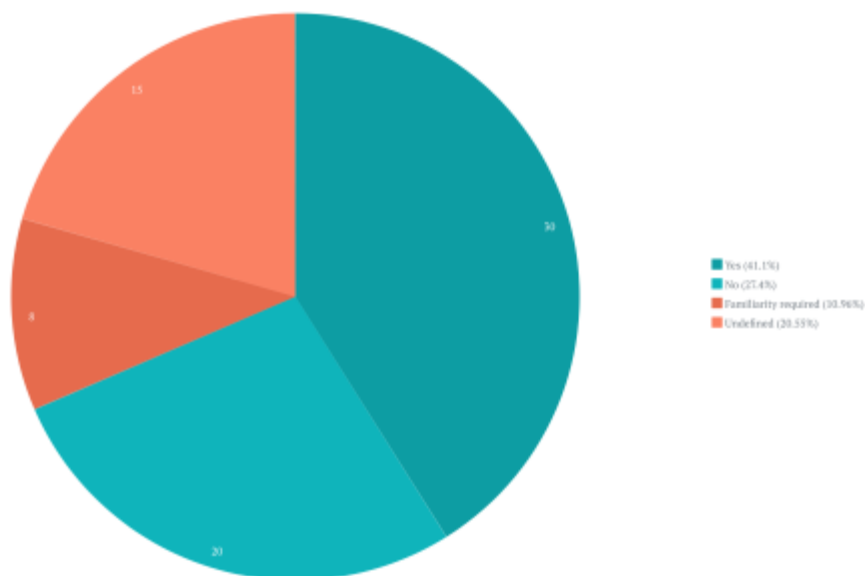

Figure 19: Usability determined whether sufficient documentation was provided to use their chosen platform. The question was whether user documentation was made available to assist the respondent in navigating the platform. Respondents who failed to answer the question were marked as 'undefined'

The result for this question was slightly lower, with 41% (n=30) of respondents answering positively. Some 11% (n=8) of respondents also suggested that the information was available, but familiarity with the platform was required to locate it. Several respondents commented that documentation was not required as the platform was easy to use and understand.

The final question in this section asked the respondent if they knew enough about the data that they could use it in their research. If the answer was no, the respondent was asked a final question to determine what they would need to use the data. There were 19 negative responses, and each respondent provided comments. The comments for this question echoed the previous comments made. Respondents struggled with a lack of metadata, documentation or licensing information, a dislike for registering to download the information, and data being in the wrong format or summarised in tables rather than providing the raw data.

3.7. User comments

Further analysis was done on user comments, and seven categories emerged. One of these categories was self-reporting. Several respondents noted that their familiarity with a platform influenced their high scores for satisfaction rates. However, one respondent who self-reported and selected extremely satisfied for the categories of ease of use features offered and pleasure using the platform selected scores of not at all satisfied and neutral for the remaining three categories. Despite being involved with the tested platform, they scored low for all categories. The same

respondent scored 'trust' at varying levels as well. Regarding bias, the respondent still critically evaluated their tested platform.

Search functionality was often mentioned. Specifically, respondents noted a desire to enhance search functionality to allow searching by data type or to use search tools such as Boolean strings. The second issue was the quantity of data. Some platforms have plentiful data, yet searches yielded no results. Non-heritage platforms (for example, government spatial platforms) lack archaeological data. In contrast, other platforms need more data for searching. Some users stated that searches could have been clearer. One user noted that the platform has good functionality but a poor user interface, while others noted that the user interface was aesthetically pleasing, but the platform lacked functionality.

A recurring theme was resentment at having to register or log in to access data. One user that resented logging in said that some files still required special access. Another respondent commented that they would like to have summary reports available for the data to help determine whether they want to proceed with registering for an account. The tested platforms approach registration in various ways. For example, the ADS does not require registration at all, while DANS has an option for depositors to offer their data as fully open or to require registration before downloading.

The platform audience was addressed several times in the comments. Some platforms have a pleasing interface but are aimed at public audiences. As a result, these platforms lacked the functionality required for academic research; another category is reusability. Several platforms had data available; however, owing to the formats provided, the data were not reusable, or the data was not downloadable, while some files were reported as corrupted.

