Ashurbanipal: A Diristry to Document Multimedia Metadata Tools for Transdisciplinary Archives

1st Anousha Athreya
Brain Health Alliance
Ladera Ranch, CA, USA
aathreya@bhavi.us

2nd S. Koby Taswell
Brain Health Alliance
Ladera Ranch, CA, USA
ktaswell@bhavi.us

3rd Andrew Kang
Brain Health Alliance
Ladera Ranch, CA, USA
akang@bhavi.us

4th Ishani Das
Brain Health Alliance
Ladera Ranch, CA, USA
idas@bhavi.us

4th Carl Taswell
Brain Health Alliance
Ladera Ranch, CA, USA
ctaswell@bhavi.us

Abstract—In historical artifact conservation, archiving objects using entity metadata plays a significant role in managing the related versions of the artifacts preserved, recorded and documented at various time points. In this paper, we discuss five fields of study to display the importance of related versions in identifying patterns over time through historical events, cultural heritage, brain health, performing arts, and fine arts. We describe our use of the *Ashurbanipal* diristry to document scholarly research on archiving tools and technologies. We highlight the importance of the provenance infosubset in tracing metadata for cultural objects managed in NPDS repositories and enabling interoperability with existing multimedia bibliographic formats including MARC and BIBFRAME.

Index Terms—PORTAL-DOORS Project, Ashurbanipal diristry, equivalent entities, related versions, provenance, archiving, artifact conservation.

I. INTRODUCTION

As cities continue to develop on par with humankind, our definition of innovation has evolved alongside, bringing rise to new iterations of infrastructure in systems. While containing different meanings in various contexts, infrastructure as an underlying framework takes on a larger term when contextualizing it to the transdisciplinary transfer of knowledge. Buetow *et al.* [1] describe database infrastructure in terms of biomedical research, indicating that cyberinfrastructure has the potential to bridge interdisciplinary communities in various biomedicine research disciplines and bring the innovation occurring at the level of an individual researcher to a 'virtual team science' with the cross-disciplinary transfer of knowledge of tools and technologies. Similarly, infrastructure can range from database infrastructure for storing metadata to civil engineering infrastructure and architecture. Without clear and informed design principles to combat the disintegration of the European sewage system during the High Middle Ages, city life fell into disarray with open ditches and pathways that halted travel across roadways. Following the plague, Joseph Bazalgette, an English civil engineer, proposed developing a system of collectors on either side of London to drain sewage that overflowed from the River Thames [2]. The creation of such infrastructure, civil and biomedical, can be attributed to the design principles that inform its use and meaning. Design principles are subject to iterations, changes, or improvements that eventually create different versions of the same entity, or the form or variant of a type or original [3].

Related versions of an object must be compared within the context of the problem-oriented domain-specific field to determine equivalence, if present, through the threshold of similar attributes. In such instances, provenance in any domain-specific field can be used to determine how a specific concept was changed, influenced, adapted, or simply transferred between its related versions. Comparison of related versions can be depicted with the Notre Dame cathedral as example. It was rebuilt numerous times during 1504-1560, and more recently, after the devastating fire in 2019 [4]. While the cathedral will be rebuilt in the same physical location, will its defining features and mediums remain the same? How does the new version of the cathedral compare to each of its ten iterations and the original design? Such questions need to be resolved to determine the threshold of relatedness for an iteration of the Notre Dame to the original. In context of fine or performing arts, artistic techniques and dances can be adapted by two different directors, different cultures and communities, or even between two artists.

In this paper, we introduce the Ashurbanipal diristry. The goal of the diristry recalls its namesake Ashurbanipal, the king of the Assyrian empire from 669-631 BC, who built one of the greatest libraries with the intent of collecting every book in the world. The diristry serves as an archive of methods and formats used by librarians, archivists, and museum curators for data and artifact conservation. Technologies may include multimedia formats such as MARC and BIBFRAME used by the Library of Congress for archiving cultural objects. There is substantial information available about evolving technologies resulting from efforts to conserve and/or recover newly emerging archaeological discoveries. We find it pertinent that a transdisciplinary archive of this information would be a necessity for the preservation of these tools.

II. IDENTIFYING PATTERNS OVER TIME

Each domain-specific field uses its own methods of tracking related versions of historical artifacts. We explore five fields of study that exemplify identifying trends in history with great diversity and various use cases: historical events, cultural heritage, brain health, performing arts, and fine arts.

