



PROJECT MUSE®

Methodological Diversity in the Evaluation of Cultural Heritage Digital Libraries and Archives: An Analysis of Frameworks and Methods / Diversité méthodologique dans l'évaluation des bibliothèques et les archives numériques du patrimoine culturel : Une analyse des cadres et des méthodes



Emily Villanueva, Ali Shiri

Canadian Journal of Information and Library Science, Volume 43, Number 3, January/Janvier 2021, pp. 316-342 (Article)

Published by University of Toronto Press

➔ For additional information about this article

<https://muse.jhu.edu/article/781390>

Methodological Diversity in the Evaluation of Cultural Heritage Digital Libraries and Archives: An Analysis of Frameworks and Methods

Diversité méthodologique dans l'évaluation des bibliothèques et les archives numériques du patrimoine culturel : Une analyse des cadres et des méthodes

Emily Villanueva
University of Alberta
evillanu@ualberta.ca

Ali Shiri
University of Alberta
ashiri@ualberta.ca

Abstract: Digital library evaluation has become increasingly important in information science, yet there has been minimal evaluative work focusing on digital cultural heritage. We report on a comprehensive review of methodologies and frameworks used in the evaluation of cultural heritage digital libraries and archives. Empirical studies are examined using Tefko Saracevic's digital library evaluation framework to identify models, frameworks, and methodologies in the literature and to categorize these past evaluative approaches. Through the classification and critique of evaluative types and trends, we aim to develop a set of recommendations for the future evaluation of cultural heritage digital libraries and archives.

Keywords: cultural heritage, digital libraries, evaluation, methodologies

Résumé : L'évaluation de bibliothèques numériques a gagné beaucoup en importance dans les sciences de l'information, toutefois, il y a eu peu de travail sur l'aspect de l'héritage culturel numérique. Cet article présente un sommaire des méthodologies, plateformes et techniques utilisées pour l'évaluation des bibliothèques et archives d'héritage culturel numérique. Des études empiriques sont examinées en utilisant le cadre d'évaluation de bibliothèques numériques de Tefko Saracevic pour identifier les modèles, cadres et méthodologies dans la littérature et les catégoriser selon leurs approches. Avec la classification et la critique des types d'approches et de tendances, nous établirons des lignes directrices et des recommandations pour les évaluations de bibliothèques d'héritage culturel futures.

Mots-clés : Héritage culturel, Bibliothèques numériques, Évaluations, Méthodologies

Introduction

The heterogeneous materials and multicultural user groups represented in the cultural heritage field pose a unique challenge for both the design and evaluation of information systems (Petras, Stiller, and Gäde 2013). Evaluation of cultural heritage digital libraries (CHDLs) typically falls under the system-centric or user-centric perspective, and these varying approaches bring to light the features and challenges associated with different evaluative techniques (Petras, Stiller, and Gäde 2013). While not specific to the cultural heritage field, Saracevic (2004) identified seven approaches to digital library (DL) evaluation, all addressing different components or goals: systems-centred, human-centred, usability-centred, anthropological, sociological, economic, and ethnographic. Saracevic (2000) is also credited for introducing the five main elements that frame DL evaluations: construct, context, criteria, measures, and methodology. Other relevant evaluation frameworks have been Nicholson's (2004) holistic matrix model, which examines systems and use from an internal and external view, and Fuhr et al.'s (2001, 2007) DELOS evaluation framework, which is centred on three approaches: system evaluation, usefulness evaluation, and usability evaluation. Ultimately, past research has deemed interface usability, system performance, and collection value to be the most agreed upon evaluation criteria (Xie 2008).

In this article, we discuss the results of a systematic review and meta-analysis on the evaluation of cultural heritage digital libraries and archives. To complete this analysis, we first located the existing literature on the evaluation of CHDLs from library and information studies databases and resources and compiled a bibliography of relevant literature. During this process, we ran into a problem that would recur throughout our review—namely, the difficulty in singularly defining CHDLs. For the purposes of this study, a CHDL was defined as an online repository of digital objects related to the cultural heritage of one or more cultural groups. To broaden the scope of this project, evaluations of digital cultural heritage articles involving institutions across the galleries, libraries, archives, and museums (GLAM) sector were included. Second, we examined each article to identify the frameworks, approaches, methodologies, and data-gathering tools that had been used in previous evaluative studies. During this stage, we removed articles that did not include specific information about their frameworks or did not include a CHDL evaluation, bringing the number of included articles from 103 to 59. Many of the articles that were identified discussed evaluation frameworks, but surprisingly few tested the frameworks on a specific CHDL, which limited the study's sample size. The remaining information was tabulated and categorized in a large spreadsheet informed by Saracevic's (2000) five elements that frame DL evaluations. Finally, we identified the specific cultural elements that were being evaluated in these DL studies. While DLs have been an increasingly important area of study in library and information studies, the specificity and sensitivity required when dealing with cultural heritage in the digital space has often been overlooked. To address this issue, this article will demonstrate the ways in which past CHDL evaluations have failed to account for cultural context and community

engagement and discuss the areas that have been focused on instead. We will then provide a series of evaluative guidelines and recommendations for future research on cultural heritage digital libraries and archives, which can also inform the development of future evaluation frameworks.

DL evaluation

DL evaluation studies are intended to ascertain the level of performance or the value of DL systems. There are different ways that DLs can be judged and different priorities or measures of importance. DLs are often evaluated on their effectiveness, efficiency, and usability. Based on our findings, the majority of DL evaluations focus on usability studies, from both a user and expert perspective. Usability can mean a number of different things, but [Nielsen \(1993\)](#) points out that they are typically encompassed within five main attributes: learnability, efficiency, memorability, errors, and satisfaction. Usability studies cover a broad range of topics and contexts, making them amenable to a variety of research methods including interviews, observation, focus groups, web log analysis, usage statistics, usability testing, and surveys ([Xie 2006](#)). The collection of user data then allows researchers to identify and understand user needs, address potential problems, and assess user satisfaction. While this study will show that there is still a clear focus on usability evaluations, CHDL evaluation studies may also examine system content and performance. These system-focused evaluations often focus on system speed and search capabilities, metadata and linked data capabilities, interface design, and content coverage ([Xie 2006](#)). Although these studies can also be from the user perspective, they are more likely to be from the expert perspective than from the usability evaluations, due to the nature of the evaluation methods.

Usability and performance-based evaluations (including interface and system-centred) encompass the vast majority of CHDL evaluations identified in this study; however, this alludes to a false dichotomy between the two contexts. In actuality, this study identified that the vast majority of CHDL evaluations combine aspects of at least two evaluation contexts and even occasionally combine all three. This is due to the overwhelming user focus of these studies, as most studies, even when talking about the library interface or system performance, took usability into account or included user testing of the system. A less common, but still identifiable, type of DL evaluation is an impact study. Impact studies examine the impact of digital library users and communities through a number of components ([Xie 2006](#)). [Xie \(2006\)](#) notes that these studies often incorporate a longitudinal component in order to identify usage and research trends in target communities over time. Impact studies are important in identifying the needs of user communities and solidifying the importance of CHDL in their communities. The trends and usage statistics gathered have implications for collections management in the digital space ([Xie 2006](#)). DL impact studies also lend themselves well to usability studies, as perceived usefulness and ease of use are both considered determinants of DL acceptance, which directly affects the potential impact of a DL ([Xie 2006](#)).

Although it is clear that there has been an increase in research on DL evaluation over time, there has been little discussion of evaluation criteria, especially when dealing with cultural heritage content in CHDLs. Most current research makes use of pre-existing evaluation criteria for traditional libraries or evaluation models made for DLs, but little research has focused on the creation of an evaluation framework specific to cultural heritage institutions (Stiller and Petras 2018; Xie 2006). Marchionini (2000) has suggested that DLs are extensions and augmentations of physical libraries; however, it may be useful to look to other frameworks surrounding digital technologies for evaluation criteria, including storage capacity, cost per operation, and response time. In many ways, DLs straddle the traditional and the contemporary, and the inclusion of evaluative criteria that prioritizes both of these sets of values is critical to the success of digital and traditional library integration.

