

# Biases within Digital Repositories:

## The Getty Research Portal

Hande Sever

### Introduction

The influence of computer science on disciplines outside of itself has constituted hybrid communities of practice, such as the digital humanities (DH) and digital art history (DAH), the latter being an offspring of the former.<sup>[1]</sup> The questions asked within these interdisciplinary fields are substantially different from those in core computer science fields, such as systems theory, language theory, and theory of computation.<sup>[2]</sup> As a result of the interdisciplinary nature of their inquiries, research questions pertaining to DH and the ethical challenges that accompany them are diametrically opposed to the empirical and often unambiguous measures of validity to which computer scientists are accustomed. This article will explore ethical questions pertaining to collections aggregation systems from the perspective of postcolonial scholarship and seek paths towards addressing the ethical questions currently facing many DH projects. To do so, I will use the Getty Research Portal (hereafter, “the Portal”) as a case study of a digital repository and draw on my personal experience as a software developer working on the project, as well as on current research on biases and ethics within DH.

Among the many challenges arising from the rapid expansion of DH and DAH, the question of how to ethically conduct projects, minimize biases in data production, and meaningfully include a diverse set of voices within the field still remains unanswered, despite ongoing efforts to develop projects with a more critical perspective, some of which will be examined in the latter sections of this paper. Since their inception, both DH and DAH have often tended to be anti-interpretive, especially when interpretation is understood as a political activity.<sup>[3]</sup> Instead, both fields have primarily aimed to archive materials, produce data, and develop software, which led to the rise of seemingly neutral software

tools that further canonized the oeuvre of dead white men.[4] According to Alan Liu, in the DH, cultural criticism—in both its interpretive and advocacy modes—has been noticeably absent, especially when compared with the “mainstream” humanities or, even more strikingly, with “new media studies” (the latter being populated by net critics, tactical media critics, and hacktivists) or Science and Technology Studies, where the cultural, behavioral, and social impact of classification systems has long been a key object of study.[5][6] While recent efforts to bridge these gaps are undeniable, digital humanists have been focused on developing tools, data, and metadata, but rarely does the field extend its core inquiries into the registers of economics, politics, or culture. Thus, a majority of DH projects have been concentrated on literary and artistic figures already considered part of the canon.[7] This dynamic is particularly visible when it comes to large-scale, museum-driven projects, which by their very nature define the boundaries of the canon.

One such project is the Portal. I will build upon my experience as a software developer working on the Portal specifically, and within the field of DH more generally, to examine the Portal as a case study of the multiple biases facing DH projects. The Portal is an online platform providing global access to digitized art history texts from an international group of contributing institutions. Institutions such as the Belvedere Research Center in Vienna, the National Gallery Library in London, the Leo Baeck Institute Library in New York, and the Zentralinstitut für Kunstgeschichte in Munich are among the project’s current contributors. The project aims to provide an online platform functioning as a centralized aggregator and search engine for digitized art history texts and claims no geographical limits, being described by the Getty Research Institute as “a global resource for the history of art of all cultures.”[8] Through my work on this project, I came to witness firsthand the methodological dominance of the Euro-American DH community and how its hegemony in the field sets the standards for the rest of the world, while at the same time being at the forefront of efforts to remedy such biases. This situation currently pushes peripheral cultures towards invisibility and favors a monopoly on the ways in which knowledge is understood and disseminated, due to the cultural, political, and linguistic bias of its digital standards, protocols, and interfaces.[9] This paper aims to investigate this ongoing struggle by uncovering biases inherent in the Portal through a categorical system developed by Batya Friedman and Helen Nissenbaum, and to identify the latent possibility of a different ethical framework.[10]

## An Overview of the Getty Research Portal

The Portal was launched in 2012 by the Getty Research Institute, an operating program of the J. Paul Getty Trust. The J. Paul Getty Trust is an international cultural and philanthropic institution devoted to the study, exhibition, protection, production, and dissemination of visual art and

visual art-related knowledge through a set of four operating programs—the Getty Museum, the Getty Conservation Institute (GCI), the Getty Foundation, and the Getty Research Institute (GRI)—which are split across two campuses: the Getty Center in West Los Angeles and the Getty Villa in Malibu. These four institutions, operating in concert, make the Getty Trust one of the biggest and richest institutions in the world of visual arts today.[11] The Getty Museum, split between the Center and the Villa, exhibits antique, medieval, modern, and contemporary art. Despite its relative difficulty of access from the city center, it is one of the most visited museums in the United States.[12] The GCI, similarly split between both campuses, is a private research institution dedicated to conservation practices.[13] The Getty Foundation, chaired in the Getty Center, is a philanthropic organization which can award up to 0.75% of the Getty Trust’s endowment in grants to individuals and institutions at its own discretion.[14] Finally, the GRI, contained entirely within the Getty Center, is a research center dedicated to producing knowledge related to the visual arts.[15]

Funded by the trust’s private capital, the GRI has greatly evolved from its beginnings as a small library—the *Los Angeles Times* described it as a “fledgling study center” after the departure of its first director in 1992[16]—into one of the prime centers for art research internationally, a key piece of the Getty Trust’s overall cultural imprint, and one of the foremost places where the worlds of exhibition-making, visual art, and academia interact. This influence has to be credited to a number of key factors, notably its private collections of millions of books, documents, and photographs and its considerable outreach, both to the academic world and to the public at large, through its scholar-in-residence program, publications, and talks and workshops, as well as through an emphasis on accessibility; while the GRI’s collection does not circulate, it extends library privileges to the public.[17] This understanding of the GRI as an institutional provider of access has driven the GRI to focus greatly on software tools, and more specifically on the development, administration, and maintenance of online databases and tools facilitating the dissemination and researching of art objects and texts.

