

Rick Mourits
Thunnis van Oort
Kay Pepping
Pascal Konings
Britt van Duijvenvoorde

Modelling the enslaved as historical persons

Extending the Persons in Context (PiCo) model to fit a 19th century slave society.

The Historical Database of Suriname and the Caribbean (HDSC) publishes open data on persons living in the former Dutch colonies in Suriname and the Caribbean, in collaboration with knowledge and heritage partners and citizen scientists throughout this region and the Netherlands (see Van Galen et al., 2023; <https://datasets.iisg.amsterdam/dataverse/HDSC>). As a result, historical person observations have been collected that can be used to reconstruct historical persons. To implement these person observations and reconstructions within a wider community and make them interoperable and accessible to the general and scholarly public, the use of a clear and systematic data model is required.

Exchange of large volumes of data between different parties puts specific requirements on the data. For example, data needs to be shared and curated on a large scale. Therefore, the Dutch Digital Heritage Network (NDE) and Ministry of Education, Culture, and Science (2021; 2024) have appointed the Semantic Web as the standard for digital heritage data. The advantage of the Semantic Web, also known as the Resource Description Framework (RDF) or Linked Data, is that each bit of information in a dataset is described with a unique URI, which can be represented as Uniform Resource Locators (URLs) to retrieve and describe entities or properties. As a result, the Semantic Web makes it possible to effectively exchange data and write software, as long as multiple parties use the same URIs.

In order to effectively communicate (historical) person data, the Dutch Center for Family History (CBG) has developed the [Persons in Context](#) (PiCo) model (Woltjer et al., 2024), on the basis of the ROAR model developed by Menno den Engelse and Leon van Wissen (2021). The PiCo model states how information on common factors in historical person observations, such as age, dates of birth, marriage, and death, occupations, relations between family members, roles in source documents, etc. should be described. We chose to implement this model for our own databases with historical person data on enslaved people, as the purpose of the PiCo model is to fit all historical person observations and to provide flexibility and guidelines for extending the model according to the requirements of specific historical data sets. The PiCo model is sustained by a PiCo maintenance group, consisting of the Center for Family History and multiple archival as well as research institutions.

The PiCo model reuses as many existing URIs as possible to describe historical person data, as it encourages communities to give the same name to the same entity. This prevents discussions on naming and describing resources that have already been authoritatively defined, such as birth dates or names. On the one hand, this makes data more findable as people can search on one common term, rather than having to waste time on resolving a

potentially endless list of variations in modelling the same concept. But, perhaps even more importantly, it also enforces stringent definitions of properties and resources that can be used for efficient knowledge interfering.

Implementing the Persons in Context model

Reuse of data is easiest when data is described concentrically, that means by reusing general, domain-specific, and specialised schemas for resources that are commonly known in the world, domain, or field in question (Meroño-Peñuela et al., 2020; Mourits et al., 2024). This makes it relatively easy to model historical person data living in slave societies, as most resources have already been described within the PiCo model, and specialised concepts can be reused from the history of slavery domain.

Within the history of slavery, [Enslaved.org](https://www.enlaved.org/) has used the Semantic Web technology to describe enslaved people (Hawthorne et al. 2023; Shimizu et al., 2023). In order to do so, they described all concepts encountered in historical databases on enslavement in their own data model. Sadly, the Enslaved.org model is not reusable for our data in its entirety, as it uses an event-based model that is not compatible with the concentric description used in the PiCo-model. Nevertheless, we are very enthusiastic about the Enslaved project's efforts to provide domain-specific terms that can enhance the interoperability of data sets relating to slavery.

By incorporating elements of the Semantic Web vocabularies of the Enslaved project in the design of the Persons in Context model, we were able to convert the 1828-1847 Paramaribo Ward Registers (Sang-Ajang et al., 2024) and the 1830-1863 Surinam Slave Registers (Rosenbaum-Feldbrügge et al., 2023) into ttl format¹. On a few occasions, we were forced to create new properties required to accurately model the data. We would prefer to refer to Enslaved.org properties in these cases instead of introducing these properties ourselves. However, we were unable to reach a consensus with Enslaved.org on how to model these properties, nor locate another authority.

Presenting a use case: re-linking identities of urban enslaved

To show the practical advantage of working with the extended data model, we wish to demonstrate at the SSHA 2025:

- a) what the problems are for modeling historical person data slave societies. There are some incongruencies between the Enslaved.org ontology/thesaurus and the inherent logic behind the PiCo model. Moreover, there seem to be some elements about which both PiCo and Enslaved.org are not very elaborate. Some of these issues are already clear after modelling the data, but others will become clear after using the data;
- b) what the added value of the Semantic Web is by linking slaveholders in the neighborhood registers and slave registers. We should be able to match graphs with standardised protocols, as all entities and properties in the graph are already standardised and should require little to no data transformation. Added advantage is that we actually test the extended data model in practice, which will no doubt highlight additional issues;

¹ see also: <https://github.com/HDSC-Nijmegen/Slavenregisters/tree/main/ttl%20conversion> and <https://github.com/HDSC-Nijmegen/Ward-registers>

- c) how the results can be shared in a data story and other online formats.