Historical Events Historians identify historical events as either time points to track key trends among centuries or spatial data (across countries, cities, regions) to identify economic, political, and social trends [5]. While history is often tracked and recorded based on the sources and evidence of the era, when there is destruction of regional evidence or none at all, certain important persons and events are lost in history. Hatchepsut is known as the Lost Pharaoh of Egypt and is regarded as one of the most successful pharaohs of the indigenous Egyptian dynasty. Following Hatchepsut's death, her figure was chiseled off statues and stone walls, obelisks recording her existence were smashed, and paper documents listing her name were destroyed. Following a hypothesis of the "Hatshepsut Problem" and observed misalignment between the pronouns and ruler at the time, her name was reinstated into history [6]. The rebuilding of the Gresham Exchange in London (The Royal Exchange) also falls under versioned examples of historical events. The building showed economic and cultural significance beyond just England, but also throughout Eurasia until it was destroyed by the Great Fire of 1666. While undergoing major architectural changes, the Royal Exchange has maintained its original goal of facilitating trade [7].

Cultural Heritage In cultural heritage, information archives are developed in museums and libraries to preserve objects for cultural heritage over time for 'information literacy', e.g. Library of Congress, National Archives and Records Administration, Franklin D. Roosevelt Presidential Library and Museum [8]. Heritage conservation means individual objects are assigned heritage value to be preserved to represent social life and identities of groups involved as well as a collection of primary, secondary, and tertiary sources of an event [9]. With the decline of the Mayan civilization, conservation discourse has centered around creating a timeline from events revealed by archaeological discoveries. Archeologist Takeshi Inomata investigated thousands of structures constructed by Mayan defense from Ceibal, Guatemala, a radiocarbon center that allows archaeologists to chronologically track population size of the Mayan civilization [10].

Such ideas of cultural heritage can even be extended to surnames. European surname laws have stated the child will inherit the fathers last name unless their mother's is preferred [11]. Through history, surnames have been created personalized to each individual, taking into account inheritance, culture, and societal group. However, surnames have evolved over time through cultural and historical changes. Often out of safety, individuals from persecuted ethnic groups have changed their surnames when fleeing to another countries to disguise their identities. During WWII for example, the Jewish folk that

Figure 1. Comparing Sequential Versions

escaped some European countries changed their names under political and social pressure to avoid persecution [12].

Brain Health In brain health, serial studies are often used in context of medical imaging scans taken at different time points in a patient's lifetime. Especially in neurodegenerative diseases, serial scans are crucial to monitor and track the progression of the disease. In Alzheimer's disease, it is not uncommon to review imaging markers for neurodegeneration during a serial imaging study of a patient. While neurobiological changes to a patient's brain occur years before symptoms appear, progressive brain atrophy wears down the brain reserve until the amnestic symptoms of Alzheimer's develop [13]. Serial scans could help identify and correlate with different thresholds of memory loss in Alzheimer's patients from simple memory disturbance to dementia. [14]. Provenance in neuroimaging centers on discovering the origins of a specific entity, and also, addressing the data quality of an image for interoperability, reproducibility, and comparison. Data analysis using provenance can perform complex analyses to gain information from large cross-sectional studies in multiple sclerosis research Aspects of data provenance can be gathered from the site of data collection, the headers of the image files, and/or the databases that have recorded the image history [15].

Performing Arts Within the realm of performing arts, two pieces can be compared at the single point in time in the piece although played in different locations or between differences in casting, a performer, composer, or even alterations to a plotline. In various bands or singers, a live performance of a song in a certain location can be compared to another live performance at a different location versus the studio recording [16]. Although most musicians put new spins on their own music or cover other's music, not many songs have been rehashed as much as the United States national anthem entitled the "The Star-Spangled Banner". It's been

performed by some of America's most famous artists from Beyonce and Whitney Houston to Aretha Franklin and Jimi Hendrix, each time undergoing some amount of evolution between performances [17]. Variations between performances vary and often depend on the musician's style with some musicians utilizing their extremely large vocal range, some performing something more solemn and traditional, and others putting their own spin on it, by introducing an old American classic to new styles of music whether Rock, Soul, or R&B. These different performances by different artists all build off the same source material however each bring something new to the table. Specific to the genre of jazz, we see musicians writing 'standards' which give a melodic idea and set of chord changes for themselves and others to improvise over and never give the exact same performance twice [18]. Variation on another's work or one's own music is not something new to modern music of the past few decades, nor is it restricted to purely musical performances. Figure 1 references *Annie: The Musical* as a sufficient use-case in comparing four versions of an originally written musical by composers Charles Strouse, Martin Charnin, and Thomas Meehan in 1977 who later on created related versions of the play. This can be compared to an original version of *Annie* based on the comics written by Harold Gray. In each version, they maintain similar plotlines and cast different versions of the characters in the musicals.