The project that both defined and provided a model for this institutional direction was Getty Vocabularies, a set of databases which emerged from the Getty Vocabulary Program, a department funded in 1987 within the Research Institute, with the goal of compiling and distributing standardized art-related terminology.[18] This led to the creation of in-house digital tools able to cope with the need for large-scale, robust and flexible databases of terminologies. Originally intended to be used internally, these tools were opened to other institutions due to widespread demand, leading to an ever-growing list of contributors.[19] Nowadays, the Getty Vocabularies set the standard internationally for terminology across most art institutions, and have grown to become powerful automatic databases dictating the acceptable proper names, dates, and classifications pertaining to most art objects, imposing a common centralized standard to art and art-adjacent knowledge production around the world.[20]

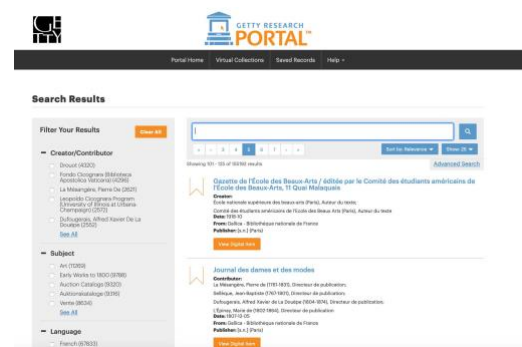


Fig. 2. The Getty Research Portal's search results

The Portal, launched in 2012 by the GRI, was significantly influenced by the Getty Vocabularies project with regard to its aim: establish a singular interface in order to standardize art historical research protocols across a variety of institutions.[21] The Portal is a free online search gateway that aggregates descriptive metadata of digitized art history texts, with links to fully digitized copies which can be downloaded free of charge. The GRI worked with a number of institutions to create the Portal: the Avery Architectural and Fine Arts Library at Columbia University, the Frick Art Reference Library, and the Thomas J. Watson Library of the Metropolitan Museum of Art in New York, as well as members of the New York Art Resources Consortium, the Biblioteca de la Universidad de Málaga in Spain, the Institut National d'Histoire de l'Art in Paris, and the Universitätsbibliothek Heidelberg.[22] Together with the GRI's Library, these art libraries have contributed nearly 100,000 digitized art history texts, which are all immediately searchable and accessible to the greater public via the Portal.

Librarians belonging to cultural institutions across the world have been actively involved in the creation of information repositories, particularly in the last few years, in a trend holding true across many fields.[23] These repositories contain a wide array of resources, including commercial databases, links to academic websites, writing resources, and institutional newsletters.[24] For instance, the UNESCO Libraries Portal Project provides access to websites of library institutions around the world.[25] A large-scale aggregator focusing on textual content, the Portal is one of these information repositories. Unlike other methods of searching for texts online, such as Google Books, every link in the Portal leads to a complete and downloadable digital surrogate. Because the Portal aggregates the metadata of the digitized texts and links to them, instead of keeping them on a server, there are no technical limitations as to how much material can be collected. However, given current restrictions on the digital dissemination of copyright materials, the Portal's content is limited to works published before 1923.[26]

Every year, an increasing number of books and journals are uploaded from the GRI's own collections and from contributing institutions' collections onto the Internet Archive—a nonprofit digital library which offers free access to its contents—and made available through the Portal. As texts are made available, metadata records pertaining to the digitized text need to be ingested by the Portal database from the relevant contributing institution. Metadata is the term applied to information that describes other information, such as labels describing objects, content, or documents.[27] Metadata is used by virtually all record keeping institutions, and specific metadata standards exist for many information fields in libraries, museums, archives, and record keeping environments. The Portal is developed to be able to take full advantage of these classification systems, in order to enable searchability. For instance, a user of the Portal might look for documents pertaining to a subject, written by a specific writer, or having belonged to a particular collection. This descriptive information will have been attached to the relevant

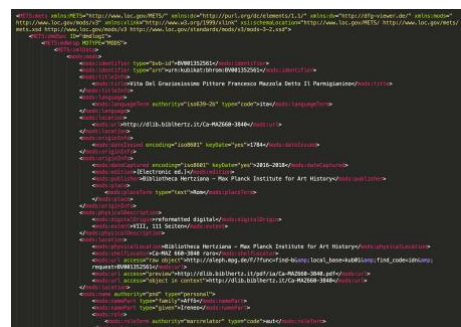


Fig. 3. Metadata of a record from the Bibliotheca Hertziana in METS format.

document as metadata, and as such enables the user to search through the Portal using one of these keywords and find the document in question.

