

# **Structuring Cultural Heritage PROVenance: The Rijksmuseum Use Case Lessons Learned from Existing Provenance Practices**

Steven Kuhnen<sup>2</sup>, Rebecca Sepers<sup>2</sup>, Chris Dijkshoorn<sup>3</sup>, Victor de Boer<sup>1</sup>, Lora Aroyo<sup>1</sup>

<sup>1</sup> Vrije Universiteit Amsterdam

<sup>2</sup> Universiteit van Amsterdam

<sup>3</sup> Rijksmuseum Amsterdam

The Cultural Heritage domain has been historically quite provenance-centric and has developed various strategies and practices for capturing, recording and maintaining provenance information of their collections. Advances in digital technologies provide limitless opportunities for performing these tasks more effectively, efficiently and at scale. Initiatives such as W3C Provenance working group have produced digital formats to help provenance modeling and representation across different domains. Different cultural heritage institutions have deployed their collection specific guidelines. However, there is little overlap between existing Web standards and the cultural heritage practices - Web standards primarily focus on 'data provenance', where heritage institutions focus on the provenance of real world objects in their collections. Alignment among different cultural heritage institutions with existing Web standards can not only help optimize the internal processes, but can also improve the work of digital humanities scholars searching and researching across different collections online.

Fifteen years ago the Rijksmuseum defined guidelines for recording provenance data of their collection within the context of the Rijksmuseum Provenance project [1]. Since then, these guidelines have been adapted a number of times to find a way to incorporate the different types of provenances encountered. Naturally, this resulted in inconsistencies or missing information in the provenance records. For example, there are currently 3.637 objects transferred from one owner to another without the actual type of provenance being known. In this paper we present the results of an initial computational assessment of some of the Rijksmuseum provenance records (e.g. we analysed the data for four provenance types - Auctions, Estate Inventories, Art markets and dealers and Inheritance (by descent, by widows)). The dataset used for this analysis consists of the provenance history between 1933-1945 of 34,202 artworks [2]. The goal is to identify and analyze the existing bottlenecks in the process and their potential impact on the interoperability of the collection, and ultimately define an approach for a solution based on existing standards, e.g. W3C PROV model<sup>1</sup>, already applied in the cultural heritage domain [3] to model the lifetime of artworks in terms of the activities and agents related to their ownership and location (i.e. the chain of custody for artworks in the Rijksmuseum). It also ensures that the artwork provenance information is provided in a structured and queryable form. We believe that the lessons

---

<sup>1</sup> <https://www.w3.org/TR/prov-primer/>

learned from this data analysis would be valuable beyond the Rijksmuseum use case alone, and they would help other heritage institutions to update and improve their practices. In summary, this paper reports on the identification and classification of **provenance issues** and their **impact rank** with respect to different tasks and cases. We identify the **bottlenecks** in addressing each of the issues and offer possible **solution strategies** for each of the issues. Reflecting on the results of this work, we can continue analysing computationally other types and in this way continue the effort to improve both practices and the resulting provenance records at the Rijksmuseum.

```
<?xml version="1.0" encoding="UTF-8"?>
- <adlibXML>
  - <recordList>
    - <record>
      <prirref>2</prirref>
      <PersistentIdentifier>RM0001.COLLECT.2</PersistentIdentifier>
      - <provenance>
        <provenance.project.txt>...; collection H.R. van Maasdijk, The Hague, 1932;{Blankert 1991, note 6; note RKD.}...; pri
          collection;{Blankert 1991, p. 507.}...; anonymous sale, The Hague (Glerum), 25 November 1991, no. 123, to th
        </provenance>
      </record>
```

**Fig. 1 Rijksmuseum Provenance Structure**

## References

- [1] Hakkareinen, M, and Koivulahti, T.: Research Into Art Looted by the Nazis - An Important International Task. *Nordisk Museologi*, vol. 1, 58-73 (2007).
- [2] Karrels, Nancy: Renewing Nazi-Era Provenance Research Efforts: Case Studies and Recommendations. *Museum Management and Curatorship*, vol. 29:4, 297-310 (2014)
- [3] Sadusky, R: Computational Provenance: DataONE and Implications for Cultural Heritage Institutions. *International Conference on Big Data*, 2016
- [4] Schöch, C: Big? Smart? Clean? Messy? Data in the Humanities. *Journal of Digital Humanities*, vol.2 (3), 2-13 (2013)