Ultimately, as a marriage between the two, Saracevic (2000) identified a set of criteria specifically for DLs: traditional library criteria—collection (purpose, scope, authority, coverage, currency, audience, cost, format, treatment, and preservation), information (accuracy, appropriateness, links, representation, uniqueness, comparability, and presentation), use (accessibility, availability, searchability, and usability), and standards; traditional information retrieval criteria—relevance (precision and recall), satisfaction, and success; and traditional human-computer interaction/interface criteria—usability, functionality, efforts, task appropriateness, and failures. As stated previously, a majority of past DL evaluations in the cultural heritage sector have been usability studies, although other studies have examined collections, systems, and impact. However, these evaluation criteria have focused on DLs as a general category and have failed to account for the unique position of CHDLs in the preservation of cultural heritage resources.

In this study, we used Saracevic's (2000) evaluation framework, which our analysis demonstrated to be one of the more widely used frameworks in evaluation research (Stiller and Petras 2018). The elements incorporated into this framework have been adapted in other DL evaluation frameworks, including DELOS (Candela et al. 2007) and MEDaL (Xie and Matusiak 2016). While these evaluation frameworks are the most relevant for our study, different approaches have been taken for other DL projects. These include the Interaction Triptych Evaluation Model, which defines users, systems, and content as the most important evaluation components; Tsakonas and Papatheodorou's (2011) DiLEO DL evaluation ontology, a research effort to model components from different frameworks in order to guide the understanding of the DL evaluation process; Gonçalves et al.'s (2004) 5S model, which evaluates DLs on the basis of five main components (streams, structures, spaces, scenarios, and societies); and Blandford et al.'s (2008) PRET A Rapporteur framework, which supports the design of evaluation user studies (Stiller and Petras 2018). Saracevic's (2000) evaluation framework introduces five elements that frame the evaluation of DLs: construct, context, criteria, measures, and methodology. This framework was chosen because there were no widely accepted, identifiable frameworks created specifically for CHDLs.

Each element represents a component of our study's evaluation and is described in the following way:

1. Construct for evaluation: what is there to evaluate; what is encompassed by a DL; and what elements are involved in the evaluation?
2. Context of evaluation: select a goal, framework, or level of evaluation: what is the level of evaluation and what is critical for the selected level?
3. Criteria reflecting performance as related to selected objectives: what parameters of performance to concentrate on and what dimension or characteristic to evaluate?
4. Measures reflecting selected criteria to record the performance: what specific measures to use for a given criterion?
5. Methodology for doing evaluation; what measuring instruments to use and what procedures to use for data collection and analysis?

The analysis in this article closely follows the elements presented by [Saracevic \(2000\)](#), focusing on evaluation studies and use cases of CHDLs. This article will thus concentrate on the frameworks and criteria used in the evaluation of CHDLs as well as what we learned from the outcomes of these studies and recommendations for future evaluation frameworks better suited for cultural heritage online environments.

Data gathering and analysis methods

For this analysis, we gathered relevant studies through a systematic search of works related to the evaluation of CHDLs. Searches were conducted on the Directory of Open Access Journals, the Association for Computing Machinery's (ACM) Digital Library, Scopus, Library and Information Science Source, and the University of Alberta's general article search feature through EBSCO and were not limited by time period or geographic area. Terms used in these searches included cultural heritage, DLs, digital archives, evaluation (user and system), methodologies, methods, approaches, and frameworks. A resulting list of 103 articles were collected before being reviewed for relevance. The collected articles were primarily limited by language, including only English and Spanish articles. Relevance was determined through further examination of the content in the collected articles. Articles were removed if DL evaluation was discussed, but no evaluation was conducted. Additionally, articles that described already existing evaluation frameworks but did not include a use case scenario of the framework were excluded. Fifty-nine articles remained in the study following this review process, indicating that these studies covered a specific CHDL evaluation study. Following this, each of the remaining articles was analysed according to Saracevic's five evaluation elements (construct, context, criteria, measures, and methodology), with the relevant material extracted and placed on a spreadsheet. For this study, the construct category resulted in a single group (cultural heritage institutions), as our focus was primarily on framework and criteria, not the specific institution involved in the evaluation. This conclusion was decided upon

through our grounded theory approach, defined by [Glaser and Strauss \(1967, 1\)](#) as ‘the discovery of theory from data—systematically obtained and analysed in social research.’ This qualitative method informed our analysis and the interpretation of the collected CHDL data. While the construct did not end up being central to this study, it is important to note that Europeana and CULTURA (CULTivating Understanding and Research through Adaptivity) were the most frequently evaluated projects, which can limit the large-scale applicability of the results of our work.¹

Analysis and discussion

In this section, we present a list of the evaluation frameworks defined and included in the evaluation studies identified for this project as well as the results of assessing the evaluations with Saracevic’s framework and further categorizing the results in the tables that follow.

Frameworks

The frameworks identified in this study cover a range of GLAM institutions, including libraries, archives, and museums. They differ in a number of ways, including priority (system, user, interface, impact), medium (print, audio, video), and focus (technology, usability). [Table 1](#) presents a compiled list of frameworks described in the aforementioned works.

Table 1. Specific frameworks or methods of evaluation that were applied in the included evaluation studies

Evaluation frameworks
Marchionini’s Digital Library Evaluation Frameworks (Marchionini 2000 ; Suire et al. 2016)
Interaction Triptych Evaluation Model (Agosti, Orio, and Ponchia 2018 ; Steiner et al. 2014)
TIME Evaluation Framework (Task, Information Model, Manipulation of Materials, Ergonomics of Visual Displays) (Dobрева and Chowdhury 2010)
Information Interaction Framework (Stiller 2014)
Museum’s Sites Evaluation Framework (Pallas and Economides 2008)
Technology Acceptance Model (Jeng 2008)
Tobacco Bag Stringing Ontology Model Evaluation (Pattulli 2011)
Culturally Aware, Multi-Method, and Multidisciplinary User Evaluation Framework (Shiri and Stobbs 2018)
Toolkit for the Impact of Digitised Scholarly Resources (Punzalan, Marsh, and Cools 2017)
Archival Metrics Toolkit (Punzalan, Marsh, and Cools 2017)
TREC Video Retrieval Evaluation (Oomen et al. 2013)

Context

As the constructs for this study have all been simplified into the large family of “cultural heritage institutions,” the first evaluation element of [Saracevic’s \(2000\)](#) model is context. The construct was kept intentionally broad as the focus of our research was on the specific ways in which CHDLs were evaluated and not on

the CHDLs themselves (although the CHDLs can be seen in [Appendix 1](#)). Following this, context looks at the perspective used in the evaluation, whether it is user centred, interface centred, or system centred. The results indicate that user-centred evaluations are the most common, which differs from [Saracevic's \(2004\)](#) assertion that system-centred contexts are the most prevalent in DL evaluations. Whether this is simply a difference in categorization or if CHDLs really differ from DLs in their evaluation priorities remains unknown. The majority of user-centred evaluations targeted individual user experiences through direct user involvement in the process, whereas interface and system-centred designs frequently combined the user experience with interface usability or system search efficiency. [Table 2](#) summarizes the number of evaluations per perspective for this study.

Table 2. Number of studies categorized by their context (studies can include more than one focus)

Type of study	Cultural heritage institutions
User centred	52
Interface centred	26
System centred	27
Number of studies	59

Methods

The methods and methodologies identified in this study were numerous and varied, but they have been categorized into four main groups, adapted from [Stiller and Petras's \(2018\)](#) study on Europeana's evaluations. To understand how best to address future CHDL evaluations, it is critical to understand what approaches have been used in the past. While a majority of the evaluations were criteria based, many of the criteria were custom designed for individual studies. Additionally, there is overlap between the criteria-based and usability studies, as many of them used usability as a criterion or emphasized usability in their studies (see [Table 3](#)).