The final category was metadata. Respondents were asked specifically about metadata, and 23 responses were received. Several respondents stated they could not locate the appropriate datasets or the accompanying metadata. In addition, some of the metadata was unusable or corrupt. Barriers to using the metadata included an assumed familiarity with the dataset, a need for more detailed contextual information and language barriers. There was also a concern over how the metadata was controlled or mediated. One respondent was concerned that platforms were not enforcing a requirement on the depositors to include detailed metadata. Another respondent wanted the metadata to be mediated by museums as it is scientific information. One respondent also suggested that the data depositor should describe how their data might be (re)used in the future.

4. Discussion

Research question one was addressed by developing a methodology for documenting the user experience when reusing digital archives and repositories. The validity of the methodology was tested by employing a survey. WG4 attempted to distribute the survey to professionals in archaeology and heritage. However, the distribution method proved limiting as most respondents were participants in the SEADDA COST Action. While contributions beyond SEADDA were welcomed and

encouraged, the time frame for collecting responses could have been better. In future, a more comprehensive distribution method could be used to collect data from a broader audience. This approach can be addressed in future survey iterations and would prove helpful. The primary drawback is that the responses presented here are less broad than WG4 anticipated.

Other limitations to consider are bias and survey length. Bias within the survey was addressed throughout the survey design process. However, the uneven uptake regionally and multiple responses from individuals limited the value of analysing geographic patterns. Additionally, with a 20% dropout rate, the survey needed to be shorter. Dropouts often occurred in the final qualitative section of the questionnaire. The survey length was considered during the design phase, but a tension existed where WG4 wanted to stay true to the structure of the ISO 25022 standard. Improvements could be made here as there was occasional duplication between questions in the user scenarios and the qualitative questions.

Feedback received during the 2022 workshop emphasised that several users felt overwhelmed by choosing which platform to test. Again, this was something that had been discussed during the design phase. Limiting the platform choice would have enabled comparative data between responses received. However, by giving respondents a choice, we collected a list of 41 platforms, highlighting one of the challenges archaeologists face when locating data. Where should you start trying to find data? While *findability* has been addressed in FAIR and by significant investments made in recent years, knowing where to start or which resource to use is still a barrier. Aggregation websites, such as ARIADNEplus, can mitigate some of these challenges but are limited to partners who want to participate.

Research question two aimed to identify the barriers to the use and reuse of archaeological archives and repositories. The main difference of our methodology, compared to the previous research on archaeological or heritage data archives, is not the focus on the quality, as software, system or data characteristic, but on the quality of interaction between the archive (including software solutions, user interface, data, metadata, paradata) and the user. Furthermore, we sought to understand how specific users can use the digital archive to meet their distinct needs and goals. Following this approach, the identification of barriers to the reuse of archaeological archives could be divided into two parts (i) the characteristics of digital archives and repositories (as measured trends), disabling the user's ability to use and reuse data; (ii) the reasons, explaining why these characteristics are developed and exist in digital archives.

Data analysis revealed several barriers that hinder a user's ability to reuse data. These five 'disabling characteristics' relate to flexibility, licensing, documentation, functionality and the quantity and quality of data.

i.   **Limited possibilities for data export**

The response rate for data flexibility provides an opportunity for improvement. While 50% of respondents could export the data, only 12% were satisfied with the file formats offered for download. One respondent astutely pointed out that

the depositor, not the platform, determines data structures. Data repositories provide freedom for depositors to decide which file formats to use. However, this freedom must be translated for data users and impacts reusability. Data flexibility scored low in the survey; only 36% of respondents were satisfied with the file formats offered. Comments focused on the need to manipulate or reconfigure data before it could be used. PDF files were the most disliked, as they offer no flexibility despite being downloadable. In addition, respondents hated having to register to access or download data. Archaeological data is diverse, with various file types and preservation needs (Richards et al. 2021). How do we encourage depositors to provide formats that data reusers need? Repositories can provide guidance, such as the *Guides to Good Practice* written by the ADS and *Digital Antiquity* (Archaeology Data Service and Digital Antiquity 2013). However, if those guidelines became enforceable mandates, it could imply that archaeologists hold no personal responsibility for data quality.

ii. **Availability of licensing information**

Many data archives and repositories have policies in place for applying licensing rules to deposited data, demonstrated by 64% of respondents being able to locate licensing information. There is also an increasing awareness of the need for licensing with initiatives such as Creative Commons. However, respondents' struggled to find licensing information in general, with information buried in linked pages, unclear details provided or different licensing terms applying to individual items in a single archive. Despite the positive response rate, more work must be done in this area.