Fine Arts Fine Arts is used to represent aesthetics and beauty relating to works of poetry, sculpture, or various types of paintings and techniques. Many versions of paintings are displayed, either through alteration from restoration, or versions created by the original artists themselves. A single artist may either iterate through different steps in the painting process on different canvases or paint different iterations of the same painting on the same canvas [19]. Leonardo Da Vinci's *The Virgin of the Rocks* has two different paintings to its name with key differences in lighting, color, and the techniques used to depict grey smoke in the painting while maintaining the same theme and character position. The painting *Supper at Emmaus* by Caravaggio in 1601 has two versions of the painting displayed in The National Gallery in London and Brera Gallery in Milan. The differences between the two paintings are evident, with different shading techniques, various subjects in the paintings, and the somber depiction of Christ's Supper in the second version [20]. Figure 1 references *La Taureau* by Pablo Picasso, which was a series of paintings he completed to display the aesthetic of a bull, beginning from a minimalistic design to the detailed drawing and displays the iterative process he undertook to create related versions of the painting.

III. Ashurbanipal Diristry

The PORTAL-DOORS Project (PDP) has developed the Nexus-PORTAL-DOORS-Scribe (NPDS) Cyberinfrastructure with lexical PORTAL registries, semantic DOORS directories, hybrid Nexus diristries, and Scribe registrars as a hierarchically distributed metadata management system [21], [22]. With the implementation of measures to track key fields of cultural heritage objects and historical events over time, the NPDS Cyberinfrastructure can be described as an *archiving system* that stores metadata relating to the provenance of scholarly research. Two main bibliographic formats used by the Library of Congress include MARC (Machine Readable Cataloguing) and BIBFRAME, in a more modern transition to semantic web technologies and multimedia. MARC21 contains fields used to record various types of provenance-related data such as location, publication, or performances that can be mapped to various metadata formats, including BIBFRAME and RIS. The existing BIBFRAME format organizes catalogued information into levels of abstraction: work, instance, and item [23].

In order to enhance an extensible version of NPDS to accommodate related versions of cultural objects, we have continued development of the NPDS Cyberinfrastructure to include import forms for mapping to the NPDS format, based on which initial import format is preferred by curators accessing PDP utilities [21]. When discussing the resource metadata of a cultural artifact, one might track from the point of its discovery, including the item name, a description of its past use and origins, and the provenance of the artifact [24]. We have regcognized the importance for NPDS diristries to manage provenance as an infosubset that is interoperable with the MARC and BIBFRAME standard currently used by librarians, archivists, and museums.

As of now, the Ashurbanipal diristry contains 83 high-quality records with references to various methods of preservation and curation of cultural artifacts, from storing information about virtual restoration systems for broken pottery to a comparison of composition styles and trends between centuries. We have demonstrated several use-cases of how NPDS will handle provenance within five domain-specific fields in the Ashurbanipal diristry; these examples can be found at the portaldoors.net website as records in the Ashurbanipal diristry. For implementation of the provenance format, we have demonstrated explicit conceptual and templated examples in the fields of art history, music, performing arts, historical events, and cultural heritage, and plan to curate others relating to provenance in architecture and computer science principles. Each templated example consists of a demonstrable example as a conceptual version of provenance in plain text, a version of provenance demonstrated in an XML schema which is our preferred format for future implementation in PDP, and a syntactic format variant in which we demonstrate provenance in some of the formats currently in use by librarians, archivists, and museum curators.

Development of the *Ashurbanipal* diristry informs an archiving system for documenting tools and technologies for the preservation and restoration of cultural objects. This includes a range of existing scholarly literature for library technologies and cataloguing techniques. Entries could include a scholarly research article discussing techniques in painting restoration, existing methods of denoising an image in photography, or techniques to restore old film. We intend to use this diristry to demonstrate the use of NPDS in tracking and logging scholarly literature relating to evolving tools, tech-

nologies, and multimedia formats used by archivists for artifact examination. The diristry has been established to input any relevant literature, organizations relating to developing archival methods, or existing techniques regarding the preservation of cultural heritage.

IV. CONCLUSION

We introduce the Ashurbanipal diristry as an archive to document methods and technologies intended for use by trans-disciplinary multimedia librarians, archivists, and museum curators for cataloguing and conserving artifacts. We discuss versioned examples in the context of documenting cultural heritage artifacts, particularly with tracking and managing the evolution of such artifacts through defined points in time. We describe several examples of related versions of artifacts in the context of several problem-oriented domain-specific fields: historical events, cultural heritage, brain health, performing arts, and the fine arts. Within this framework, we consider provenance a key factor in determining the relationships between two things as same or different, and relevant to answering the question of whether and when is a cultural artifact a separate unrelated entity, a derived entity, equivalent but distinct entity, or a related version of the same original entity?