My experience working as a software developer assigned to enabling the addition of metadata records onto the Portal's project provides direct insight into the processes used by the Portal to ingest metadata and their consequences on the project as a whole. The workflow of ingesting records from the project's contributing institutions entails normalizing the datasets as they are fed through a data transformation code that enables them to be uploaded to the Portal once they conform to the Portal's standardized structure. For instance, records that do not include a title will be omitted until further information is received from the contributing institution. Data transformation, the process of converting data from one format or structure into another format or structure, is an integral aspect of most data management and integration, and the necessary backbone to an inter-institutional project such as the Portal. The data transformation required by the Portal is performed via a series of automated and semi-manual steps. The Portal is able to automatically transform records provided in standardized metadata formats such as the Machine Readable Cataloging (MARC), Metadata Encoding and Transmission Standard (METS), Metadata Object Description Schema (MODS), and Dublin Core (DC).<sup>[28]</sup>

Despite this standardization, metadata often comes to the Portal in radically different formats, requiring a constant use of automated data transformation. Data transformation is, in most cases, a batch process, as developers write code or implement transformation rules in a data integration tool and then execute those tools on large volumes of data that would not be processable in a realistic time frame through entirely manual processes. This holds true for the Portal. The lines of code governing these processes, known as transformation codes, are modified to suit the needs of each contributing institution. For instance, my personal responsibilities on the project included developing a new transformation code for records encoded in “METS with MODS” format, as the Warburg Institute in London, initially a METS contributor, asked to continue its contributions through the “METS with MODS” format. The Portal also transforms CSV (comma-separated values) files to facilitate contributions from institutions that may not have access to standardized metadata, in order to open participation to institutions less concerned with metadata generation. This option was specifically developed to include the Tokyo National Research Institute for Cultural Properties as a contributor to the project, and to enable a more diverse set of institutions to join the project in the future. However, regardless of their format (MARC, MODS, DC, or CSV), in order to be properly ingested, all contributions must be in Unicode format—one of the most popular formats for encoding text—when exporting from catalogue systems, as the Portal does not support non-Unicode formats.

Fig. 4. A list of records from the Tokyo National Research Institute for Cultural Properties in CSV

Fig. 5. A record from the Tokyo National Research Institute for Cultural Properties, as seen on the Portal website.

## Digital Epistemology

The Portal project was developed with a dependency on standardized classification systems such as MARC, METS, MODS, and DC, as it requires communication and resource sharing between different institutions. However, the politics of the development and standardization of these classification systems and its effects on DH and DAH projects still remain under-researched. These questions are central to the acknowledged aims of the Portal project, which claims to be “a global resource for the history of art of all cultures.”<sup>[29]</sup> This type of universalist claim needs to be examined in the context of the geographical diversity of contributors, in order to assess the actual state of subordination and bias from which non-English-speaking cultural institutions suffer. For instance, as of February 2020, the Portal does not include a single contributor from the African continent or Southwest and Central Asia.<sup>[30]</sup> The framework proposed within this paper is not to be understood only as a criticism of these results, but as a tool intended to help interpret probable causes and remedies.

Biases inherent to the Portal result from its status as a digital repository in which a social institution, algorithmic tools, and users interact. Professors in Information Science Batya Friedman and Helen Nissenbaum provide a solid theoretical framework for analyzing injustices in algorithmic systems by focusing on the structure of discovered bias.<sup>[31]</sup> They outline three categories of bias: preexisting, technical, and emergent. Preexisting bias is rooted in social institutions, practices, and attitudes that predate the system, and lives independently of the system itself. Technical bias arises when various technical properties of the systems created with false assumptions about their use are applied. Emergent bias arises in direct contact with users. Emergent bias was not intentionally designed into the system, nor was it created through false assumptions by the designers—it emerges out of the interaction between the system and its users.

The founding contributors of the Portal are based in the United States, France, Germany, and Spain, and their contributions are all in Western European languages. The Portal can nonetheless depict characters outside of the Latin alphabet, due to its use of the Unicode format, which enables the display, for instance, of contributions from the Tokyo National Research Institute for Cultural Properties, a non-founding contributor which uses Japanese scripts for most of its contributions. Following the model provided by Friedman and Nissenbaum, two forms of biases can be discovered in the Portal’s interaction with contributing institutions: preexisting and technical. Since its inception, the GRI has dedicated most of its funding to developing ties with European institutions, undervaluing the cultural heritage of the non-European world—a fact plainly visible in the choice of the Portal project’s inaugural contributors.<sup>[32]</sup> As Domenico Fiormonte argues in *Towards a Cultural Critique of Digital Humanities*, “Even though so much effort has been expended in making existing DH more international, the

impression remains the same: a solid Anglo-American stem onto which several individuals of mostly European countries are grafted.”[33] This institutional bias unfortunately holds true for the Portal project as well, at least in its inception. A second bias pertaining to the Portal is perhaps inherited from the operating program it originates from. The GRI is a physical library, and the Portal, following this model, is an essentially bibliographic database, limiting itself to “the printed literature of art.”[34] This excludes much present-day scholarly work, such as recorded lectures, online projects, or digitally published research and catalogues. While the Portal does in fact link to some online-only resources, such as the catalogues produced by the Getty Foundation’s own Online Scholarly Catalogue Initiative, it links to these projects generally, rather than to specific essays within a project, and award-winning digital projects such as Smarthistory are altogether missing.[35]