Table 3. Evaluation methods used and the number of studies that applied them

Method	Description	Number of studies
Criteria-based study	Certain criteria were determined to assess a service or algorithm	32
Log file analysis	Evaluation uses an automatically created log file of user interactions	7
Usability study	Evaluation incorporates several methods to assess usability of a service but does not provide specific criteria	19
Impact study	Study uses an expert assessment of the overall value of a service within one or more specific areas	6

Criteria

Many of the studies included in this analysis did not provide a thorough list of the criteria used in their given study. However, by using [Stiller and Petras's \(2018\)](#) criteria outline, we have categorized the criteria in the studies based on the language used around the evaluation and its outcomes. We have added one additional criterion to the list, which is cultural content. This can be seen as an extension of the coverage criterion and is relevant in cases where the CHDL has made a point about its focus on cultural content directly relevant to its target population or institution. The categorization that we have created provides a general overview of the priorities of DL evaluations regardless of context, which is helpful for the implementation of future DL evaluation frameworks (see [Table 4](#)).

Table 4. Criteria applied in evaluation studies and the number of studies that applied them (studies can apply more than one criterion)

Criteria	Description	Number of studies
Accessibility	Covers the forms of access used to navigate content	25
Coverage	Focuses on collection aspects, like size, media type, and linguistic, thematic, or geographic coverage	17
Data quality	Describes all criteria used to identify the quality of data and metadata	26
Error rate	Covers the quantification of errors of workflows or in data quality	3
Impact criteria	Covers all criteria that express change on users or society	6
Performance evaluation	Measures how easy workflows and algorithms are executed (mainly system focused)	12
Usability	Covers established criteria in usability research such as efficiency and effectiveness as well as ease of use and related criteria	34
Usage statistics and patterns	Covers user behaviour such as paths through the system and items clicked or viewed	6
User satisfaction	Subjective criteria of users' perception of the system	23
Cultural content	Extension of coverage criteria that specifically focuses on cultural content	14

Cultural elements

As this study is focused on CHDLs, we have included an additional table for the cultural content criterion. [Table 5](#) specifies the articles that focus on cultural heritage as well as the cultural element that they prioritize. Additionally, all of the information gathered during this review has been aggregated and is presented alphabetically by author surname in [Appendix 1](#).

[Saracevic's \(2000\)](#) framework has allowed us to analyse a number of CHDL evaluations as well as to identify where gaps may occur that should be addressed. Through our analysis of 59 evaluation studies, we have identified a number of gaps in the identified CHDL evaluations, some of which are unique to cultural

Table 5. Specific cultural elements that were focused on, or evaluated in, the evaluation studies

	Cultural elements
Bow (2019)	Cultural coverage; ethnography
Farnel et al. (2017)	Culturally relevant metadata
Feinberg (2013)	Cultural interplay and authorial voices in metadata
Galani and Kidd (2019)	Cultural heritage “experience”
Ibrahim and Ali 2018	Cultural learning
Komlodi, Caidi, and Wheeler (2004)	Cross-cultural usability
Liew (2005)	Cross-cultural design and usability
Liew and Chowdhury (2016)	Advocacy; community engagement; equity; cultural sensitivity; literacy
Núñez and Repiso (2019)	Cultural heritage interoperability
Shiri (2018)	Culturally driven methodological framework
Shiri and Stobbs (2018)	Cultural context; community engagement
Szabo, Lacedelli, and Pompanin (2017)	Cultural coverage; co-creation
Van den Akker et al. (2013)	Cultural and historical metadata
Yelmi, Kuşcu, and Yantaç (2016)	Cultural soundscape

heritage institutions and some of which could be applicable to both DLs and CHDLs:

- The identified studies strongly lean towards well-known CHDLs, such as CULTURA and Europeana, which provides a limited view of CHDL evaluation since it may be missing out on important conversations by privileging certain institutions.
- There is a strong focus on user-centred perspectives. While the user is a critical component of CHDLs, especially in an evaluative capacity, it once again provides a lack of other perspectives.
- Many evaluation studies suffer from a lack of description in both their methodologies and in their criteria, making it difficult to fully analyse, re-use, or replicate studies.
- In terms of cultural heritage, many CHDL evaluations fail to mention or account for culture in any capacity. Few studies have talked about community collaboration for a specific cultural group or the creation of a culturally driven framework.

Although it is clear that there have been large-scale evaluations of CHDLs, they are often evaluated in the same way as DLs, which is void of the cultural element. Not every CHDL study may need to have a culturally driven evaluation framework; however, the importance of culture and community should not be entirely ignored as it is in most of these studies. This is particularly important when speaking of DLs made for historically disenfranchised communities, which have often been excluded from academic narratives about their own culture. In these cases, community engagement and cultural awareness are critical not only for the library’s development but also for any future evaluations, whether that be an analysis of community satisfaction, the collection size of cultural material, or the presence of items in local languages and dialects. No evaluative framework for cultural

content in a digital environment has been determined as of this study, and, in the following sections, we make some recommendations for possible inclusions.

Guidelines and recommendations

It has been argued that simplicity is a fundamental principle of building search interfaces (Aula and Käki 2005; Buttenfield 1999). This is why interface analysis from the user perspective is one of the most prevalent methods of CHDL evaluation identified in this study as well as one of the primary focuses of the human-computer interaction field. The guidelines and recommendations reported here are of great importance to those involved in the creation and study of CHDLs since they are intended to offer alternatives to current CHDL iterations as well as to reveal problems with prior evaluative work and considerations for future evaluations and evaluation frameworks (Gaona-García, Martin-Moncunill, and Montenegro-Marin 2017). While the included guidelines centre on cultural knowledge management as opposed to culturally appropriate evaluation methods, our belief is that these two concepts are intertwined. Knowledge management and evaluation are reciprocal processes, and if more CHDLs begin to work closely with relevant populations during testing and development stages, there will be a greater drive to account for these contexts and populations during later evaluations. CHDLs should not just inform communities but be informed by them. Our guidelines for more effective knowledge management in cultural heritage repositories, inspired by DL interface guidelines created by Gaona-García, Martin-Moncunill, and Montenegro-Marin (2017), are as follows:

- include the cultural community in CHDL creation, including working with the community to identify what is important to them in a cultural heritage setting and implementing those findings where possible (that is, culturally appropriate and sensitive metadata, coverage of specific concepts and events, and locally tested interface);
- define methods for linking data objects through cultural topics or knowledge areas;
- define an enriched culturally relevant language within the appropriate knowledge representation scheme to facilitate mapping processes to external ontologies hosted in other cultural heritage repositories;
- Use linked data processes to enable interoperability between heterogeneous CHDL repositories;
- provide an easily navigable interface to allow users to gain an overview of the cultural areas of expertise or interest represented; and
- integrate metadata with site navigation to make thematic site exploration easier.

As the alteration of CHDL management practices will not necessarily affect CHDL evaluation frameworks, we also propose a number of recommendations to consider before undertaking future evaluations of CHDLs and other cultural heritage repositories. These recommendations are in response to the current lack of culturally driven CHDL evaluation and include the following:

- include the cultural community in CHDL testing and evaluation;
- identify the number of relevant digital resources retrieved in each query process;
- examine the relevance of metadata attached to digital resources retrieved in each query process;
- track the usage of digital resources according to their use case—academic, scientific, public, and targeted cultural group;
- identify the number of associated digital resources for a cultural topic or thematic area; and
- identify the number of multilingual digital resources as well as the system’s multilingual search capabilities.

Conclusion

This analysis and critique of 59 CHDL evaluation studies provides only a brief overview of the variety of evaluative frameworks, methods, and criteria that have been presented and adapted over the years. While all of the studies included in this article have been helpful in the formation of this overview, they have also shed light on the shortcomings of the current field of evaluative literature in CHDL. Namely, the lack of a cultural component in many CHDL evaluations seems like a vast oversight by many. The adoption of general DL evaluation frameworks, much like the one used in this study, has been part of the problem, as these pre-formulated frameworks were not intended for a cultural heritage audience. This does not mean that these formative frameworks must be scrapped but, rather, that adjustments and adaptations need to be made. For example, what is the involvement of the cultural community in the creation and testing of the CHDL? Is that information being reported? Are cultural communities being included in the evaluative process? Are their needs being met, whether that is through coverage, interface design, or metadata vocabulary? Ultimately, we are strongly recommending the inclusion of specific cultural components to future CHDL evaluations in order to solidify their distinction from other DLs.

