Enhancing Provenance Research with Linked Data: A Visual Approach to Knowledge Discovery

Sarah Binta Alam Shoilee[1,*], Annastiina Ahola[2], Heikki Rantala[2], Eero Hyvönen[2,3], Victor de Boer[1], Jacco van Ossenbruggen[1] and Susan Legene[1]

[1]*Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, the Netherlands*
[2]*Aalto University, Semantic Computing Research Group (SeCo), Konemiehentie 2, 02150 Espoo, Finland*
[3]*University of Helsinki, Helsinki Centre for Digital Humanities (HELDIG), Unioninkatu 40, 00170 Helsinki, Finland*

Abstract

Provenance research is critical for understanding the historical trajectories of cultural objects housed in museums, yet it is often hindered by fragmented, ambiguous, or missing data. With the increasing adoption of Linked Data (LD) in cultural heritage, new possibilities emerge for analysing provenance metadata. This paper presents the PM-Sampo demonstrator, a structured approach to analysing provenance data through Linked Data methodologies and visualisation techniques. By connecting historical events, places, and actors to object collections and analysing data with visualisation tools, PM-Sampo aims to facilitate large-scale provenance analysis, enabling domain researchers to detect patterns, inconsistencies, and hidden connections that could otherwise go unnoticed. A case study on objects from Dutch museums associated with the Aceh War (1873–1914), an armed conflict between the Netherlands and the Muslim sultanate of Aceh, illustrates the functionalities of the demonstrator, revealing gaps in acquisition records, unexpected geographical distributions, and acquisition timelines extending well beyond the formal end of the conflict. The establishment of actor-connections further brings to the surface overlooked relationships between individuals and institutions, while provenance visualisation highlights the need for more comprehensive provenance documentation by domain experts. The study underscores the opportunities of data-driven approaches in provenance research, demonstrating how visualisation tools can aid in knowledge discovery and exploring knowledge gaps.

Keywords

Linked Data, Visualisation, Provenance Research, Cultural Heritage, Knowledge Discovery

1. Introduction

Provenance research is a critical discipline in cultural heritage studies, focussing on the trace of origins, ownership history, and movement of objects across time and space [1]. It provides essential insights into ethical, legal, and historical contexts of collections, particularly in cases of contested ownership, colonial acquisitions, and restitution claims [2]. By reconstructing an object's past, provenance research helps museums, scholars, and policymakers make informed decisions about collections, ensuring transparency, accountability, and historical justice. Beyond ethical and legal considerations, provenance research plays a crucial role in deepening our understanding of the historical, social, and economic forces that shaped the circulation of cultural objects.

However, provenance research is inherently complex due to the fragmented nature of historical records, inconsistent documentation practices, and the large scale of data sources involved. Provenance information is often dispersed across archival materials, museum databases, and historical texts, making it a time-consuming job to consolidate and analyse systematically. With the increasing adoption of Linked Data (LD) in cultural heritage, new possibilities are emerging for structuring, connecting, and

SEMDH'25: Second International Workshop of Semantic Digital Humanities, June 01–02, 2025, Portoroz, Slovenia
*Corresponding author.
s.b.a.shoilee@vu.nl (S. B. A. Shoilee); annastiina.ahola@aalto.fi (A. Ahola); heikki.rantala@aalto.fi (H. Rantala); eero.hyvonen@aalto.fi (E. Hyvönen); v.de.boer@vu.nl (V. d. Boer); jacco.van.ossenbruggen@vu.nl (J. v. Ossenbruggen); s.legene@vu.nl (S. Legene)
0000-0001-9458-8105 (S. B. A. Shoilee); 0009-0008-6369-4712 (A. Ahola); 0000-0002-4716-6564 (H. Rantala); 0000-0003-1695-5840 (E. Hyvönen); 0000-0001-9079-039X (V. d. Boer); 0000-0002-7748-4715 (J. v. Ossenbruggen); 0000-0002-2826-9541 (S. Legene)

analysing provenance metadata at scale. LD enables the integration of heterogeneous datasets, linking historical events, places, and actors to object collections across multiple institutions. This interconnected approach facilitates a more comprehensive analysis of provenance information, allowing researchers to detect patterns, inconsistencies, and previously overlooked connections. This research aims to address some of the provenance research challenges through Linked Data visualisation techniques, offering interactive tools, such as timeline analysis, redirect links, geospatial mapping, and faceted search to untangle the intricate relationships between objects, collectors, places, events, and time periods.