iii. **Limited explanation of archive functionality**

Platforms should provide user documentation or other mechanisms to assist users in exploiting the full functionality available to them. Unfortunately, the response rate for locating this user documentation was much lower than the licensing question, with only 41% of respondents able to locate appropriate documentation. In addition, 11% of respondents commented that familiarity with the platform was required to find the required documentation. Finally, several respondents commented that documentation was unnecessary as the platform was easy to use and understand.

iv. **Malfunctioning user interfaces**

This survey aimed to explore the functionality of underlying search infrastructure for digital archives and repositories. One barrier to testing the functionality of these interfaces was that respondents struggled to find an appropriate dataset. For the 3D Pottery scenario, half of the respondents could not locate an appropriate dataset. This result varied across the different scenarios. For example, in the GIS #2 scenario, 75% of respondents could locate a dataset, with the remaining 25% responding 'partially' to this

question. The search tools to find data were also considered lacking. Search functionality was mentioned in the comments. Respondents desired to enhance search functionality to allow searching by data type or to use search tools such as Boolean strings. One user noted that the platform has good functionality but a poor user interface, while others indicated that the user interface was aesthetically pleasing but lacked functionality. Many platforms require users to register to view or export data. Respondents resented registering or logging in to access data, with some files requiring special access. Another respondent commented that they would like to have summary reports available for the data to help determine whether they want to proceed with registering for an account. Within archaeology, there can be a need to protect sensitive data, such as geographic information, for sites at risk of looting (Sheehan 2015, 188). Data can adhere to FAIR without being completely open, recognising that open is only sometimes appropriate (Wilkinson *et al.* 2016). The tested platforms approach registration in various ways. For example, the ADS does not require registration at all, while DANS has an option for depositors to offer their data as fully open or to require registration before downloading.

v. **Limited quantity and quality of data/metadata/paradata**

The quantity of data available would appear quite large, given that respondents tested 41 distinct platforms. However, there is a perception of a lack of available data. This perception could be attributed to the inability to locate data, as discussed above. While some platforms have plentiful data, searches yielded no results. In contrast, respondents commented that some platforms have adequate search functionality but need more data for searching. In addition, non-heritage platforms (for example, government spatial platforms) lack archaeological data. The audience for the platform also influences the type of data available, affecting both the aesthetics of the platform, as well as the flexibility of the data.

There was no specific measure for data quality. Trust can be used as an approximation for the perceived quality of the data and the metadata. Anecdotal evidence would suggest that trust in other people's data is relatively low. However, the levels of trust reported by respondents were much higher than their satisfaction with the platforms tested. For the first question, 'Was there enough metadata for you to understand the dataset?', 60% of respondents were highly confident. The scores vary between user scenarios, with 90.9% of C14 respondents being satisfied with the metadata available but less so with GIS respondents, where 37.5% were not satisfied. Despite this, the GIS #3 respondents were still willing to use the dataset without the requisite metadata. One interpretation of this response is that there had been a shift in trust.

Researchers previously preferred to get data from colleagues or based on an individual's reputation (Faniel *et al.* 2013, 301). Data repositories now carry authority based on their reputations, resulting in a shift in trust from the individual to the institution. It was reported that some users in the community

expect that the data they find in repositories has been peer-reviewed. While journals such as the *Journal of Open Archaeology Data* and *Internet Archaeology* publish peer-reviewed data articles, this is not the case for all data held in digital archives or repositories. Data repositories can apply for Core Trust Seal accreditation that evaluates data repositories against 16 requirements (CoreTrustSeal 2017). Again, this measure does not apply to all archives or repositories. Alternative interpretations need to be offered for this shift in trust and need to be explored further.

Testing a range of scenarios aimed to discover if the type of data respondents worked with impacted the user experience when searching in digital archives. The highest response rate was for the GIS scenarios. These respondents also scored highly on locating data, effectiveness, efficiency, satisfaction and trust. Comparing the use case scenarios, the main 'disabling characteristics' are:

1. **GIS use case scenarios #1, #2 and #3**

   Respondents testing the GIS scenarios had the most difficulty downloading copies of the data without registration, finding and exporting the data in a reusable format, filtering searches by topic (for example, archaeological site type, sampled material, user-defined area or chronology), finding clear licensing information for data reuse, and finding information (discovery metadata) and adding and using Web Services.