Note

- 1 Europeana, <https://www.europeana.eu/en> (accessed December 15, 2020); CULTivating Understanding and Research through Adaptivity, <http://www.cultura-strep.eu/> (accessed December 15, 2020).

References

- Abdallah, Samer, Emmanouil Benetos, Nicholas Gold, Steven Hargreaves, Tillman Weyde, and Daniel Wolff. 2017. “The Digital Music Lab: A Big Data Infrastructure for Digital Musicology.” *ACM Journal on Computing and Cultural Heritage* 10 (1): 1–21. <https://doi.org/10.1145/2983918>.
- Agosti, Maristella, Marta Manfioletti, Nicola Orio, and Chiara Ponchia. 2013. “Evaluating the Deployment of a Collection of Images in the CULTURA Environment.” In *Research and Advanced Technology for Digital Libraries. TPDL 2013: Lecture Notes in Computer Science*, ed. T. Aalberg, C. Papatheodorou, M. Dobрева, G.

- Tsakonas, and C.J. Farrugia, 180–91. Vol. 8092. Berlin: Springer. https://doi.org/10.1007/978-3-642-40501-3_18.
- Agosti, Maristella, Nicola Orio, and Chiara Ponchia. 2014. "Guided Tours across a Collection of Historical Digital Images." In *Proceedings of the 3rd AIUCD Annual Conference on Humanities and Their Methods in the Digital Ecosystem (AIUCD '14)*, ed. F. Tomasi, R. Rosselli Del Turco, and A.M. Tammaro, 7:1–7:6. New York: Association for Computing Machinery. <https://doi.org/10.1145/2802612.2802632>.
- . 2018. "Promoting User Engagement with Digital Cultural Heritage Collections." *International Journal on Digital Libraries* 19: 353–66. <https://doi.org/10.1007/s00799-018-0245-y>.
- Aletras, Nikolaos, Mark Stevenson, and Paul Clough. 2012. "Computing Similarity between Items in a Digital Library of Cultural Heritage." *Journal on Computing and Cultural Heritage* 5 (4): 16:1–16:19. <https://doi.org/10.1145/2399180.2399184>.
- Anderson, Ian G. 2007. "Pure Dead Brilliant?: Evaluating the Glasgow Story Digitisation Project." *Program: Electronic Library and Information Systems* 41 (4): 365–85. <https://doi.org/10.1108/00330330710831585>.
- Aula, Anne, and Mika Käki. 2005. "Less Is More in Web Search Interfaces for Older Adults." *First Monday* 10 (7). <https://doi.org/10.5210/fm.v10i7.1254>.
- Bailey, Eoin, Seamus Lawless, Cormac Hampson, Alexander O'Connor, Mark Sweetnam, Owen Conlan, and Vincent Wade. 2012. "CULTURA: Supporting Enhanced Exploration of Cultural Archives Through Personalisation." In *Proceedings of the 2nd International Conference on Humanities, Society and Culture*. Vol. 51. Hong Kong: Centre for the Humanities and Social Change. <http://www.tara.tcd.ie/handle/2262/67053>.
- Blandford, Ann, Anne Adams, Simon Attfield, George Buchanan, Jeremy Gow, Stephann Makri, Jon Rimmer, and Claire Warwick. 2008. "The PRET a Rapporteur Framework: Evaluating Digital Libraries from the Perspective of Information Work." *Information Processing and Management* 44 (1): 4–21. <https://doi.org/10.1016/j.ipm.2007.01.021>.
- Bonacini, Elisa. 2019. "Engaging Participative Communities in Cultural Heritage: Using Digital Storytelling in Sicily (Italy)." *International Information and Library Review* 51 (1): 42–50. <https://doi.org/10.1080/10572317.2019.1568786>.
- Borgman, Christine L., Gregory H. Leazer, Anne J. Gilliland-Swetland, and Rich Gazan. 2001. "Iterative Design and Evaluation of a Geographic Digital Library for University Students: A Case Study of the Alexandria Digital Earth Prototype (ADEPT)." In *Research and Advanced Technology for Digital Libraries. ECDL 2001: Lecture Notes in Computer Science*, ed. P. Constantopoulos and I. T. Sølvberg, 390–401. Vol. 2163. Berlin: Springer. https://doi.org/10.1007/3-540-44796-2_33.
- Bow, Catherine. 2019. "Diverse Socio-Technical Aspects of a Digital Archive of Aboriginal Languages." *Archives and Manuscripts* 47 (1): 94–112. <https://doi.org/10.1080/01576895.2019.1570282>.
- Buttenfield, Barbara. 1999. "Usability Evaluation of Digital Libraries." *Science and Technology Libraries* 17 (3–4): 39–59. https://doi.org/10.1300/j122v17n03_04.
- Candela, Gustavo, Pilar Escobar, and Manuel Marco-Such. 2017. "Semantic Enrichment on Cultural Heritage Collections: A Case Study Using Geographic Information." In *Proceedings of the 2nd International Conference on Digital Access to Textual Cultural Heritage (DATeCH2017)*, 169–74. New York: Association for Computing Machinery. <https://doi.org/10.1145/3078081.3078090>.

- Candela, Leonardo, Donatella Castelli, Nicola Ferro, Yannis E. Ioannidis, Georgia Koutrika, Carlo Meghini, Pasquale Pagano, Seamus Ross, and Dagobert Soergel. 2007. *The DELOS Digital Library Reference Model: Foundations for Digital Libraries, Version 0.96*. Pisa, Italy: DELOS Network of Excellence on Digital Libraries.
- Crane, Gregory, and Clifford Wulfman. 2003. "Towards a Cultural Heritage Digital Library." In *Proceedings of the 3rd ACM/IEEE-CS Joint Conference on Digital Libraries*, 75–86. Washington, DC: Institute of Electrical and Electronics Engineers Computer Society. <https://doi.org/10.1109/JCDL.2003.1204845>.
- Dobрева, Milena, and Sudatta Chowdhury. 2010. "A User-Centric Evaluation of the Europeana Digital Library." In *Proceedings of the Role of Digital Libraries in a Time of Global Change, and 12th International Conference on Asia-Pacific Digital Libraries*, ed. Gobinda Chowdhury, Chris Koo, and Jane Hunter, 148–57. Berlin: Springer. https://doi.org/10.1007/978-3-642-13654-2_19.
- Dorward, Jim, Derek Reinke, and Mimi Recker. 2002. "An Evaluation Model for a Digital Library Services Tool." In *Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries*, 322–23. New York: Association for Computing Machinery. <https://doi.org/10.1145/544220.544298>.
- Farnel, Sharon, Ali Shiri, Sandra Campbell, Cathy Cockney, Dinesh Rathi, and Robyn Stobbs. 2017. "A Community-Driven Metadata Framework for Describing Cultural Resources: The Digital Library North Project." *Cataloging & Classification Quarterly* 55 (5): 289–306. <https://doi.org/10.1080/01639374.2017.1312723>.
- Feinberg, Melanie. 2013. "Comparative Appraisal: Systematic Assessment of Expressive Qualities." In *Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries*, 115–24. New York: Association for Computing Machinery. <https://doi.org/10.1145/2467696.2467700>.
- Fenlon, Katrina. 2013. "Exploring Evaluative Methods for Large-Scale Local History." In *Proceedings of the 76th ASIS&T Annual Meeting: Beyond the Cloud: Rethinking Information Boundaries*, 153:1–153:4. Silver Springs, MD: American Society for Information Science and Technology. <https://doi.org/10.1002/meet.14505001152>.
- Freire, Nuno, Borbinha José, and Pável Calado. 2012. "An Approach for Named Entity Recognition in Poorly Structured Data." In *The Semantic Web: Research and Applications. ESWC 2012: Lecture Notes in Computer Science*, ed. E. Simperl, P. Cimiano, A. Polleres, O. Corcho, and V. Presutti, 718–32. Vol. 7295. Berlin: Springer. https://doi.org/10.1007/978-3-642-30284-8_55.
- Fuhr, Norbert, Preben Hansen, Michael Mabe, Andras Micsik, and Ingeborg Sølberg. 2001. "Digital Libraries: A Generic Classification and Evaluation Scheme." In *Research and Advanced Technology for Digital Libraries. ECDL 2001: Lecture Notes in Computer Science*, ed. P. Constantopoulos and I.T. Sølberg, 187–99. Vol. 2163. Berlin: Springer. https://doi.org/10.1007/3-540-44796-2_17.
- Fuhr, Norbert, Giannis Tsakonas, Trond Aalberg, Maristella Agosti, Preben Hansen, Sarantos Kapidakis, Claus-Peter Klas, László Kovács, Monica Landoni, Andrés Micsik, Christos Papatheodorou, Carol Peters, and Ingeborg Sølberg. 2007. "Evaluation of Digital Libraries." *International Journal of Digital Libraries* 8 (1): 21–38. <https://doi.org/10.1007/s00799-007-0011-z>.
- Galani, Areti, and Jenny Kidd. 2019. "Evaluating Digital Cultural Heritage 'in the Wild': The Case for Reflexivity." *ACM Journal on Computing and Cultural Heritage* 12 (1): 5:1–5:15. <https://doi.org/10.1145/3287272>.
- Gaona-García, Paul A., David Martin-Moncunill, and Carlos Montenegro-Marin. 2017. "Trends and Challenges of Visual Search Interfaces in Digital Libraries and