A technical bias arises from the use of the Unicode standard as a way to depict non-Latin alphabet characters, and from the Portal’s insistence on accepting records in Unicode format. The Unicode standard is developed by the Unicode Consortium, a nonprofit organization “devoted to developing, maintaining, and promoting software internationalization standards and data, particularly the Unicode Standard, which specifies the representation of text in all modern software products and standards.”[36] However, as examined by Antonio Perri, there are technical biases embedded in the use of the Unicode standard, as it makes incorrect or subpar assumptions with regard to some encoding problems.[37] In his research, Perri considered a number of encoding solutions proposed by the Unicode consortium for texts written in Indian sub-continental scripts, Chinese, Arabic, and Hangul. In all cases, in addition to an excessive dependence on visualization software, which makes portability and bandwidth use serious problems, he demonstrated that Unicode’s solutions were rooted in a hyper-typographic understanding of writing; that is to say, based on a model of Western writing structured by its logical sequencing. By neglecting their visual features, this understanding of text had overlooked important functional aspects of many writing systems. Perri gives a striking example of this bias when discussing Unicode’s treatment of ligatures and the position of vowel characters in the Devanagari script. In many Indian systems, visual aspects of a text as a whole will prevail over the graphemes’ reading order. The Unicode encoding, however, makes a faulty assumption in privileging the logical scheme of the text, ignoring typographic and visual elements that are directly constitutive of the application of the writing scheme.

Another example of technical bias contained within the Unicode standard is provided by Yoshiki Mikami et al. in *Language Diversity on the Internet: An Asian View*. [38] The Mongolian language can be written either in Cyrillic script or through its own traditional graphemes, of which more than eight different codes and fonts have been identified. No standardization of these fonts has been provided, causing inconsistency, encoding errors, and mistranslation between different programs attempting to parse texts in Mongolian language. As a result, some Mongolian webpages have to resort to using image files for



their typographic content, which increases load time and bandwidth use. According to Mikami et al., many Indian web pages face the same challenge. Some Indian news providers resort to the use of proprietary fonts for Hindi scripts, while some use heavy image files for textual content. These faults in the standards of encoding methods prevent information from many groups using non-Anglo-American scripts to be read, exchanged, and parsed, and as such lead to a consequent digital language divide. This technical bias is reproduced by the Getty Portal through its use of the Unicode standard as an obligatory encoding method. For instance, while Japanese alphabets such as those used by the Tokyo National Research Institute for Cultural Properties are adequately served by the Unicode standard, institutions desiring to provide records in Mongolian languages or Indian subcontinental scripts would face rejection or expect distortions and inconsistencies.

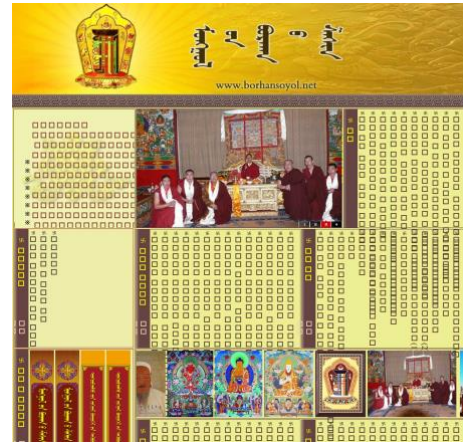


Fig. 6: Website using the traditional Mongolian script, accessed from a standard browser.

Apart from these technical biases inherent to the Portal's own decisions with regard to encoding, the conceptual structure of metadata itself poses a problem for the Portal's avowed claim of being a global resource. In *Technologies of Social Regulation*, the authors argue that metadata standards such as MARC—one of the standards used by the Portal—are, first and foremost, commercial entities within the production system of advanced capitalism, and that their emphasis on socially virtuous goals, such as ease of access and freedom of information, obscures the fact that these technologies are not socially neutral or benign, but operate solely within a capitalist framework.[39] Developed by the Library of Congress, MARC first appeared in the mid-1960s, became a national standard in the United States in 1971, and then an international one in 1973.[40] MARC has thoroughly given cataloging a distinctly Fordist quality—the privileging of standards and consistency over the needs of specific communities of users—and libraries that could not afford MARC-compliant equipment and technology quickly became marginalized.[41]

The problem of technical biases inherent to metadata standards is not specific to MARC, but applies to metadata standards as a whole. Metadata systems identify, sort, and make knowledge accessible. It is this expansive research power which makes the Portal possible, but this power comes at the cost of an overdetermination of the models within which knowledge within a field is conceived. A set of expectations of what knowledge looks like and what it does is produced by the use of classification systems, and therefore it is of little surprise that these systems would be more accessible to specific groups and limited in use to others, or that they would replicate existing structures of dominance and exploitation within the epistemic field.[42] As Johanna Drucker writes, “No classification system is value neutral, objective, or self-evident, and all classification systems bear within them the ideological imprint of their production.”[43]

In *How We Construct Subjects: A Feminist Analysis*, Hope Olson asserts that metadata is developed by the West and does not apply to all forms of knowing.[44] She argues that



classification schemes used in cataloging, indexing, and metadata standards are grounded in a hierarchy grown from traditional logic which excludes alternative forms of thought. For instance, Olson argues that the body of feminist thought which identifies women's knowledge of the world "as an interconnected web offers a radically different model from the hierarchical structure of traditional logic."<sup>[45]</sup> Other scholars have offered similar criticism of Western-centric knowledge classification. Donald Fixico writes, "'Indian Thinking' is 'seeing' things from a perspective emphasizing that circles and cycles are central to the world and that all things are related within the universe."<sup>[46]</sup> Linda Tuhiwai Smith similarly sees the hierarchies of classification as a tool of imperialism and as exemplifying a positivist approach to knowledge in general, and research in particular.<sup>[47]</sup> Regardless of encoding-based technical solutions, in relying on metadata standards the Portal already limits the forms of knowing which it can replicate, leaving outside of its purview alternative forms ungraspable by its classification systems.