This current research is conducted within the framework of Pressing Matter[1], a project investigating the ownership, value, and historical significance of colonial heritage in museums. In this context, provenance research is approached through Actor-Network Theory (ANT) [3], which conceptualises objects as relational entities interconnected with people, places, and events that shape their histories. In a broader context, this research contributes to ongoing efforts to establish a framework for dealing with colonial heritage in museums, moving beyond traditional notions of ownership to examine broader themes of exchange, violence, evidence, and historical accountability.

This paper introduces the PM-Sampo demonstrator, a structured approach that leverages Linked Data methodologies and visualisation techniques to support large-scale provenance analysis. The aim of PM-Sampo is to enable researchers in the domain to explore the relationships between objects, collectors, institutions, and historical contexts in a more systematic and interactive manner. The demonstrator provides tools such as geospatial mapping, timeline analysis, and connection visualisations to enhance provenance interpretation and knowledge discovery. To illustrate its functionalities, this study applies PM-Sampo to a case study on objects from Wereldmuseum[2] associated with the Aceh War (1873–-1914), a significant colonial conflict between armed Dutch military power and the Sultanate of Aceh, Indonesia.

Concretely, this paper contributes: (1) a structured approach for querying provenance information across people, places, events, time, and objects through faceted search, (2) a demonstration of visualisation interfaces, such as geospatial-object networks, acquisition timelines, and connection visualisation among entities in provenance records, and (3) a Linked Data-driven web application to enhance object provenance reconstruction and to aid knowledge discovery. This research aims to support museums, researchers, and policy makers in addressing the complexities of provenance studies.

The remainder of this paper is structured as follows: Section 2 reviews related work, Section 3 describes the dataset used, Section 4 details the implementation of PM-Sampo, and Section 5 explores its functionality in provenance research. Finally, Section 6 concludes the paper.

2. Related Work

Knowledge Discovery for provenance research intersects with multiple fields, including Knowledge Discovery in Databases (KDD), Digital Humanities, and Computational Cultural Heritage. The concept of KDD, defined by Fayyad et al., is a process of identifying valid, novel, and potentially useful patterns in data through data mining techniques such as classification, clustering, and association rule mining [4]. In the context of Linked Data, knowledge discovery can be approached through graph data mining, where data patterns are explained through interconnected links [5]. Another relevant approach involves utilising background knowledge to navigate networks and uncover new information [6]. Both methods are valuable for digital humanities, where computational techniques can complement expert knowledge to facilitate discovery and interpretation in provenance research.

Heritage object metadata is inherently complex, making knowledge graphs (KGs) a suitable representational framework for structuring and analysing these relationships. However, it remains an open challenge to assess the extent to which relational learning models can support knowledge discovery in cultural heritage [7]. While humanities research has traditionally been skeptical of automated and quantitative approaches [8], recent studies highlight the benefits of data-driven methodologies in expanding research perspectives and scalability [9].

[1]Pressing Matter project homepage: https://pressingmatter.nl

[2]Wereldmuseum is a Dutch museum where a large portion of their housed collections has embedded colonial past.

The increasing availability of cultural heritage datasets from museums, libraries and archives has fuelled the integration of computational approaches within digital humanities [10, 11]. Big Data methodologies now allow for large-scale analysis of historical records, shifting digital humanities research beyond simple data visualisation to more autonomous problem solving tools [7]. In the context of provenance research, relational search (RS) or uncovering meaningful semantic associations between entities could be a promising approach. RS has been applied in diverse domains, including national security [12], medical research [13], and cultural heritage studies [14, 15]. By employing RS techniques, provenance research can move beyond static metadata exploration to reveal dynamic relationships among objects, people, and events.

In existing work, the Sampo model [16] provides a Linked Data-driven approach to cultural heritage research. The Sampo-UI framework [17, 18] offers a semantic portal framework that allows easy integration of heterogeneous datasets, enabling advanced search, visualisation, and analysis across interconnected entities. Previous implementations, such as WarSampo (for World War II data), BiographySampo (for biographical data), and MMM Sampo (for mapping manuscript migrations), demonstrate its applicability in structuring and exploring cultural heritage data through knowledge graphs [16, 15]. By employing semantic technologies, Sampo model enables relational searches, facilitating the discovery of hidden connections between people, places, and events. Current research is based on the Sampo model, adapting its principles for provenance analysis to uncover patterns in object histories, acquisition networks, and collector relationships. We here present a Linked Data-driven demonstrator, PM-Sampo, aimed at uncovering the various ways that heritage objects could potentially be associated with specific historical events, places, and times through their collectors' acquisition patterns.