   Other GIS-related 'disabling characteristics' noted by respondents were:

   - Products such as BaseMaps did not work well, as the map looked cluttered
   - Data could not be integrated as a web service without purchasing additional GIS software
   - Viewing data was limited to specific parameters
   - Downloading files individually, rather than being able to bulk download

2. **Radiocarbon (C14) use case scenario**

   Respondents testing the C14 scenario had difficulty achieving the primary task of downloading digital files containing stratigraphic relationships between stratigraphic units with associated C14 dating evidence. In addition, respondents could not search or find specific radiocarbon sample data within the digital platform using sample numbers. Respondents did not identify other radiocarbon-related specific 'disabling characteristics'.

3. **3D Pottery use case scenarios #1 and #2**

   Regarding the tasks in these scenarios, respondents had the most difficulty with exporting 3D data in a reusable file format; filtering or searching by chronology, material, geography, or typology for pottery; finding high-

resolution 3D models of archaeological pottery; and using 3D visualisation platforms to view and store their 3D data

Other 3D Pottery-related 'disabling characteristics' are:

- o It was possible to find the models as records, but the illustrative materials are limited to print screens of the model, as opposed to the real models
- o Primarily, 3D models were only available for fast visualisation purposes in formats such as 3D PDF files or SketchFab embedded files.

5. Future Directions

The Quality-in-Use survey aimed to collect data on the qualitative experiences of data users. Some limitations were a result of sticking closely to the ISO 25022 standard. The feedback from the survey was mixed. When the results were presented to SEADDA members at a workshop in Braga, the conclusion was made that the methodology works, but the execution could be improved. The limitations included the survey format, user-friendliness, overwhelming participants with choice, distribution method, and a need for comparative data.

Participants found choosing a platform for testing overwhelming. WG4 had made a conscious choice not to single out individual repositories or narrow the scope of results received. However, in future iterations of the survey, this should be reconsidered. In addition to user comfort, narrowing the choice of platforms for testing would provide more opportunities to compare survey results, which was lacking in this first survey. The survey could have been more user-friendly, and at times there was duplication between the tasks asked of respondents. Standardising the tasks across new data types or making tasks more complex could address this. There is an opportunity to perform a second survey and advance the development of the Quality-in-Use methodology.

Another limitation of the survey was the distribution method. The survey was primarily distributed to SEADDA members, thus failing to address the variety of user communities in archaeology. Future directions should be discovering who uses archaeological data and what data reuse means to them. Qualitative data should be collected to understand their experiences. However, to facilitate this, there is a need to move away from the original format of a survey. Future directions should acknowledge the feedback provided and collect data on a more personal level, such as through workshops or interviews. Guided workshops are a way to walk participants through the methodology, where they can then test a dataset and answer survey questions. This approach provides an opportunity to address different user communities, for example, by asking a group of commercial archaeologists or museum visitors to participate and understand how their expectations vary. Further opportunities to develop the methodology would be to expand and test other data formats, expand on use cases by increasing the complexity of tasks, evaluate data quality, and broaden audience participation.

6. Conclusion

The Quality-in-Use survey evaluated the user experience when searching for archaeological data. The survey provided interesting results, particularly concerning shifts in trust, the ability to locate data and the differences between working with different data types.

To address research question one, WG4 examined the methodologies and best practices found in scholarly literature and reports related to understanding and documenting the quantitative and qualitative reuse of archaeological archives. The decision was made to interpret reuse as a quality feature of the archaeological archive. Conceptually, product quality can be recognised and defined differently (Garvin 1984). However, in this case, WG4's interest in the use and reuse of digital archaeological data and the exploration of how digital archaeological archives can better respond to user needs provides a basis to focus research on a user-based conceptual product quality approach (Garvin 1984 27) and to relate the methodology to the information system quality measurement standardised approach defined as '...calculation, performed to combine two or more values of quality measure elements…' (International Organisation for Standardisation 2019 3). The standardised approach to information system quality measurement comprises different measurement frameworks such as Quality-in-Use, system and software product quality, data quality, and IT service quality (International Organisation for Standardisation 2019 v). Seeking to better respond to this research challenge, WG4 created a methodology to understand and document the reuse of archaeological archives based on a Quality-in-Use measurement standardised framework (International Organisation for Standardisation 2016). This framework also responds to the need for the complexity of research to investigate quantitative and qualitative kinds of reuse.