- Repositories." *The Electronic Library* 35 (1): 69–98. <https://doi.org/10.1108/el-03-2015-0046>.
- Glaser, Barney G., and Anselm L. Strauss. 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Aldine de Gruyter.
- Gonçalves, Marcos André, Edward A. Fox, Layne T. Watson, and Neill A. Kipp. 2004. "Streams, Structures, Spaces, Scenarios, Societies (5S): A Formal Model for Digital Libraries." *ACM Transactions on Information Systems* 22 (2): 270–312. <https://doi.org/10.1145/984321.984325>.
- Goodale, Paula. 2016. "Individual Differences in Exploration and Content Curation Activities Within a Cultural Heritage Digital Library." In *Proceedings of 1st International Workshop on Accessing Cultural Heritage at Scale*, ed. P. Clough, P. Goodale, M. Agosti, and S. Lawless. Vol. 1611. New York: Association for Computing Machinery. Accessed December 15, 2020. <http://ceur-ws.org/Vol-1611/paper8.pdf>.
- Goodale, Paula, Paul Clough, Mark Hall, Mark Stevenson, Kate Fernie, and Jillian Griffiths. 2014. "Supporting Information Access and Sensemaking in Digital Cultural Heritage Environments." In *Theory and Practice of Digital Libraries 2013 Selected Workshops: Communications in Computer and Information Science*, ed. Ł. Bolikowski, V. Casarosa, P. Goodale, N. Houssos, P. Manghi, and J. Schirrwagen, 143–54. Vol. 416. Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-08425-1_13.
- Gordea, Sergiu. 2014. "An Image Similarity Search for the European Digital Library and Beyond." In *Theory and Practice of Digital Libraries 2013 Selected Workshops: Communications in Computer and Information Science*, ed. Ł. Bolikowski, V. Casarosa, P. Goodale, N. Houssos, P. Manghi, and J. Schirrwagen, 190–96. Vol. 416. Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-08425-1_21.
- Hall, Mark M., Samuel Fernando, Mark Stevenson, Paul D. Clough, Aitor Soroa, and Eneko Agirre. 2014. "Evaluating Hierarchical Organisation Structures for Exploring Digital Libraries." *Information Retrieval* 17 (4): 351–79. <https://doi.org/10.1007/s10791-014-9242-y>.
- Hampson, Cormac, Séamus Lawless, Eoin Bailey, Owen Conlan, Christina M. Steiner, and Eva-Catherine Hillemann. 2014. "Metadata-Enhanced Exploration of Digital Cultural Collections." *International Journal of Metadata, Semantics and Ontologies* 9 (2): 155–67. <https://doi.org/10.1504/ijmso.2014.060342>.
- Harris, Martyn, Mark Levene, Dell Zhang, and Dan Levene. 2018. "Finding Parallel Passages in Cultural Heritage Archives." *ACM Journal on Computing and Cultural Heritage* 11 (3): 15:1–15:24. <https://doi.org/10.1145/3195727>.
- Harsányi, Zoltán, Viera Rozinajová, and Nadeža Andrejčíková. 2012. "Identifying Semantic Relationships in Digital Libraries of Cultural Heritage." In *Progress in Cultural Heritage Preservation. EuroMed 2012: Lecture Notes in Computer Science*, ed. M. Ioannides, D. Fritsch, J. Leissner, R. Davies, F. Remondino, and R. Caffo, 738–45. Vol. 7616. Berlin: Springer. https://doi.org/10.1007/978-3-642-34234-9_78.
- Hill, Linda L., Larry Carver, Mary Larsgaard, Ron Dolin, Terence R. Smith, James Frew, and Mary-Anna Rae. 2000. "Alexandria Digital Library: User Evaluation Studies and System Design." *Journal of the American Society for Information Science* 51 (3): 246–59. [https://doi.org/10.1002/\(sici\)1097-4571\(2000\)51:3%3C246::aid-asi4%3E3.0.co;2-6](https://doi.org/10.1002/(sici)1097-4571(2000)51:3%3C246::aid-asi4%3E3.0.co;2-6).
- Hu, Xiao, Eric M. Ho, and Chen Qiao. 2017. "Digitizing Dunhuang Cultural Heritage: A User Evaluation of Mogao Cave Panorama Digital Library." *Journal of Data and Information Science* 2 (3): 49–67. <https://doi.org/10.1515/jdis-2017-0014>.