3. Data Description

The PM-Sampo demonstrator primarily utilises object collection data from the Wereldmuseum [3], published through the Colonial Collections Data Hub [4] initiative. The Wereldmuseum is a conglomerate comprising Wereldmuseum Leiden, Wereldmuseum Amsterdam, and Wereldmuseum Rotterdam. A large part of the collections housed within these museums originate from colonial-era commissions by the Dutch Ministry of Colonies and scientific expeditions to former Dutch colonial territories. Many of these objects were acquired as part of the larger Colonial Institute in the late 19th and early 20th centuries.

The Colonial Collections Data Hub previously published Wereldmuseum datasets in Linked Data format along with the publication of the SPARQL endpoint[5], making object metadata more accessible and structured based on a shared domain ontology, that is, CIDOC-CRM ontologies [19] and following the Linked Art recommendations[6]. The dataset published through the portal originates from the Wereldmuseum's internal collection management systems and has been refined for structured analysis for provenance research. It captures diverse aspects of provenance, historical context, and object metadata through multiple structured graphs.

The PM-Sampo demonstrator builds upon this openly available dataset, simplifying and enriching it where necessary to improve usability for large-scale provenance analysis. To integrate existing dataset into the PM-Sampo demonstrator and to make query time faster, lightweight ontologies ("facet ontologies") were developed to shorten and simplify paths of key concepts, such as acquisition records, to historical events and to their geographic locations[7]. Furthermore, data enrichment processes were applied, i.e., GeoNames data extraction, to enhance geospatial mapping by associating latitude and longitude coordinates with relevant production places. The new version of the data can be accessed through the SPARQL endpoint: http://ldf.fi/pm-sampo/sparql

[3]Wereldmuseum webpages: https://\{amsterdam/leiden.rotterdam\}.wereldmuseum.nl
[4]Colonial Collection Hub data portal: https://data.colonialcollections.nl
[5]Wereldmuseum Linked Data endpoint: https://api.colonialcollections.nl/datasets/nmvw/collection-archives/sparql
[6]Linkedart webpage: https://linked.art
[7]The data conversion process is documented in the GitHub repository along with the current schema: https://github.com/Shoilee/PM-SampoDataManager

4. PM-Sampo Implementation

The data service is separated completely from the PM-Sampo demonstrator; only the external SPARQL endpoint is used to access the data within the portal. The development of PM-Sampo semantic portal is supported by the Sampo-UI framework [17, 18]. The Sampo-UI is a framework that offers software developers a starting base to build a JavaScript web application, which can be customised with minimal effort to create LD applications. On the technical side, the framework consists of two main components: (1) a client-side interface built using the well-established React[8] and Redux[9] libraries and (2) a Node.js[10] back-end developed with the Express[11] framework.

To extract the benefit from the Sampo-UI framework, developing of PM-Sampo demonstrator was started from an existing portal demonstration[12] whose configurations are then modified declaratively to meet the requirements of provenance analysis and the data models of the current dataset. In our case, a first small demo that provided a search and browsing interface to inspect the data was actually implemented during the first project meeting. Subsequent improvement and alteration have been made to meet the requirements of the provenance researchers obtained from previous research [20]. In the PM-Sampo application, the user first lands on the *landing page* with several *application perspectives* to the data. Perspectives are based on target entity classes of the underlying knowledge graph, i.e., Objects, Provenance Events, Historical Events, and Actors. The PM-Sampo demonstrator is available on the Web at https://pmsampo.demo.seco.cs.aalto.fi/en/ and the source code for the demonstrator was published on GitHub: https://github.com/Shoilee/PM-Sampo/releases/tag/v1.0

Following the Sampo model principles, the usage cycle of each perspective can be divided into two steps: 1) filter instances of the class(es) corresponding to the perspective and 2) create different visualisations to analyse the result instances. The data is filtered using the faceted semantic search [21] tools provided by the portal where the properties of the perspective class are used as facets. The results and facet options including hit counts are updated after each selection of a facet, making it possible for the user to precisely filter the end-result entities by different properties. The hit counts help the end user to direct the search towards promising facet selections and prevent the user from ending up in dead ends with no results (hits). After filtering the data to a desired subset, the user can analyse the result set, i.e., a set of instances of the class corresponding to the application perspective, with integrated data-analytic tools available as tabs on the application perspective page. In the same way, data-analytic tabs can be integrated with instance pages that aggregate information about the individual entities of the application perspective.