Lessons learned and way forward

Using the Enslaved.org vocabularies we extended the Persons in Context model, so that it can also be used to describe persons data model person data for a 19th century slave society. Following the methodology of concentric description, we reused existing terminology on the Semantic Web to describe our data. Our goal is to extend it even further and make the model also suitable for slave societies in the 17th and 18th centuries, as well as slave societies beyond the Atlantic. However, before doing so, we first need to design a workflow to deal with concepts that did not seem to be readily available, which currently present a conundrum on how best to proceed.

We reached out to the PiCo maintenance group to ask for feedback on how best to proceed. We explained how we extended the PiCo model by reusing Enslaved.org URIs to describe our data, and requested whether PiCo could implement a few proposed properties as these were not provided by Enslaved.org, at least not in the way we want. The feedback² from the PiCo maintenance group was that extensions to the PiCo model should come from authoritative concepts in specialised fields.

- The extension of the PiCo model is a best practice because it is reusing a specialized vocabulary.
- PiCo also wants to recommend such extensions in the long run if they are used more widely in a field.
- PiCo itself will not introduce new terminology if a specialized field does not provide it.

In other words, PiCo will not come up with property until there is consensus within slavery history. Therefore, we have made the first steps to collaborate with the GLOBALISE and ESTA projects, so that we can move to a consensus on how to model historical person data in the context of enslavement.

At the SSHA 2025 we hope to show the data model and discuss how it works in practice with a use case. Our goal here is fourfold: modelling enslaved persons as persons whose lives can be reconstructed by combining multiple sources, showcasing PiCo's current limits for modelling the lives of enslaved people, showcasing how existing data models can be extended using community standards, and giving a first glimpse of the reconstructed life course of the enslaved. This would only be the beginning, as the wide range of colonial sources that we want to support, will provide more insight into the lives of enslaved as well as raise new modelling issues.

Literature

Den Engelse, M., & van Wissen, L. (2019). *Reconstructions and observations in archival resources: Ontology specification*. <https://w3id.org/roa>

² see: <https://github.com/CBG-Centrum-voor-familiegeschiedenis/PiCo/issues/30>

- Hawthorne, Walter, Daryle Williams, and Dean Rehberger (Directors). (2023). *Enslaved: Peoples of the Historical Slave Trade* (Enslaved.org), Version 2. [Digital Project]. Michigan State University: Matrix Center for Digital Humanities & Social Sciences.
- Meroño-Peñuela, A., De Boer, V., Van Erp, M., Zijdemans, R. L., Mourits, R. J., Melder, W., Rijpma, A., & Schalk, R. (2020). *CLARIAH: Enabling interoperability between humanities disciplines with ontologies*. In G. Cota, M. Daquino & G. L. Pozzato (Eds.), *Applications and practices in ontology design, extraction, and reasoning* (pp. 73–90). <https://doi.org/10.3233/SSW200036>
- Ministerie van Onderwijs, Cultuur en Wetenschap & Netwerk Digitaal Erfgoed (2024). *Nationale strategie digitaal erfgoed 2025-2028*. <https://open.overheid.nl/documenten/454c3aac-2a77-407c-b979-807b66ba6719/file>
- Mourits, R.J., Riswick, T., Stapel, R. (2025). *Common Language for Accessibility, Interoperability, and Reusability in Historical Demography*. In: Steffen, B. (eds) *Bridging the Gap Between AI and Reality*. AISoLA 2023. Lecture Notes in Computer Science, vol 14129. Springer, Cham. https://doi.org/10.1007/978-3-031-73741-1_2
- Netwerk Digitaal Erfgoed & Ministerie van Onderwijs, Cultuur en Wetenschap (2021). *Nationale strategie digitaal erfgoed*. <https://open.overheid.nl/documenten/ronl-cec5f78e-8659-4a02-8fb6-4923ab26d95b/pdf>
- Rosenbaum-Feldbrügge, M., Mourits, R.J., A.B., M., Janssen, J., Quanjer, B., Van Oort, T., Kok, J. & Van Galen, C. (2023). *Suriname Slave and Emancipation Registers Dataset Version 1.1*. IISH Data Collection. <https://hdl.handle.net/10622/CSPBHO>
- Sang-Ajang, J., Altink, N., Dikland, P., Jonkers, C., Kariomengolo, Valies, C. & Van Oort, T. (2024). *Paramaribo Ward Registers 1828-1847*. IISH Data Collection. <https://hdl.handle.net/10622/VLN8FD>
- Shimizu, C. et al. (2023). *The Wikibase Approach to the Enslaved.Org Hub Knowledge Graph*. In: Payne, T.R., et al. *The Semantic Web – ISWC 2023*. ISWC 2023. Lecture Notes in Computer Science, vol 14266. Springer, Cham. https://doi.org/10.1007/978-3-031-47243-5_23
- Van Galen, C.W., Mourits, R. J., Rosenbaum-Feldbrügge, M., A.B., M., Janssen, J., Quanjer, B., Van Oort, T., & Kok, J. (2023). *Slavery in Suriname. A Reconstruction of Life Courses, 1830–1863*. *Historical Life Course Studies*, 13, 191-211. <https://doi.org/10.51964/hlcs15619>
- Woltjer, P. E., Zandhuis, I., Coret, B., Lindeman, M., Balkenende, J. D., Zijdemans, R. L., & Mourits, R. J. (2024). *Persons in Context. A Model to Represent Observations and Reconstructions of Historical Persons in Linked Data*. *Historical Life Course Studies*, 14, 105-125. <https://doi.org/10.51964/hlcs19312>