The discussion section of this article responded to research question two by outlining the barriers to the use and reuse of digital archaeological archives and repositories. The methodology defined quality as the interaction between the archive, software solutions, user interface, data, metadata, paradata and the user. Following this approach, the identification of barriers to the reuse of archaeological archives could be divided into two parts (i) the characteristics of the archives that limit the use and reuse of archaeological data; (ii) the reasons why these characteristics are developed and exist in digital archives. The identified characteristics that create barriers to data reuse relate to the 14 measures set out in ISO 25022. These barriers fall into the categories of flexibility, content completeness and user trust.

i. Limited possibilities for data export (data could not be exported or downloaded at all or can be exported in copy-paste or single selection (not in bulk) mode only; there is a limit to the number of records a user can download at one time or even one file only at a time can be downloaded; there is no possibility to find or export the raw data in popular open or industry formats (e.g. JSON, GeoJSON, SHP, CSV, XML); data could be exported but in 'close' (e.g. PDF) file format only; for data export user must follow the included link to the source).

ii. Availability of licensing information (no licensing information at all or information is 'buried' in the linked page; information not clear, not related to the generally known licences or clicking the copyright icon opens useless information (e.g. general

comments on the register of monuments, sources of presented data and penalties for illegal excavations); not all the data archived is licensed under the same terms; licensing information is not displayed at the item level; very restrictive licence granting – all copyright to all content).

iii. Limited explanation of archive functionality (no user documentation, explanation, introductory video, and helpdesk contacts; no information on how to export the files).

iv. Malfunctioning user interfaces (the language barrier (no English translation available); not so efficient and rapid search (lack of filtering criteria, direct search, the possibility of Boolean queries); malfunctioning search; need of special permission to be able to access (some of) the files (without explaining how you can get this permission); not intuitive user interface; some buttons are presenting but not activated; the functionality is dependent on licensing and does not work without copyright licensing software).

v. Limited quantity and quality of data/metadata/paradata (lack of corresponding/context-related archaeological and technical data/metadata/paradata; data are not attributed to sites; data and metadata are insufficient or described in general terms only; no information exists as to metadata correspondence with metadata standards; a minimal number of datasets available online; only a couple of datasets in the archive are relevant to archaeology; is low quality/resolution of visual 2D and 3D data).

The reasons why these 'disabling characteristics' are developed and exist in digital archives are related to the evolution of digital archaeological archives in the context of the change in the field of scholarly archaeology, digital information technologies and related societal, regulatory, and institutional structures. The main reasons are:

1. The institutional 'ownership' of digital content, understanding the 'ownership' as the kind of organisational culture opposite to sharing, open access, participatory or crowdsourcing approaches. This relates to the authoritarian ('top-down') curatorial approach, in which the creators of the technological solution are perceived as experts who know the public's needs and create the product meant for them. This approach encourages institutions to limit digital data access, submit minimum datasets online, or submit strongly copyrighted, low-resolution or watermark-protected visual data or data in 'close' (not reusable) formats.

2. The pre-digital era regulatory framework provides the statutory basis for defining specific institutional and societal roles in digital archaeology. In the age of ubiquitous computing and networking, this regulatory framework can be seen as a 'design failure' because of its misalignment with the current technology capabilities and technology-related practices. Digital data, metadata and paradata in different countries are treated in various ways as intellectual property objects, intellectual property licences are interpreted differently (e.g., Creative Commons), and there are different regulations of items such as so-called 'orphan works'. The application of the General Data Protection Regulation is also related to this regulatory framework. This encourages the submission of limited, unclear, or not precise licensing and licence-related information and limits access to data hypothetically sensitive in the case of GDPR.

3. The transfer of data, metadata and paradata from one system to another, using different metadata schemes and their mapping models (e.g. transfer of data to Europeana from local systems, mapped with the EDM), impacts the completeness and quality of digital content. The link between the original metadata set and the end user is partially lost during this process. The quality of the data, metadata and paradata received by the user depends on the mapping quality at the time of metadata transfer.