A key enhancement in PM-Sampo was an advanced faceted search that extended filtering beyond direct entity properties to include other related entities. For example, objects could be filtered based not just on object attributes but also on associated historical events and actors involved in acquisitions, offering a more context-aware exploration of provenance data. In addition to that, targeted data analytics visualisations were integrated, including a faceted search results table, a summarisation pie chart on facet filters, a production places map for the object perspective, a provenance tab for object instances listing provenance events with the aim of creating an object biography, and a provenance events timeline for chronological insights.

Another notable addition was the *Related Tab*, which inferred new relationships not explicitly defined in the knowledge graph. Two key new connections were introduced: (1) actor-to-actor links through shared objects and (2) actor-to-historical event links, established through SPARQL queries when actors contributed to object acquisition and objects are associated with historical events. These relationships are visualised through lists and explained through intermediary objects.

[8]React webpage: https://reactjs.org
[9]Redux webpage: https://redux.js.org
[10]Node.js webpage: https://nodejs.org/en
[11]Express webpage: https://expressjs.com
[12]Sampo-UI information page: https://seco.cs.aalto.fi/tools/sampo-ui/

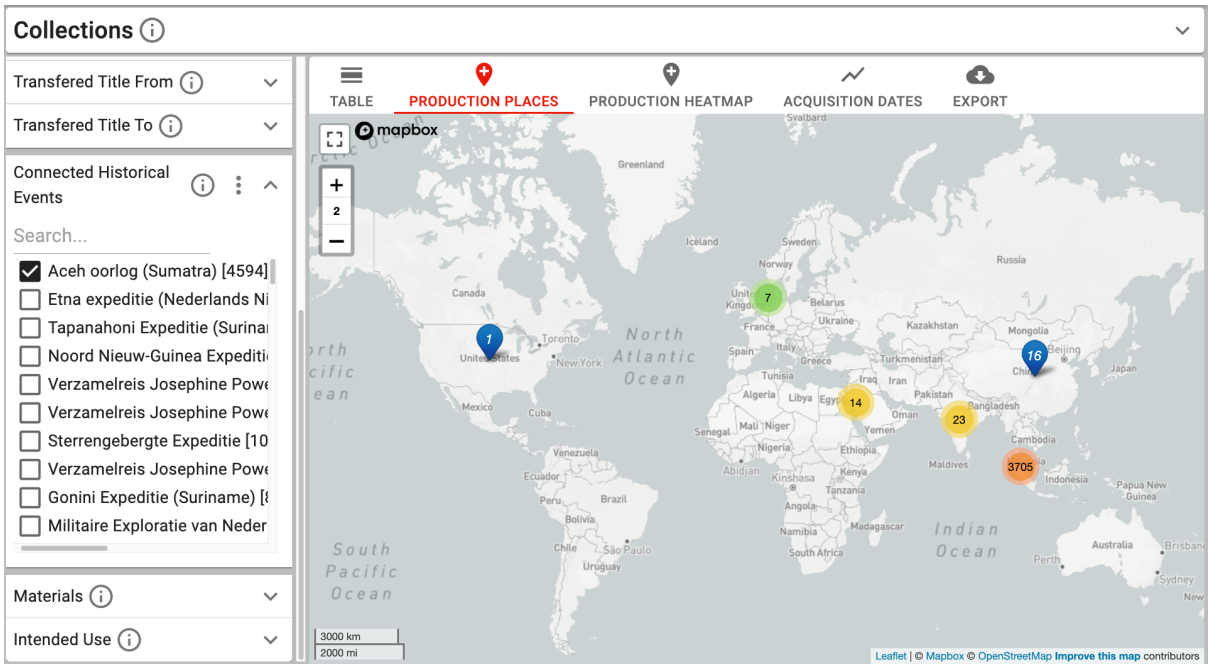


Figure 1: A place visualisation of all objects which has connection with Aceh Oorlog (Sumatra).

5. Use-cases

PM-Sampo demonstrator provides a structured way to analyse provenance data by linking and visualising historical events, places, and actors to object collections. Objects related to the historical event "Aceh Oorlog (Sumatra)" (Aceh War 1873–1914) serve as an ideal case for demonstrating the various functions of PM-Sampo. This subset of data has been thoroughly studied, containing high-quality metadata that aids in identifying connections with the war in Aceh, Indonesia in the 19th and 20th centuries.

5.1. Geographic Distribution of Objects

One of the initial observations through PM-Sampo demonstrator is the geographic spread of objects related to the Aceh War, which is visualised in Figure 1. Although the war itself occurred in Aceh, the dataset shows that the objects' production places are from diverse locations, including India and the United States. A visual representation using a place-object mapping function communicates these interesting connections, highlighting the possible circulation of objects produced beyond Aceh. This challenges common assumptions that the war loot is localised to the geographical location and emphasises the necessity for investigation into object migration patterns.