As for emergent bias, uncovering the ways in which the interaction between the Getty Research Portal produces biases separate from the biases inherent in either preexisting institutional practices or assumptions made when developing the system would require a study of the real-life patterns of use of the Portal. This would necessitate obtaining accurate data on user experience and the impact of the Portal on projects employing it as a tool, both of which fall outside the purview of this paper. However, more research on the emergent biases created by the interaction between private and institutional actors and collections aggregate systems is needed, and would undoubtedly produce striking results.

## Solutions

Without question, preexisting bias with regard to the geographic selection of contributors should be addressed by reaching out more consistently to institutions outside of the Global North. To this end, Getty-led collaborative events and exhibitions have already proven to be a strong driver of inclusion. *Pacific Standard Time: LA/LA*, a series of exhibitions organized collaboratively across Southern California exhibition spaces to highlight the work of Latin American artists, concluded with multiple Latin American institutions, such as the Fundación Espigas in Buenos Aires, joining the Portal. The success of *Pacific Standard Time: LA/LA* with regard to these new inclusions demonstrates the efficacy of collaborative programming in building long-term partnerships which can diversify the list of contributing institutions.<sup>[48]</sup> Furthermore, new online projects and online-only scholarly work is being produced every year, including by the GRI itself, and the Portal should broaden its scope to include them, to produce an accurate view of the current state of art history as a field.<sup>[49]</sup>

As for technical biases, the Portal team cannot directly solve the problems inherent in existing metadata standards such as MARC and in information technology standards such as

Unicode, but it can attempt to provide a much more flexible barrier of entry for the data provided by contributing institutions and accept datasets which are not formatted through these standards. Efforts have already been made in that direction, notably through the recent development of methods for ingesting CSV files. Further efforts can be undertaken jointly with contributors, as the Getty continues financially and technically supporting institutions unable (for monetary or technical reasons) to provide the data structures necessary for their records to be included. If the Portal wants to claim its status as “a global resource for the history of art of all cultures,” it needs to provide adequate support to institutions for which participation would be, all things left equal, either impossible or extremely difficult, and it needs to do so by developing methods able to accommodate ingesting a more diverse set of records.

It should be noted that these solutions directly reinforce each other: the inclusion of diverse institutions, if the needs of these institutions are met and supported through the Getty’s extensive financial and technical resources, can be a driver in addressing biases inherent in the Portal in a self-reinforcing, virtuous loop. For example, as discussed earlier in this paper, the inclusion of the Tokyo National Research Institute for Cultural Properties drove the development of features enabling the Portal to ingest CSV files.

As for the technical biases inherent to metadata standards, while limits to the knowledge which can be represented fairly through the classification systems inherent to the Portal exist, recent efforts by similar DH projects to account for non-Western-centric models of knowledge provide examples of museum-led aggregation systems accounting for diverse subjectivities. One such example is the Collections Online platform of the Museum of New Zealand Te Papa Tongarewa (Te Papa).<sup>[50]</sup> Te Papa’s Collections Online attempts to present an accessible record of its extensive and highly diverse collections, ranging from fossils to Taonga Maori (Maori cultural treasures), a collection constituted through collaborative processes with Maori and Pacific Islander artists, academic, and communities.<sup>[51]</sup> Te Papa’s Collections Online has attempted to correct a legacy of colonial ethnographical museology, replacing it with a bicultural museology which accounts for Maori and Pacific Islander models of knowledge. This has driven the technology, curation, and display mechanisms of Collections Online. For instance, Te Papa does not include images of Taonga in Collections Online without consulting the *iwi* (tribe) from which they originate. Similarly, images of Taonga Maori are not copyright approved to be used outside of research, study, personal, or educational purposes.<sup>[52]</sup> As for technical solutions, Collections Online makes a point of supporting the macron, a relatively rare diacritical mark that is used extensively in written Maori languages, for both display and search.<sup>[53]</sup> Collections Online also uses extensive culturally specific categories in their metadata to facilitate accessibility and provide an accurate representation of Maori and Pacific Islander cultural objects within the context in which they are understood by their originators. In doing so, Collections Online demonstrates that, even within a centralized platform

**Item details**

<b>Name</b>	Umu pack
<b>Production</b>	Unknown; circa 2008
<b>Classification</b>	boxes
<b>Materials</b>	cardboard
<b>Dimensions</b>	350mm, 300mm, 400mm
<b>Registration Number</b>	FE012635
<b>Credit line</b>	Gift of Reverend Iamanu Amaama, 2011

**Overview**

**Umu pack**  
 This small rectangular cardboard box is an umu pack. In the Samoan language, the term umu refers to food cooked in an earth oven. This specially made box is used by air travellers who wish to bring foodstuff that was cooked in an umu into New Zealand.

*Fig. 7. Te Papa’s Collections Online page, providing item details for an Umu pack.*

aggregating diverse and culturally heterogeneous content, technical and institutional solutions can be devised which enable the deployment of diverse models of knowledge, an example which could prove beneficiary to the Portal as it diversifies its contributors. As for emergent biases, future research will surely uncover unforeseen biases resulting from ongoing interactions between the Portal and its users. As such, a flexible and vigilant approach to updating the Portal's systems as these biases are discovered will prove necessary.