4. Institutional traditions also impact the quantity and quality of data/metadata/paradata. Libraries, museums, archives, and scholarly and heritage protection institutions often have long histories. The collections they house, the language they use to describe these collections, and the data and metadata standards used by institutions differ as products of that historical legacy.

5. Obsolescence of technology and related solutions affect the quality of data, possibilities of use and reuse, malfunctioning user interfaces and limited explanation of archive functionality. Most modern technological solutions that heritage institutions create are WEB1 and WEB2 based and related to WEB1-based thinking on how technologies work in society. This kind of thinking encourages the perception of the digital archaeology dataset as a text-orientated (less visual), tabular (not networked), SQL relational structure (not, e.g. graph) based archive with no need for a good and well-designed user interface, helpdesk or good relationships with social network structures. This kind of thinking (in some cases) is supported by project-based financing facilities, for whom the 'clear and transparent' waterfall software development lifecycle model is much more preferable than, e.g. Agile.

6. Lack of understanding of why the metadata and paradata are needed is due to the limited quantity and quality of data/metadata/paradata and malfunctioning user interfaces. People working in libraries, museums, archives, and scholarly and heritage protection institutions lack knowledge of information management and information work. This shortcoming is also mirrored in the information systems they design. This kind of misunderstanding (in some cases) is supported by a lack of communication and shared knowledge between heritage and IT professionals.

The final research question was to recommend how archaeological archives can better respond to user needs. The solutions are related to eliminating the reasons for 'disabling characteristics' or developing digital tools to reduce the negative impacts of these reasons.

Maximising the discovery, use and reuse of archaeological data is imperative for both heritage professionals and non-professionals. Crucially, archaeological archives need to optimise search capabilities and user interfaces to respond better to contemporary society. The barriers have been mentioned here, but there is also some worldview-related baggage. Part of the barriers may emerge from the perceived loss of control over the use of digital objects, possibly prompting archaeological institutions to limit the access and use of their digital assets, thus hampering the realisation of wider societal value digitisation is expected to bring and decreasing the possibilities of creating additional values via 'sharing economy' models. Most digital archives created by archaeological institutions are based on an authoritarian, 'top-down', curatorial approach, according to which the creators of the archive are perceived as the expert who knows the public's needs and creates the product offered to them. Suppliers and customers are in strongly fixed positions, where archive creators are active suppliers and members of archaeology-related professions, while non-professional communities are passive consumers of archives without the ability to use participatory tools, reuse data or co-create digital content. Participatory models need to change worldview-related characteristics and use democratic, 'bottom-up', open, crowdsourced archives. Marttila and Botero (2017 97) express that '...co-design activities and infrastructure strategies in relation to a broader interest in advocating not only the preservation of and access to digital cultural heritage, but, more importantly, enabling collaboration, to support the emerging practices of diverse user groups, and to contribute to cultural commons…' which will strengthen the pluralistic consensus-based paradigm in the digital

archaeology sector. To achieve this, understanding Quality-in-Use can be useful both at a conceptual level, as a driving concept, and as a measurement tool.

This survey started at the end of one process and the start of another, after data have been archived and before reuse. The problem is multi-faceted, and other methods can be used to try and understand the problem. Despite remaining work, the survey demonstrates that archaeological data are available and that the search tools archaeologists have at their disposal do work. However, challenges persist, particularly in knowing where to look. This problem is not the same as findability, as defined by FAIR principles. Further development and training will address this issue. The Quality-in-Use methodology developed here could be developed into a quality assurance workshop, in a more controlled environment with more narrow parameters. Special consideration should be given to the target audience, the workshop format and the platforms being tested. Platform familiarity is crucial for increasing usability, and professional development and training can enhance user comfort and proficiency.

Acknowledgements

Data

quantitative-survey-results
This CSV file contains the quantitative data generated as a result of the SEADDA WG4 Quality-in-Use Survey.

key-legend
This CSV file describes the fields and numeric values expressed in quantitative-survey-results

task-weighting
This CSV file lays out the weighted values assigned to each task in the Use Case Scenarios that were used to calculate effectiveness and efficiency.

Bibliography

Archaeology Data Service and Digital Antiquity 2013 *Caring for Digital Data in Archaeology: A Guide to Good Practice*, Oxford: Oxbow Books.