5.2. Temporal Analysis: Acquisition Timeline

Although the Aceh War formally ended in 1914, the PM-Sampo demonstrator reveals that object collection activities associated with this historical event extended until 2010. This finding raises significant questions about post-war circulation and reassessment of war-related artifacts. By visualising an acquisition timeline (in Figure 2), PM-Sampo allows researchers to observe trends in collection histories, revealing inconsistencies with common assumption and continuities in provenance data.

Notably, the timeline highlights significant spikes in acquisition events in the years 1907 and 1959. A closer examination of the data through the portal reveals that in 1907, a substantial proportion of these acquisitions —550 out of 564 objects (97%)— originated from Theodorus Jacobus Veltman. This finding aligns with archival records indicating that in 1907, Veltman, a former soldier in the Dutch colonial forces, sold 753 objects, primarily from Indonesia, to the Museum Volkenkunde (the predecessor of Wereldmuseum). The provenance of these objects raises further questions: while 550 of the 753 objects

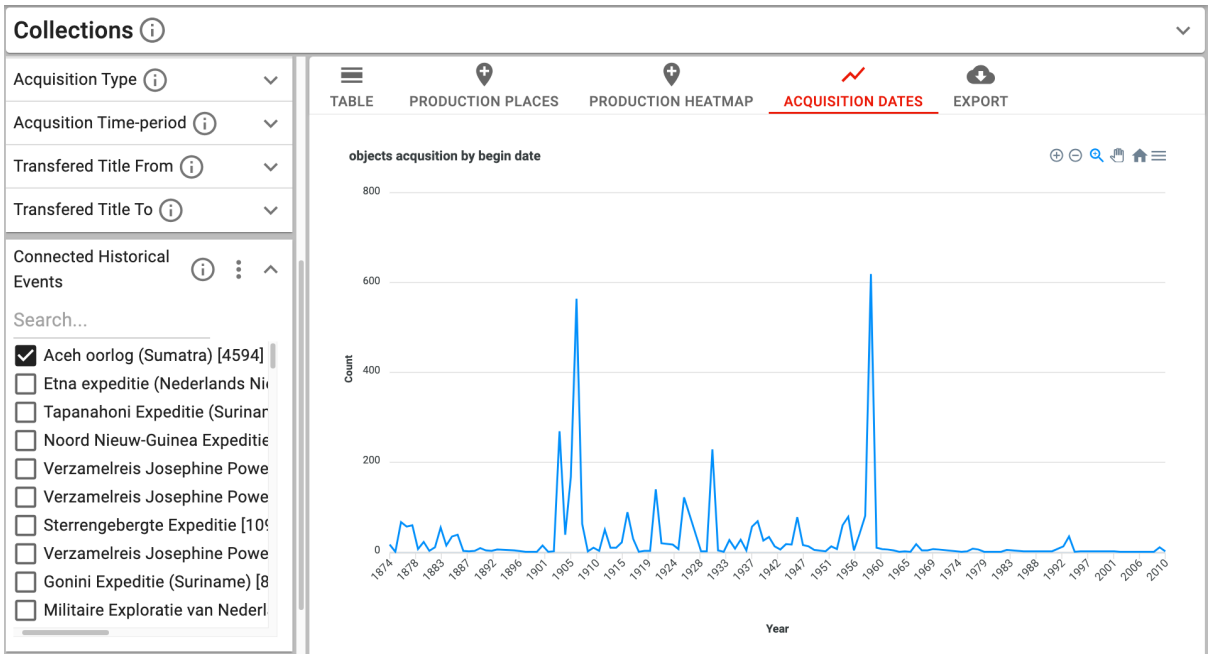


Figure 2: A timeline visualisation of begin date of objects acquisition events and objects has connection with Aceh Oorlog (Sumatra).

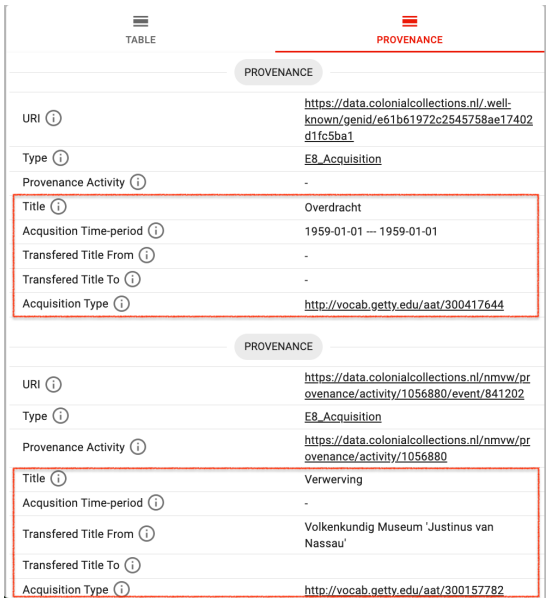
are linked to the Aceh War, the remaining 203 objects from the same acquisition remain unaccounted in this context. This discrepancy invites further investigation into the documentation of these objects, as well as the broader mechanisms of collection and documentation in colonial-era acquisitions.