- Hu, Xiao, Jeremy Ng, and Shengping Xia. 2018. "User-Centered Evaluation of Metadata Schema for Nonmovable Cultural Heritage: Murals and Stone Cave Temples." *Journal of the Association for Information Science and Technology* 69 (12): 1476–87. <https://doi.org/10.1002/asi.24065>.
- Hug, Charlotte, and Cesar Gonzalez-Perez. 2012. "Qualitative Evaluation of Cultural Heritage Information Modeling Techniques." *Journal on Computing and Cultural Heritage* 5 (2): 8:1–8:20. <https://doi.org/10.1145/2307723.2307727>.
- Ibrahim, Nazrita, and Nazlena M. Ali. 2018. "A Conceptual Framework for Designing Virtual Heritage Environment for Cultural Learning." *ACM Journal on Computing and Cultural Heritage* 11 (2): 11:1–11:27. <https://doi.org/10.1145/3117801>.
- Jeng, Judy. 2008. "Evaluation of the New Jersey Digital Highway." *Information Technology and Libraries* 27 (4): 17–24. <https://doi.org/10.6017/ital.v27i4.3238>.
- Komlodi Anita, Nadia Caidi, and Kristin Wheeler. 2004. "Cross-Cultural Usability of Digital Libraries." In *Digital Libraries: International Collaboration and Cross-Fertilization. ICADL 2004: Lecture Notes in Computer Science*, ed. Z. Chen, H. Chen, Q. Miao, Y. Fu, E. Fox, E. Lim, 584–93. Vol. 3334. Berlin: Springer. https://doi.org/10.1007/978-3-540-30544-6_67.
- Liew, Chern L. 2005. "Cross-Cultural Design and Usability of a Digital Library Supporting Access to Maori Cultural Heritage Resources." In *Design and Usability of Digital Libraries: Case Studies in the Asia Pacific*, ed. Y. Theng and S. Foo, 285–97. Hershey, PA: IGI Global. <https://doi.org/10.4018/978-1-59140-441-5.ch016>.
- Liew, Chern L., and Gobinda G. Chowdhury. 2016. "Digital Cultural Heritage and Social Sustainability." *Journal of Community Informatics* 12 (3): 173–96. Accessed December 15, 2020. <http://ci-journal.net/index.php/ciej/article/view/1278/1238>.
- Marchionini, Gary. 2000. "Evaluating Digital Libraries: A Longitudinal and Multifaceted View." *Library Trends* 49 (2): 304–33. Accessed December 15, 2020. https://www.ideals.illinois.edu/bitstream/handle/2142/8336/librarytrendsv49i2g_opt.pdf?sequence=1.
- Marketakis, Yannis, Nikos Minadakis, Haridimos Kondylakis, Konstantina Konsolaki, Georgios Samaritakis, Maria Theodoridou, Giorgos Flouris, and Martin Doerr. 2017. "X3ML Mapping Framework for Information Integration in Cultural Heritage and Beyond." *International Journal on Digital Libraries* 18 (4): 301–19. <https://doi.org/10.1007/s00799-016-0179-1>.
- Matthews, Paul, and Judith Aston. 2012. "Interactive Multimedia Ethnography: Archiving Workflow, Interface Aesthetics and Metadata." *Journal on Computing and Cultural Heritage* 5 (4): 14:1–14:13. <https://doi.org/10.1145/2399180.2399182>.
- Melucci, Massimo, and Nicola Orio. 2004. "Combining Melody Processing and Information Retrieval Techniques: Methodology, Evaluation, and System Implementation." *Journal of the American Society for Information Science* 55 (12): 1058–66. <https://doi.org/10.1002/asi.20058>.
- Nicholson, Scott. 2004. "A Conceptual Framework for the Holistic Measurement and Cumulative Evaluation of Library Services." *Journal of Documentation* 60 (2): 164–82. <https://doi.org/10.1108/00220410410522043>.
- Nielsen, Jakob. 1993. *Usability Engineering*. Cambridge, MA: AP Professional.
- Núñez, María E., and Virginia Ortiz Repiso. 2019. "Colecciones Digitales del Patrimonio Cultural de Canarias: Situación Actual y Propuesta de Actuación." *Revista General de Información y Documentación* 29 (1): 141–65. <https://doi.org/10.5209/rgid.64552>.
- Oomen, Johan, Paul Over, Wessel Kraaij, and Alan F. Smeaton. 2013. "Symbiosis between the TRECVID Benchmark and Video Libraries at the Netherlands Institute for Sound and Vision." *International Journal of Digital Libraries* 13 (2): 91–104. <https://doi.org/10.1007/s00799-012-0102-3>.

- Pallas, John, and Anastasios A. Economides. 2008. "Evaluation of Art Museums' Web Sites Worldwide." *Information Services and Use* 28 (1): 45–57. <https://doi.org/10.3233/isu-2008-0554>.
- Pattulli, M. Cristina. 2011. "Modeling a Domain Ontology for Cultural Heritage Resources: A User-Centered Approach." *Journal of the American Society for Information Science and Technology* 62 (2): 314–42. <https://doi.org/10.1002/asi.21453>.
- Petras, Vivien, Juliane Stiller, and Maria Gäde. 2013. "Building for Success? Evaluating Digital Libraries in the Cultural Heritage Domain." In *Recent Developments in the Design, Construction, and Evaluation of Digital Libraries: Case Studies*, ed. C. Cool and K. B. Ng, 141–63. Hershey, PA: IGI Global. <https://doi.org/10.4018/978-1-4666-2991-2.ch009>.
- Punzalan, Ricardo L., Diana E. Marsh, and Kyla Cools. 2017. "Beyond Clicks, Likes, and Downloads: Identifying Meaningful Impacts for Digitized Ethnographic Archives." *Archivaria* 84: 61–102. Accessed December 15, 2020. <https://archivaria.ca/index.php/archivaria/article/view/13614>.
- Saracevic, Tefko. 2000. "Digital Library Evaluation: Toward an Evolution of Concepts." *Library Trends* 49 (3): 350–69. Accessed December 15, 2020. https://www.ideals.illinois.edu/bitstream/handle/2142/8343/librarytrendsv49i2i_opt.pdf.
- . 2004. "Evaluation of Digital Libraries: An Overview." Paper presented at the DELOS Workshop on the Evaluation of Digital Libraries, Padua, Italy. http://dlib.ionio.gr/wp7/WS2004_Saracevic.pdf.
- Shiri, Ali. 2018. "Methodological Considerations in Developing Cultural Heritage Digital Libraries: A Community-Driven Framework." In *Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries*, 383–4. New York: Association for Computing Machinery. <https://doi.org/10.1145/3197026.3203893>.
- Shiri, Ali, and Robyn Stobbs. 2018. "Community-Driven User Evaluation of the Inuvialuit Cultural Heritage Digital Library." *Proceedings of the Association for Information Science and Technology* 55 (1): 440–9. <https://doi.org/10.1002/pr2.2018.14505501048>.
- Skevakis, Giannis, Konstantinos Makris, Varvara Kalokyri, Polyxeni Arapi, and Stavros Christodoulakis. 2014. "Metadata Management, Interoperability and Linked Data Publishing Support for Natural History Museums." *International Journal on Digital Libraries* 14 (3): 127–40. <https://doi.org/10.1007/s00799-014-0114-2>.
- Steiner, Christina M., Maristella Agosti, Mark S. Sweetnam, Eva-C. Hillemann, Nicola Orio, Chiara Ponchia, Cormac Hampson, Gary Munnely, Alexander Nussbaumer, Dietrich Albert, and Owen Conlan. 2014. "Evaluating a Digital Humanities Research Environment: The CULTURA Approach." *International Journal of Digital Libraries* 15 (1): 53–70. <https://doi.org/10.1007/s00799-014-0127-x>.
- Stiller, Juliane. 2014. "From Curation to Collaboration: A Framework for Interactions in Cultural Heritage Information Systems." PhD diss., Humboldt-University. Accessed December 15, 2020. <https://edoc.hu-berlin.de/handle/18452/17596>.
- Stiller, Juliane, Maria Gäde, and Vivien Petras. 2013. "Multilingual Access to Digital Libraries: The Europeana Use Case." *Information: Wissenschaft und Praxis* 64 (2–3): 86–95. <https://doi.org/10.1515/iwp-2013-0014>.
- Stiller, Juliane, and Vivien Petras. 2018. "Learning from Digital Library Evaluations: The Europeana Case." *ABI Technik* 38 (1): 37–45. <https://doi.org/10.1515/abitech-2018-0006>.
- Suire, Cyrille, Axel Jean-Caurant, Vincent Courboulay, Jean-Christophe Burie, and Pascal Estraillier. 2016. "User Activity Characterization in a Cultural Heritage Digital Library System." In *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries*, 257–58. New York: Association for Computing Machinery. <https://doi.org/10.1145/2910896.2925459>.