## Conclusion

This paper provides an account of the Portal project, identifies preexisting and technical biases which have prevented it from reaching its avowed aims, and suggests solutions and alternative frameworks through which these aims could be reached. In general, a sensitivity to hidden biases is crucial for individuals working on projects that engage with standardized classification systems, since an attention to the cultural impact of classification systems, and the systems of domination they can uphold, ends up demonstrating both the power these systems hold as well as the ideologies that produce their fault lines. As Johanna Drucker writes, "A sensitivity to these issues is not only important, but enlightening in its own right, since the cross-cultural or cross-constituency perspective demonstrates the power of classification systems, but also, our blindspots."<sup>[54]</sup> The claim of the Portal to be "a global resource for the history of art of all cultures" contrasts with the fact that the Portal still does not have a single contributing institution from Africa or Southwest and Central Asia. This absence disregards the ongoing digitization efforts that are flourishing in these regions, efforts undertaken despite the subordination from which non-English-speaking digital humanists suffer, and the technical biases they have to overcome.

Looking at the design and development, both technical and institutional, of the Portal, the project exhibits, like many similar DH and DAH projects, a double bias: a technical one and a preexisting one. It is of note that these two biases are entangled, and that it is difficult to discern where the technological choice begins and where the cultural prejudice ends. Undervaluing the cultural heritage of the non-Euro-American world leads to faulty assumptions about the technical solutions needed to access the knowledge produced by it, and these same faulty assumptions reinforce the undervaluation.<sup>[55]</sup> This case study poses conclusively the urgent need to elaborate a different guiding model for the Portal, one based on the concept of knowledge as underwritten by international inclusive commons and the cultivation of cultural margins, and likewise opposed to the DH's present obsessions with large-scale digitization projects and "archiving fever" entirely focused on Euro-American heritage.<sup>[56]</sup> Fortunately, as preexisting and technical biases are entangled, so are cultural and technical solutions. Efforts to support both technically and culturally a diverse range of contributors, as well as to include different

models of knowledge within the boundaries of the project, can begin to trace a different future for the Portal. The Getty Trust has a significant track record of directing its immense cultural and financial capital towards diverse local efforts, notably in antique conservation.<sup>[57]</sup> A similar effort targeted towards addressing biases in projects such as the Portal would prove an invaluable benefit for Digital Art History, and for heritage and conservation projects more generally.

## Biography

Hande Sever is an artist, technologist, and theorist working at the intersections of postcolonialism, vernacular architecture, and emerging technologies. Her writing has been published by the Getty Research Journal, Art Institute Review, Journal of Arts & Communities, and X-TRA Contemporary Art Quarterly, among others. She is currently a PhD candidate in the Art History, Theory, and Criticism program at the University of California San Diego (UCSD). Prior to joining UCSD she was a visiting faculty member of the California Institute of the Arts' (CalArts) Photography and Media program.

---

[1] Susan Schreibman, Raymond George Siemens, and John Unsworth, eds., *A Companion to Digital Humanities* (Malden, MA: Blackwell Pub, 2004); Claire Bishop, "Against Digital Art History," *International Journal for Digital Art History*, no. 3 (July 27, 2018): 123–124, <https://doi.org/10.11588/dah.2018.3.49915>.

[2] Matthew Kirschenbaum, "What Is Digital Humanities and What's It Doing in English Departments?," in *Debates in the Digital Humanities*, ed. Matthew K. Gold (Minneapolis: University of Minnesota Press, 2012), 3–11.

[3] Bishop, "Against Digital Art History," 127–128.

[4] Daniel Allington, Sarah Brouillette, and David Golumbia, "Neoliberal Tools (and Archives): A Political History of Digital Humanities," *Los Angeles Review of Books* (2016), accessed March 26, 2020, <https://lareviewofbooks.org/article/neoliberal-tools-archives-political-history-digital-humanities/>.

[5] Alan Liu, "Where Is Cultural Criticism in the Digital Humanities?," in *Debates in the Digital Humanities*, Matthew K. Gold (Minneapolis: University of Minnesota Press, 2012), 490–509.

[6] See, for instance, Geoffrey C. Bowker and Susan Leigh Star, *Sorting Things out: Classification and Its Consequences*, 1st paperback ed., Inside Technology (Cambridge, MA, and London: The MIT Press, 2000).

[7] Allington, Brouillette, and Golumbia, "Neoliberal Tools (and Archives)."