The spike in acquisitions recorded in 1959 can be attributed to the incorporation of the *Ethnographic Museum Justinus van Nassau* in Breda as a subsidiary of the Volkenkunde Museum in Leiden (a predecessor of Wereldmuseum). When the museum in Breda was permanently closed in 1993, its ethnographic collection was transferred to the Volkenkunde Museum, explaining the recorded acquisition events from this period. However, it is crucial to note that while metadata documents the transfer of objects between institutions, it does not capture the historical context of displacement of these objects from their places of origin. One of the insightful features in PM-Sampo is provenance visualisation (cf. Figure 3a) that acquisition, which make this gap in the record apparent. One other thing becomes evident though this visual is that two attributes of acquisition event time and actor are never part of the same instance, which is clearly a gap in the record (further explanation is given in Section 5.5). Using a visual interface, such as provenance timelines, the demonstrator enables researchers to systematically track how objects moved through various hands, institutions, and contexts.

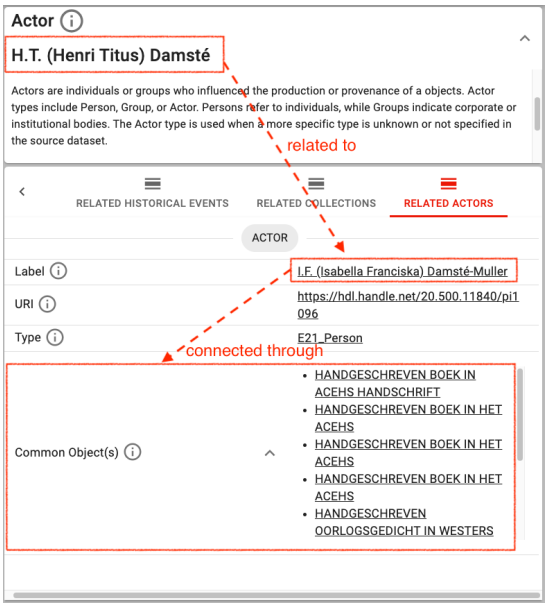
5.3. Actor-Network Visualisation

PM-Sampo plays a crucial role in establishing and analysing actor networks, which is valuable in provenance research. For example, objects collected by specific individuals, such as H.T. (Henri Titus) Damsté, are well documented in relation to historical events due to his connection with Aceh war. Objects associated with his wife, I.F. (Isabella Franciska) Damsté-Muller, may not be categorised with the same historical context. (However, in this particular case, both Mr. and Mrs. Damsté are explicitly connected to the Aceh War.) Their connection, which is established through the chain of custody within PM-Sampo (see Figure 3b), underscores the demonstrator's ability to reveal hidden relationships, such as family bonds, which could otherwise remain undetected in standard provenance metadata.

Future Work. Additionally, creating a visualisation of the actor network could significantly enhance the understanding of these connections. A network representation would allow researchers to identify hidden relationships, detect inconsistencies in metadata, and analyse the broader context of object circulation more effectively. This approach would provide a more intuitive way to explore provenance



(a) Provenance visualisation of an object from Volkenkunde Museum Justinus van Nassau. The acquisition date and acquisition actor is not attributed to same acquisition event. Metadata lacks history prior to museum acquisition.



(b) Visualisation of Actor-Actor relationship. Through shared objects, PM-Sampo established link between H.T. (Henri Titus) Damsté and his wife I.F. (Isabella Franciska) Damsté-Muller and visualised with *Related Actors* tab.

Figure 3: Instance page visualisation for (a) object and (b) actor in PM-Sampo

data, ensuring a clearer representation of how individuals, institutions, and objects are interlinked over time. Thus, implementing such a visualisation will be a key focus for future development.