- Sulé Duesa, Andreu, Assumpció E. Rius, and Jesús Gascon García. 2011. "Evaluación de las Interfaces de Consulta de las Colecciones Digitales Patrimoniales Españolas." *Anales de Documentacion* 14 (2): 1–21. Accessed December 15, 2020. <https://revistas.um.es/analesdoc/article/view/113931>.
- Szabo, Victoria, Stefania Z. Lacedelli, and Giacomo Pompanin. 2017. "From Landscape to Cities: A Participatory Approach to the Creation of Digital Cultural Heritage." *International Information and Library Review* 49 (2): 115–23. <https://doi.org/10.1080/10572317.2017.1314141>.
- Tsakonas, Giannis, and Christos Papatheodorou. 2011. "An Ontological Representation of the Digital Library Evaluation Domain." *Journal of American Society for Information Science and Technology* 62 (8): 1577–93. <https://doi.org/10.1002/asi.21559>.
- Van den Akker, Chiel, Ardjan van Nuland, Lourens van der Meij, Marieke van Erp, Susan Legêne, Lora Aroyo, and Guus Schreiber. 2013. "From Information Delivery to Interpretation Support: Evaluating Cultural Heritage Access on the Web." In *Proceedings of the 5th Annual ACM Web Science Conference*, 431–40. New York: Association for Computing Machinery. <https://doi.org/10.1145/2464464.2464491>.
- van Hooland, Seth, Ruben Verborgh, Max De Wilde, Johannes Hercher, Erik Mannens, and Rik Van de Walle. 2013. "Evaluating the Success of Vocabulary Reconciliation for Cultural Heritage Collections." *Journal of the Association for Information Science and Technology* 64 (3): 464–79. <https://doi.org/10.1002/asi.22763>.
- Vila-Suero, Danila, Mariano Rico, Iuliana Botezan, and Asunción Gómez-Pérez. 2019. "Evaluating the Impact of Semantic Technologies on Bibliographic Systems: A User-Centered and Comparative Approach." *Journal of Web Semantics* 59: 1–17. <https://doi.org/10.2139/ssrn.3351106>.
- Wang Shenghui, Antoine Isaac, Valentine Charles, Rob Koopman, Anthi Agoropoulou, and Titia van der Werf. 2013. "Hierarchical Structuring of Cultural Heritage Objects Within Large Aggregations." In *Research and Advanced Technology for Digital Libraries. TPDL 2013: Lecture Notes in Computer Science*, ed. T. Aalberg, C. Tapatheodorou, M. Dobрева, G. Tsakonas, and C.J. Farrugia, 247–59. Vol. 8092. Berlin: Springer. https://doi.org/10.1007/978-3-642-40501-3_25.
- Xie, Iris. 2006. "Evaluation of Digital Libraries: Criteria and Problems from Users Perspectives." *Library & Information Science Research* 28 (3): 433–52. <https://doi.org/10.1016/j.lisr.2006.06.002>.
- . 2008. "Users' Evaluation of Digital Libraries (DLs): Their Uses, Their Criteria, and Their Assessment." *Information Processing and Management* 44 (3): 1346–73. <https://doi.org/10.1016/j.ipm.2007.10.003>.
- Xie, Iris, and Krystyna K. Matusiak. 2016. *Discover Digital Libraries: Theory and Practice*. Amsterdam: Elsevier.
- Yelmi, Pinar, Hüseyin Kuşcu, and Asım E. Yantaç. 2016. "Towards a Sustainable Crowdsourced Sound Heritage Archive by Public Participation: The Soundsslike Project." In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, 71:1–71:9. Vol. 71. New York: Association for Computing Machinery. <https://doi.org/10.1145/2971485.2971492>.

Appendix 1. Summary of the meta-analysis of the 59 evaluations studied

Authors and year of source	Name	Framework/Approach	Method	Criteria	Perspective	Data type	User centred	Interface	System centred
Abdallah et al. (2017)	British Library, CHARM, Mazurka, I Like Music	Digital Music Library Framework	Criteria-based study	Usability, data quality, performance evaluation	User	Quantitative	x		x
Agosti et al. (2013)	CULTURA	Development of a new adaptive and dynamic environment, IPSA Cultura	Usability study	Usability, accessibility	User	Qualitative	x	x	
Agosti, Orio, and Ponchia (2014)	CULTURA	Personalized embedded narratives/guided collection tours	Usability study	Accessibility, usability, user satisfaction	User	Qualitative	x		
Agosti, Orio, and Ponchia (2018)	CULTURA	Interaction Triptych Model	Usability study	Usability, user satisfaction, performance evaluation	Expert/user	Qualitative	x	x	x
Aletras, Stevenson, and Clough (2012)	Europeana	Knowledge-based and corpus-based approach	Criteria-based study	Accessibility, data quality, performance evaluation	Expert	Quantitative	x		x
Anderson (2007)	The Glasgow Story	Impact assessment	Impact study	Impact criteria, usability, user satisfaction, usage statistics and patterns, accessibility	Expert/user	Both	x		

(continued on next page)

(Continued)

Authors and year of source	Name	Framework/Approach	Method	Criteria	Perspective	Data type	User centred	Interface	System centred
Bailey et al. (2012)	Trinity College Dublin, University of Padua, CULTURA	Experimental approach	Log file analysis	Usage statistics and patterns	Expert/user	Quantitative	x		
Bonacini (2019)	#iziTRAVELSicilia	Participatory community approach	Impact study	Impact criteria, user satisfaction	User	Quantitative	x		
Borgman et al. (2001)	Alexandria Digital Earth ProtoType	Bottom-up approach to DL design	Usability study	Usability, data quality, coverage	Expert/user	Qualitative	x		
Bow (2019)	The Living Archive of Aboriginal Languages	Knowledge infrastructure through a socio-technical lens	Usability study	Usability, user satisfaction, coverage, cultural content, accessibility	Expert/user	Qualitative	x		
Candela, Escobar, and Marco-Such (2017)	Biblioteca Virtual Miguel de Cervantes	Biblioteca Virtual Miguel de Cervantes ontology	Usability study	Accessibility, data quality, usability, user satisfaction	User	Both	x	x	x
Crane and Wulfman (2003)	Perseus digital library, National Science Digital Library	Neglect versus need, new domains versus disciplinary, using versus creating digital collections	Log file analysis	Usage statistics and patterns	Expert	Quantitative	x		x
Dobrea and Chowdhury (2010)	Europeana	TIME Evaluation Framework	Usability study	User satisfaction, coverage, usability	User	Qualitative	x	x	x

(continued on next page)

(Continued)

Authors and year of source	Name	Framework/Approach	Method	Criteria	Perspective	Data type	User centred	Interface	System centred
Dorward, Reinke, and Recker (2002)	SMETE Open Federation Digital Library	Instructional Architect	Usability study	Usability	Expert/user	Qualitative	x	x	
Farnel et al. (2017)	Digital Library North	Community-driven metadata framework	Usability study	Cultural content, coverage, accessibility, usability, data quality	User	Qualitative	x	x	
Feinberg (2013)	Personal digital collection (that is, Pinterest, YouTube playlists)	Comparative appraisal	Criteria-based study	Data quality, cultural content, coverage	User	Both (mainly qualitative)	x	x	
Fenlon (2013)	Digital Public Library of America	Multimodal pilot study	Criteria-based study	Data quality, error rate, performance evaluation	Expert	Qualitative			x
Freire, José, and Calado (2012)	Europeana	Conditional random field models	Criteria-based study	Data quality, error rate, performance evaluation	Expert	Quantitative			x
Galani and Kidd (2019)	With New Eyes I See, Rock Art on Mobile	Multimodality and reflexivity	Impact study	Accessibility, impact criteria, cultural content, user satisfaction	User	Both (mainly qualitative)	x		
Goodale (2016)	Europeana	PATHS system	Criteria-based study	Usability, user satisfaction	Expert/user	Both	x	x	

(continued on next page)

(Continued)

Authors and year of source	Name	Framework/Approach	Method	Criteria	Perspective	Data type	User centred	Interface	System centred
Goodale et al. (2014)	Europeana	PATHS system; Interactive Information Retrieval Evaluation Framework	Log file analysis	Usage statistics and patterns, usability, user satisfaction	Expert/ user	Both	x		x
Gordea (2014)	Europeana	Image similarity search service	Criteria-based study	Performance evaluation, data quality	Expert	Quantitative			x
Hall et al. (2014)	Europeana, Wikipedia	Evaluation of four key activities that hierarchies support	Usability study	Data quality, coverage	User	Both	x	x	
Hampson et al. (2014)	CULTURA, Biodiversity Heritage Library, Bayerische Staatsbibliothek, Europeana, Rijksmuseum cultural resources,	Metadatabase-enhanced exploration.	Usability study/criteria-based study	Usability, user satisfaction, accessibility	User	Qualitative	x		
Harris et al. (2018)	MOSAICA project Search and Mining Tools with Linguistic Analysis system	Probabilistic approach	Criteria-based study	Accessibility, data quality, performance evaluation	User	Quantitative	x		x
Harsányi, Rozinajová, and Andrejčíková (2012)	Art museums; national libraries of several countries	Semantic interoperability	Criteria-based study	Data quality	Expert	Qualitative			x