The British Standards Institution 2016 *ISO/IEC 25022:2016: Systems and Software Engineering - Systems and Software Quality Requirements and Evaluation (SQuaRE) - Measure of Quality in Use*, BSI Standards Limited. https://www.iso.org/standard/35746.html#:~:text=ISO%2FIEC%2025022%3A2016%20defines,to%20product%20or%20system%20quality

CoreTrustSeal 2017 *Data Repositories Requirements*. https://www.coretrustseal.org/why-certification/requirements/ [Last accessed: 20 October 2022].

Faniel, I., Kansa, E., Kansa, S.W., Barrera-Gomez, J. and Yakel, E. 2013 'The challenges of digging data: a study of context in archaeological data reuse' in J.S. Downie *et al.* (eds) *Proceedings of the 13th ACM/IEEE-CS joint conference on digital libraries*, 22 July 2013, New York: Association for Computing Machinery. 295–304. https://doi.org/10.1145/2467696.2467712

Garvin, D.A. 1984 'What does 'Product Quality' really mean?', *Sloan Management Review* **26**(1), 25–43.

Geser, G., Richards, J.D., Massara, F. and Wright, H. 2022 'Data management policies and practices of digital archaeological repositories', *Internet Archaeology* **59**. https://doi.org/10.11141/ia.59.2.

Huggett, J. 2018 'Reuse remix recycle: repurposing archaeological digital data', *Advances in Archaeological Practice* **6**(2), 93–104. https://doi.org/10.1017/aap.2018.1

Kansa, E.C., Kansa, S.W. and Benjamin, A. 2014 'Publishing and pushing: mixing models for communicating research data in archaeology', *International Journal of Digital Curation* **9**(1), 57–70. http://www.ijdc.net/article/view/9.1.57/341

International Organisation for Standardisation 2016 *ISO/IEC 25022:2016 Systems and software engineering — Systems and software quality requirements and evaluation (SQuaRE) — Measurement of quality in use*. https://www.iso.org/standard/35746.html [Last accessed: 21 February 2023].

International Organisation for Standardisation 2017 *ISO/IEC 25010:2011 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models*. https://www.iso.org/standard/35733.html [Last accessed: 21 February 2023].

International Organisation for Standardisation 2019 *ISO/IEC 25020:2019 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Quality measurement framework.* https://www.iso.org/standard/72117.html [Last accessed: 21 February 2023].

Lee, J. W., Jones, P.S., Mineyama, Y. and Zhang, X.E. 2002 'Cultural differences in responses to a Likert scale', *Research in Nursing & Health* **25**(4), 295–306. https://doi.org/10.1002/nur.10041

Marttila, S. and Botero, A. 2017 'Infrastructuring for Cultural Commons', *Computer Supported Cooperative Work: CSCW: An International Journal* **26**(1-2), 97–133. https://doi.org/10.1007/s10606-017-9273-1

Marwick, B. and Pilaar Birch, S.E. 2018 'A standard for the scholarly citation of archaeological data as an incentive to data sharing', *Advances in Archaeological Practice*, **6**(2), 125–43.

Richards, J.D., Jakobsson, U., Novák, D., Štular, B. and Wright, H. 2021 'Digital archiving in archaeology: the state of the art. Introduction', *Internet Archaeology* **58**. https://doi.org/10.11141/ia.58.23

Sheehan, B. 2015 'Comparing digital archaeological repositories: tDAR versus Open Context', *Behavioural and Social Sciences Librarian* **34**(4), 173–213. https://doi.org/10.1080/01639269.2015.1096155

Sobotkova, A. 2018 'Sociotechnical obstacles to archaeological data reuse', *Advances in Archaeological Practice* **6**(2), 117–121. https://doi.org/10.1017/aap.2017.37

Wilkinson, M.D., Dumontier M., Aalbersberg I.J., Appleton G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., Bonino da Silva Santos, L., Bourne, P.E., Bouwman, J., Brookes, J. A., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. and Mons, B. 2016 'The FAIR guiding principles for scientific data management and stewardship', *Scientific Data* **3**, 160018. https://doi.org/10.1038/sdata.2016.18.

Wright, H. 2020 *WG1 Exploratory Workshop Results!.* https://www.seadda.eu/?p=1417 [Last accessed: 26 September 2022].