5.4. Scaling Up Provenance Research

The ability to conduct large-scale provenance analysis and hypothesis generation is one of the key objectives of PM-Sampo. Without the support of data analysis, provenance research often remains limited to individual case studies due to the overwhelming amount of historical data. By leveraging linked data methodologies and existing metadata, PM-Sampo makes it possible to analyse thousands of objects simultaneously, visualising patterns, inconsistencies, and raising new research questions.

One of the critical advantages of this approach is its ability to visualise trends that might otherwise go unnoticed. For instance, a significant increase in object acquisitions occurring simultaneously with the onset of a historical event in a specific geographic region may raise questions about how these artifacts entered museum collections and whether their provenance has been accurately documented. For example, while the Aceh War formally began in 1873, provenance data reveals a considerable spike in acquisition events from museums in Aceh, Lombok, and Bali around the same period. Even if these objects are not explicitly documented as being connected to the Aceh War, the visualisation of acquisition patterns allows researchers to identify potential correlations that may not be immediately evident in textual metadata. This capability is particularly valuable for domain experts, as it enables them to assess whether spikes in acquisitions align with contemporary historical events and to investigate possible links between these events and the circulation of cultural objects.

Future work. Additionally, predicting historical events based on multidimensional patterns can provide insights into undocumented occurrences by analysing correlations between persons, places, time periods, and events, ultimately refining existing provenance narratives. Furthermore, deducing historical events from geographical and temporal trends allows reconstructing historical occurrences. Although visualisation techniques facilitated by PM-Sampo can assist in inferring new knowledge and generate hypotheses, the usefulness of these inferences is still subject to further study.

5.5. Missing or Ambiguous Data

A major challenge in provenance research is the ambiguity and incompleteness of historical records. PM-Sampo communicate this through facets that require categorisation of objects based on property values. Analysing this property value, it becomes apparent that same granularity level have not been maintained across dataset. For example, some objects are only linked to a broad origin (e.g., "Indonesia") rather than a precise location (e.g., "Aceh"). Similarly, certain collectors can only be identified at the organisational level rather than individually. By providing an overview of the distribution of determined versus undetermined provenance attributes, a pie chart visualisation could communicate the extent of missing data, guiding further research priorities.

Visualisation of provenance data reveals several significant gaps and inconsistencies in acquisition records. In particular, acquisition sites are almost never present, making it difficult to trace the geographical trajectory of objects before their entry into museum collections. Additionally, acquisition events mainly document when objects were acquired by museums, rarely capturing their prior movements through different hands before that. When examining provenance linked to actors, it becomes evident that these records predominantly list organisations on the receiving end of acquisition rather than individual collectors, further obscuring the specific pathways through which objects circulated. In addition, acquisition dates and acquisition actors are never attributed to the same acquisition instance, creating a disconnect in metadata that complicates efforts to establish a comprehensive and continuous chains of custody for these objects. Through visual communication, PM-Sampo makes it easy to identify which metadata needs to be recorded for better findability and accessibility of provenance record to foster future research.

6. Conclusion

The PM-Sampo demonstrator shows it potential to be a valuable tool for provenance research, offering new ways to explore historical objects' provenance, acquisition patterns, and hidden relationships among collectors, events, and places. The use-case analysis reveals critical gaps in acquisition records, unexpected geographical distributions of collected objects, and acquisition timelines extending well beyond the formal end of the conflict. Additionally, the establishment of actor connections uncovers previously unnoticed relationships between individuals and institutions involved in the movement of these objects. The findings underscore the importance of structured provenance data and highlight the need for more comprehensive documentation practices.

By visualising provenance gaps, linking actors, and uncovering inconsistencies, PM-Sampo facilitates a more comprehensive understanding of colonial-era collections. These insights provide a strong foundation for further historical analysis and ethical reassessment of museum holdings related to contested colonial histories. At the same time, by demonstrating the advantages of publishing provenance metadata as Linked Open Data (LOD), this paper advocates for its wider adoption. Through visualisation, we illustrate how structured, open data can support provenance research, encourage interdisciplinary analysis, and contribute to the broader cultural heritage domain.

Acknowledgments. This research was supported by the NWA-funded project Pressing Matter (NWA.1292.19.419), by the Research Council of Finland FIN-CLARIAH funding from the European Union NextGenerationEU instrument, and by the Aalto Science Institute (ASCI) Visiting Doctoral Researcher Programme. Computing resources provided by the CSC – IT Center for Science were used in our work.

Declaration on Generative AI. Authors acknowledge the use of Writefull's AI tools embedded in Overleaf environment during the writing of the manuscript solely for grammar and spelling checks.

References

[1] A. Tompkins, Provenance Research Today, Lund Humphries, 2021. doi:10.1111/cura.12528.