(continued on next page)

(Continued)

Authors and year of source	Name	Framework/Approach	Method	Criteria	Perspective	Data type	User centred	Interface	System centred
Hill et al. (2000)	Alexandria Digital Library	Geolibrary	Criteria-based study	Usability, accessibility	User	Both	x	x	
Hu, Ho, and Qiao (2017)	Mogao Cave Panorama Digital Library	Framework of evaluation criteria includes the aspects of effectiveness, efficiency, satisfaction, and interactivity	Usability study/criteria-based study	Usability, user satisfaction	Expert/user	Both	x		
Hu, Ng, and Xia (2018)	General cultural heritage work in China (references Mogao Cave)	User-centered design, information representation design, and grounded theory	Criteria-based study	Usability, data quality, accessibility	Expert/user	Qualitative	x		x
Hug and Gonzalez-Perez (2012)	Institute of Heritage Sciences (Incipit) at the Spanish National Research Council	Incipit information system design framework	Criteria-based study	Data quality, usability	Expert	Qualitative	x		x
Ibrahim and Ali (2018)	Malay house virtual heritage environment	(1) Information design; (2) information presentation; (3) navigation mechanism; and (4) environment setting	Criteria-based study	Usability, cultural content, user satisfaction	Expert/user	Both (mainly qualitative)	x	x	

(continued on next page)

(Continued)

Authors and year of source	Name	Framework/Approach	Method	Criteria	Perspective	Data type	User centred	Interface	System centred
Jeng (2008)	New Jersey Digital Highway	Usefulness assessment; Technology Acceptance Model	Criteria-based study	Usability, user satisfaction	Expert/user	Qualitative	x	x	
Komlodi, Caidi, and Wheeler (2004)	Biblioteca Italiana, Early Canadiana Online, Gallica, Library of Congress National Digital Library, Neumann House, Proyecto Biblioteca Digital Argentina	Four usability design guidelines: language and visual representations, content selection, and content organization	Criteria-based study	Usability, coverage, accessibility, cultural content	User	Qualitative	x	x	
Liew (2005)	Ranfurly Collection, Niupepa Collection	Cross-cultural usability	Usability study	Usability, accessibility, data quality, cultural content, coverage	User	Qualitative	x		
Liew and Chowdhury (2016)	New Zealand Electronic Text Collection, Kete Horowhenua, New Zealand History Online	Economic sustainability, social sustainability, and environmental sustainability	Criteria-based study	Cultural content, coverage, accessibility, usability	User	Qualitative	x		
Marchionini (2000)	Perseus digital library	Evaluation as a research and problem-solving endeavour	Impact study	Impact criteria	User	Both	x		

(continued on next page)

(Continued)

Authors and year of source	Name	Framework/Approach	Method	Criteria	Perspective	Data type	User centred	Interface	System centred
Marketakis et al. (2017)	ResearchSpace, ARIADNE	Synergy Reference Model and the X3ML mapping definition language	Criteria-based study	Performance evaluation	Expert	Quantitative			x
Matthews and Aston (2012)	Wendy James's anthropological archive	Narrative-based approach	Criteria-based study	Accessibility, usability, data quality	User	Qualitative	x	x	x
Melucci and Orío (2004)	Digital Archive for the Venetian Music of the Eighteenth Century	System for Music Information Retrieval Environments	Log file analysis	Data quality	Expert	Quantitative			x
Núñez and Repiso (2019)	87 digital collections of cultural heritage of the Canary Islands	Interoperability	Criteria-based study	Data quality, accessibility, coverage, cultural content	Expert	Both	x	x	
Oomen et al. (2013)	Netherlands Institute of Sound and Vision	TREC Video Retrieval Evaluation	Log file analysis	Usage statistics and patterns, user satisfaction	User	Both	x		x
Pallas and Economides (2008)	210 web sites of art museums from all over the world	Museum's Sites Evaluation Framework	Criteria-based study	Usability, data quality, user satisfaction, coverage	Expert	Both	x	x	x
Pattueli (2011)	Library of the University of North Carolina at Chapel Hill (Tobacco Bag Stringing)	Development of the Tobacco Bag Stringing ontology followed Methontology as the general methodology framework	Usability study	User satisfaction, usability, accessibility	User	Qualitative	x	x	

(continued on next page)

(Continued)

Authors and year of source	Name	Framework/Approach	Method	Criteria	Perspective	Data type	User centred	Interface	System centred
Punzalan, Marsh, and Cools (2017)	American History and Anthropology Museums	Toolkit for the Impact of Digitised Scholarly Resources, and Archival Metrics.	Impact study	Impact criteria	Expert/ user	Qualitative	x		
Shiri (2018)	Digital Library North	Multi-disciplinary participatory methodological framework	Criteria-based study	Cultural content, coverage, accessibility, usability	Expert/ user	Qualitative	x		
Shiri and Stobbs (2018)	Digital Library North	Culturally aware, multi-method and multidisciplinary user evaluation framework	Usability study/criteria-based study	Usability, user satisfaction, cultural content, accessibility	User	Qualitative	x	x	
Skevakis et al. (2014)	Natural Europe project	Natural Europe Cultural Environment (NECE)	Usability study	Usability, data quality, user satisfaction	Expert	Qualitative	x	x	
Steiner et al. (2014)	CULTURA	Interaction Triptych Model	Usability study	Usability, user satisfaction, data quality, coverage, performance evaluation	Expert/ user	Both	x	x	x

(continued on next page)

(Continued)

Authors and year of source	Name	Framework/Approach	Method	Criteria	Perspective	Data type	User centred	Interface	System centred
Stiller (2014)	Brooklyn Museum, British Library, National Archief, Europeana, Historypin, International Children's Digital Library	Grounded theory approach; information interaction framework	Criteria-based study	Accessibility	Expert	Both	x		x
Stiller, Gäde, and Peiras (2013)	Europeana	Multilingual access in CHDL	Criteria-based study	Accessibility	Expert	Qualitative	x	x	x
Suire et al. (2016)	Experimental corpus of 240 cultural heritage documents	Marchionini's framework	Log file analysis	Usage statistics and patterns	User	Quantitative	x		x
Sulé Dueso, Rius, and García (2011)	31 Spanish heritage repositories	Alvite Diez's cultural heritage evaluation	Criteria-based study	Usability, coverage, user satisfaction, data quality	Expert	Qualitative	x	x	x
Szabo, Lacedelli, and Pompanin (2017)	Dolom.it, the virtual museum of Dolomites landscape	Interpretative framework	Impact study	Impact criteria, coverage, cultural content	User	Qualitative	x		
Van den Akker et al. (2013)	Agora	Digital hermeneutics	Criteria-based study/usability study	Usability, coverage, cultural content, accessibility, data quality	Expert/user	Qualitative	x	x	

(continued on next page)

(Continued)

Authors and year of source	Name	Framework/Approach	Method	Criteria	Perspective	Data type	User centred	Interface	System centred
Van Hooland et al. (2013)	Powerhouse Museum in Sydney	Linked data approach	Criteria-based study	Data quality	Expert	Both			x
Vila-Suero et al. (2019)	Online Public Access Catalogues (catalogo.bne.es, datos.bne.es)	Semantic web technologies/linked data	Log file analysis/usability study	Usability, user satisfaction, performance evaluation, error rate	Expert	Both	x	x	x
Wang et al. (2013)	Europeana	Three part framework: (1) fast clustering; (2) hierarchical structuring; and (3) focal semantic clusters	Criteria-based study	Data quality, performance evaluation	User	Both	x		x
Xie (2006)	Library of Congress American Memory Project, ACM Digital Library, Electronic Poetry Center at State University of New York, Buffalo	User-centered approach	Criteria-based study	Accessibility, usability, coverage, user satisfaction, performance evaluation	User	Both	x	x	x
Yelmi, Kuşcu, and Yantaç (2016)	Soundscape of Istanbul project, Soundsslike project	User-centered design approach	Criteria-based study	Accessibility, data quality, cultural content	User	Qualitative	x	x	