- [8] "Getty Research Institute Launches Gateway to the World's Art Libraries | News from the Getty," About the Getty, May 10, 2012, <http://news.getty.edu/getty-research-portal.htm>.
- [9] Domenico Fiormonte, "Towards a Cultural Critique of the Digital Humanities," *Historical Social Research / Historische Sozialforschung* 37, no. 3 (141) (2012): 59–76.
- [10] Batya Friedman and Helen Nissenbaum, "Bias in Computer Systems," *ACM Transactions on Information Systems* 14, no. 3 (July 1996): 330–347, <https://doi.org/10.1145/230538.230561>.
- [11] The J. Paul Getty Trust, "Financial Statements. June 30, 2019 and 2018 (With Independent Auditors' Report Thereon)," accessed March 29, 2020, <https://www.getty.edu/about/governance/pdfs/auditedfinancials.pdf>.
- [12] The J. Paul Getty Trust, "Getty Facts," 2019, accessed March 29, 2020, [https://www.getty.edu/about/whoweare/pdfs/getty\\_fact\\_book\\_2019.pdf](https://www.getty.edu/about/whoweare/pdfs/getty_fact_book_2019.pdf).
- [13] "Overview | Getty Conservation Institute," accessed March 30, 2020, <https://www.getty.edu/conservation/about/overview.html>.
- [14] Suzanne Muchnic, "Director Wants a Wider Reach for Getty Grants : Art: Deborah Marrow Aims to Broaden the Program's International Constituency. It Has Awarded \$20 Million in Its First Five Years.," *Los Angeles Times*, March 5, 1990, accessed March 24, 2020, <https://www.latimes.com/archives/la-xpm-1990-03-05-ca-1444-story.html>.
- [15] "About the Research Institute (Getty Research Institute)," accessed March 30, 2020, <https://www.getty.edu/research/institute/>.
- [16] Suzanne Muchnic, "Getty Center's Kurt Forster Resigns Post – Los Angeles Times," *Los Angeles Times*, March 20, 1992, accessed March 24, 2020, <https://www.latimes.com/archives/la-xpm-1992-03-20-ca-4303-story.html>.
- [17] "Library Access and Reader Privileges (Getty Research Institute)," accessed March 30, 2020, <https://www.getty.edu/research/library/using/access/>.
- [18] "Getty Vocabularies (Getty Research Institute)," The Getty Research Institute, accessed March 29, 2020, <https://www.getty.edu/research/tools/vocabularies/>.
- [19] "Contributors to Getty Vocabularies (Getty Research Institute)," The Getty Research Institute, accessed March 29, 2020, <http://www.getty.edu/research/tools/vocabularies/contributors.html>.
- [20] Patricia Harpring, "Development of the Getty Vocabularies: AAT, TGN, ULAN, and CONA," *Art Documentation: Journal of the Art Libraries Society of North America* 29, no. 1 (2010): 67–72.

[21] Kathleen Salomon, "Facilitating Art-Historical Research in the Digital Age: The Getty Research Portal," *Getty Research Journal*, no. 6 (2014): 137–141, <https://doi.org/10.1086/675796>.

[22] "Getty Research Portal™ (Getty Research Institute)," The Getty Research Institute, accessed March 29, 2020, <http://www.getty.edu/research/tools/portal/index.html>.

[23] Katherine McNeill, "Repository Options for Research Data," in *Making Institutional Repositories Work*, eds. Burton B. Callicott, David Scherer, and Andrew Wesolek (West Lafayette: Purdue University Press, 2016), 15–30, <https://doi.org/10.2307/j.ctt1wf4drq.7>, 15.

[24] Leckie, Given, and Campbell, "Technologies of Social Regulation," 239.

[25] "UNESCO Libraries Portal | Library of Congress," Library of Congress, accessed March 31, 2020, <https://www.loc.gov/item/lcwaN0020019/>.

[26] The cut-off line for texts to be considered part of the public domain in the United States is 1923, even though not all texts available through the Portal are in the public domain. See "FAQS - Getty Research Portal," Getty Research Portal, accessed March 29, 2020, <https://portal.getty.edu/faq>.

[27] Standardized bibliographic metadata on library records typically include at least title, author, publisher, place of publication, date, and some description of the contents, the physical features, and other attributes of the object. See Tony Gill, "Metadata and the Web," in *Introduction to Metadata: Third Edition*, Murtha Baca, with assistance from Joan Cobb, Nathaniel Deines, and Moon Kim (Getty Publications, 2016), 21.

[28] These formats are developed and maintained by the Library of Congress in Washington, D.C., with the exception of Dublin Core, which was developed during the 1995 invitational OCLC/NCSA Metadata Workshop in Dublin, Ohio. See "Standards | Librarians and Archivists | Library of Congress," web page, Library of Congress, Washington, D.C., 20540 USA, accessed March 29, 2020, <https://www.loc.gov/librarians/standards> and "DCMI: DCMI History," Dublin Core Metadata Initiative, accessed March 29, 2020, <https://dublincore.org/about/history/>.

[29] "Getty Research Institute Launches Gateway to the World's Art Libraries | News from the Getty."

[30] "Contributors - Getty Research Portal," Getty Research Portal, accessed March 29, 2020, <https://portal.getty.edu/contributors>.

[31] Friedman and Nissenbaum, "Bias in Computer Systems."

[32] "Contributors - Getty Research Portal," Getty Research Portal, accessed March 29, 2020, <https://portal.getty.edu/contributors>.

[33] Fiormonte, "Towards a Cultural Critique of the Digital Humanities," 64.

[34] "FAQS - Getty Research Portal."



[35] "About Smarthistory – Smarthistory," Smarthistory, accessed May 20, 2020, <https://smarthistory.org/about/>.

[36] "Unicode Consortium," Unicode.org, accessed March 29, 2020, <https://unicode.org/consortium/consort.html>.

[37] Antonio Perri, "Al di là della tecnologia, la scrittura. Il caso Unicode," *ode. Annali dell'Università degli Studi Suor Orsola Benin II* (2009): 725–748.

[38] Yoshiki Mikami et al., "Language Diversity on the Internet: An Asian View," in *Measuring Linguistic Diversity on the Internet* (Paris: UNESCO Report 142186, 2005), 91–103, accessed March 29, 2020, [http://uis.unesco.org/sites/default/files/documents/measuring-linguistic-diversity-on-the-internet-ict-2005-en\\_0.pdf](http://uis.unesco.org/sites/default/files/documents/measuring-linguistic-diversity-on-the-internet-ict-2005-en_0.pdf). See also S. Nandasara et al., "An Analysis of Asian Language Web Pages," *International Journal on Advances in ICT for Emerging Regions (ICTer)* 1 (March 26, 2009), <https://doi.org/10.4038/icter.v1i1.448>.