[2] F. Sarr, B. Savoy, The Restitution of African Cultural Heritage, Ministère de la Culture, 2018.
[3] B. Latour, Reassembling the social: an introduction to actor-network-theory, Clarendon lectures in management studies, Oxford University Press, Oxford ; New York, 2005. OCLC: ocm58054359.
[4] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, From data mining to knowledge discovery in databases, AI Magazine 17 (1996) 37. doi:10.1609/aimag.v17i3.1230.
[5] I. Tiddi, M. d'Aquin, E. Motta, Data patterns explained with linked data, in: A. Bifet, M. May, B. Zadrozny, R. Gavalda, D. Pedreschi, F. Bonchi, J. Cardoso, M. Spiliopoulou (Eds.), Machine Learning and Knowledge Discovery in Databases, Springer International Publishing, Cham, 2015, pp. 271–275.
[6] L. Zuckerman, Tracking looted art with graphs., Graphs and Networks in the Humanities 2022 Conference, February 3–4, 2022 (2022).
[7] E. Hyvönen, Using the semantic web in digital humanities: Shift from data publishing to data-analysis and serendipitous knowledge discovery, Semantic Web 11 (2020) 187–193.
[8] M. G. Kirschenbaum, The remaking of reading: Data mining and the digital humanities, in: The National Science Foundation symposium on next generation of data mining and cyber-enabled discovery for innovation, Baltimore, MD, volume 134, 2007.
[9] M. Falkenthal, J. Barzen, U. Breitenbücher, S. Brügmann, D. Joos, F. Leymann, M. Wurster, Pattern research in the digital humanities: how data mining techniques support the identification of costume patterns, Computer Science - Research and Development 32 (2017) 311–321. doi:10.1007/s00450-016-0331-6.
[10] W. McCarty, Humanities Computing, International series of monographs on physics, Palgrave Macmillan London, 2005.
[11] E. Gardiner, R. G. Musto, The Digital Humanities: A Primer for Students and Scholars, Cambridge University Press, 2015.
[12] A. Sheth, B. Aleman-Meza, I. B. Arpinar, C. Bertram, Y. Warke, C. Ramakrishnan, C. Halaschek, K. Anyanwu, D. Avant, F. S. Arpinar, K. Kochut, Semantic association identification and knowledge discovery for national security applications, Journal of Database Management on Database Technology 16 (2005) 33–53.
[13] V. Viswanathan, K. Ilango, Ranking semantic relationships between two entities using personalization in context specification, Information Sciences 207 (2012) 35–49.
[14] E. Hyvönen, Using the semantic web in digital humanities: Shift from data publishing to data-analysis and serendipitous knowledge discovery, Semantic Web 11 (2020) 187–193. doi:10.3233/SW-190386.
[15] E. Hyvönen, H. Rantala, Knowledge-based relational search in cultural heritage linked data, Digital Scholarship in the Humanities (DSH) 36 (2021) 155–164. doi:https://doi.org/10.1093/llc/fqab042.
[16] E. Hyvönen, Digital humanities on the Semantic Web: Sampo model and portal series, Semantic Web journal 14 (2022) 729–744. doi:10.3233/SW-223034.
[17] E. Ikkala, E. Hyvönen, H. Rantala, M. Koho, Sampo-UI: A full stack JavaScript framework for developing semantic portal user interfaces, Semantic Web 13 (2022) 69–84. doi:10.3233/SW-210428.
[18] H. Rantala, A. Ahola, E. Ikkala, E. Hyvönen, How to create easily a data analytic semantic portal on top of a SPARQL endpoint: introducing the configurable Sampo-UI framework, in: Proceedings of 8th International Workshop on the Visualization and Interaction for Ontologies and Linked Data co-located with the 22nd International Semantic Web Conference (ISWC 2023) in Athens, Greece, CEUR Workshop Proceedings, Vol. 3508, 2023. URL: https://ceur-ws.org/Vol-3508/paper3.pdf.
[19] M. Doerr, The cidoc conceptual reference module: an ontological approach to semantic interoperability of metadata, AI Magazine 24 (2003) 75–75.
[20] S. B. A. Shoilee, V. de Boer, J. van Ossenbruggen, Polyvocal knowledge modelling for ethnographic heritage object provenance, in: Knowledge Graphs: Semantics, Machine Learning, and Languages, volume 56, IOS Press, Leipzig, Germany, 2023, pp. 127–143.
[21] D. Tunkelang, Faceted search, Synthesis Lectures on Information Concepts, Retrieval, and Services, Morgan & Claypool, 2009.