REFERENCES

[1] K. H. Buetow, "Cyberinfrastructure: Empowering a" third way" in biomedical research," *Science*, vol. 308, no. 5723, pp. 821–824, 2005.

[2] G. De Feo, G. Antoniou, H. F. Fardin, *et al.*, "The historical development of sewers worldwide," *Sustainability*, vol. 6, no. 6, pp. 3936–3974, 2014.

[3] S. Gançarski and G. Jomier, "Managing entity versions within their contexts: A formal approach," in *International Conference on Database and Expert Systems Applications*, Springer, 1994, pp. 400–409.

[4] A. Sawkins, "Architecture, politics and the rebuilding of the cathedral of notre-dame at senlis, 1504-1560," 1998.

[5] P. D. Allison, *Event history analysis: Regression for longitudinal event data*, 46. Sage Publishers, 1984.

[6] C. H. Roehrig, R. Dreyfus, and C. A. Keller, *Hatshepsut: from queen to pharaoh*. Metropolitan Museum of art, 2005.

[7] G. Massiot and cie, *Royal Exchange, London: View of front facade with pediment sculpture by Westmacott*. Jan. 1910.

[8] K. Baker, *Information literacy and cultural heritage: developing a model for lifelong learning*. Elsevier, 2013.

[9] K. V. Korostelina, "Understanding values of cultural heritage within the framework of social identity conflicts," *Values in Heritage Management: Emerging Approaches and Research Directions*, p. 83, 2019.

[10] T. Inomata, D. Triadan, J. MacLellan, *et al.*, "High-precision radiocarbon dating of political collapse and dynastic origins at the maya site of ceibal, guatemala," *Proceedings of the National Academy of Sciences*, vol. 114, no. 6, pp. 1293–1298, 2017.

[11] C. Scapoli, E. Mamolini, A. Carrieri, *et al.*, "Surnames in western europe: A comparison of the subcontinental populations through isonymy," *Theoretical Population Biology*, vol. 71, no. 1, pp. 37–48, 2007.

[12] K. Fermaglich, "what's uncle sam's last name? jews and name changing in new york city during the world war ii era," *The Journal of American History*, vol. 102, no. 3, pp. 719–745, 2015.

[13] G. B. Frisoni, N. C. Fox, C. R. Jack, *et al.*, "The clinical use of structural mri in alzheimer disease," *Nature Reviews Neurology*, vol. 6, no. 2, pp. 67–77, 2010.

[14] P. Lin, D. Correa, T. J. Kean, *et al.*, "Serial transplantation and long-term engraftment of intra-arterially delivered clonally derived mesenchymal stem cells to injured bone marrow," *Molecular Therapy*, vol. 22, no. 1, pp. 160–168, 2014.

[15] A. J. MacKenzie-Graham, J. D. Van Horn, R. P. Woods, *et al.*, "Provenance in neuroimaging," *Neuroimage*, vol. 42, no. 1, pp. 178–195, 2008.

[16] S. Wang, S. Ewert, and S. Dixon, "Robust joint alignment of multiple versions of a piece of music," *arXiv preprint arXiv:1604.08516*, 2016.

[17] R. Stone, *The most memorable super bowl national anthems*, 2021.

[18] T. Gioia, *The jazz standards: A guide to the repertoire*. Oxford University Press, 2012.

[19] M. Orozco, "Catalogue raisonné of picasso posters," 2019.

[20] C. Scribner III, "In alia effigie: Caravaggio's london supper at emmaus," *The Art Bulletin*, vol. 59, no. 3, pp. 375–382, 1977.

[21] C. Taswell, "DOORS to the semantic web and grid with a PORTAL for biomedical computing," eng, *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 2, pp. 191–204, 2 Mar. 2008, In the Special Section on Bio-Grid published online 3 Aug. 2007, ISSN: 1089-7771. DOI: 10.1109/TITB.2007.905861.

[22] S. Choksi, P. Hong, S. Mashkoor, *et al.*, "NPDSLINKS: Nexus-PORTAL-DOORS-Scribe Learning Intelligence aNd Knowledge System," DOI: 10.1109/TransAI49837.2020.00027.

[23] A. Xu, K. Hess, and L. Akerman, "From marc to bibframe 2.0: Crosswalks," *Cataloging & Classification Quarterly*, vol. 56, no. 2-3, pp. 224–250, 2018.

[24] M. Patel, M. White, N. Mourkoussis, *et al.*, "Metadata requirements for digital museum environments," *International Journal on Digital Libraries*, vol. 5, no. 3, pp. 179–192, 2005.