[39] Gloria Leckie, Lisa Given, and Grant Campbell, "Technologies of Social Regulation: An Examination of Library OPACs and Web Portals," *Information Technology in Librarianship: New Critical Approaches* (Santa Barbara, CA: ABC-CLIO, 2009), 221–259.

[40] Lois Mai Chan and Theodora Hodges, *Cataloging and Classification: An Introduction*, 3rd ed. (Lanham, MD: Scarecrow Press, 2007), 448.

[41] Leckie, Given, and Campbell, "Technologies of Social Regulation," 229.

[42] See, for instance, Celeste Tu'ò'ng Vy Sharpe and Timothy B. Powell, "Making Digital Humanities Tools More Culturally Specific and More Culturally Sensitive," in *Teaching with Digital Humanities*, eds. Jennifer Travis and Jessica DeSpain, Tools and Methods for Nineteenth-Century American Literature (Champaign, IL: University of Illinois Press, 2018), 167–184, <https://www.jstor.org/stable/10.5406/j.ctv8bt13m.16>; Kate Holterhoff, "From Disclaimer to Critique: Race and the Digital Image Archivist," *Digital Humanities Quarterly* 011, no. 3 (August 28, 2017); and Marianna Hovhannisyan, "Speculative Classification: Tracing a Disputed Portrait between the Archives of Malvina Hoffman and Sergey Merkurov," *InterActions: UCLA Journal of Education and Information Studies* 15, no. 2 (2019), <https://escholarship.org/uc/item/9dq9c7s9>.

[43] Johanna Drucker, "Classification Systems," in *Introduction to Digital Humanities Course Book*, by Johanna Drucker et al. (Los Angeles: UCLA Center for Digital Humanities, 2014), 20–23, [http://dh101.humanities.ucla.edu/wp-content/uploads/2014/09/IntroductionToDigitalHumanities\\_Textbook.pdf](http://dh101.humanities.ucla.edu/wp-content/uploads/2014/09/IntroductionToDigitalHumanities_Textbook.pdf).

[44] Hope A. Olson, "How We Construct Subjects: A Feminist Analysis," *Library Trends* 56, no. 2 (Fall 2007): 509–541.

[45] *Ibid.*, 522.

[46] Donald Fixico, *The American Indian Mind in a Linear World: American Indian Studies and Traditional*

*Knowledge* (Abingdon, UK: Routledge Taylor & Francis Group, 2003), 1.

[47] Linda Tuhiwai Smith, *Decolonizing Methodologies: Research and Indigenous Peoples* (London: Zed Books, 1999), 25, 42–43.

[48] The Getty, “Pacific Standard Time at the Getty,” Pacific Standard Time, accessed March 29, 2020, <http://pacificstandardtime.org>.

[49] See, for instance, the new digital scholarships produced by the GRI here: “Digital Art History (Getty Research Institute),” The Getty Research Institute, accessed May 20, 2020, [https://www.getty.edu/research/scholars/digital\\_art\\_history/index.html](https://www.getty.edu/research/scholars/digital_art_history/index.html).

[50] “Home | Collections Online - Museum of New Zealand Te Papa Tongarewa,” Museum of New Zealand Te Papa Tongarewa, accessed March 29, 2020, <https://collections.tepapa.govt.nz/>.

[51] Sean Mallon, “Agency and Authority: The Politics of Co-Collecting,” in *Curatopia*, eds. Philipp Schorch and Conal McCarthy, 1st ed., Museums and the Future of Curatorship (Manchester University Press, 2019), 279–295, <https://www.jstor.org/stable/j.ctvnb7j6j.24>.

[52] “About Collections Online,” Museum of New Zealand Te Papa Tongarewa, Wellington, NZ, August 16, 2017, <https://www.tepapa.govt.nz/discover-collections/collections-online/about-collections-online>.

[53] Fiona Moorhead, “Te Papa’s New Collections Site Launches – Get Hunting,” *Te Papa’s Blog* (blog), September 14, 2017, <https://blog.tepapa.govt.nz/2017/09/15/te-papas-new-collections-site-launches-get-hunting/>.

[54] Johanna Drucker, “Ontologies and Metadata,” in *Introduction to Digital Humanities Course Book*, Johanna Drucker et al. (Los Angeles: UCLA Center for Digital Humanities, 2014), 24, [http://dh101.humanities.ucla.edu/wp-content/uploads/2014/09/IntroductionToDigitalHumanities\\_Textbook.pdf](http://dh101.humanities.ucla.edu/wp-content/uploads/2014/09/IntroductionToDigitalHumanities_Textbook.pdf).

[55] Fiormonte, “Towards a Cultural Critique of the Digital Humanities,” 72–74.

[56] Jacques Derrida, *Archive Fever: A Freudian Impression* (Chicago: University of Chicago Press, 1996), quoted in Fiormonte, “Towards a Cultural Critique of the Digital Humanities,” 74.

[57] Julia Jacobs, “Getty Trust to Invest \$100 Million in Saving Threatened Antiquities,” *The New York Times*, September 17, 2019, sec. Arts, <https://www.nytimes.com/2019/09/17/arts/design/getty-trust-james-cuno.html>.