

## **Mapping and aligning person name elements in LOD using the Person Name Vocabulary (PNV)**

DHBenelux 2019

Poster

Lodewijk Petram (Huygens Institute for the History of the Netherlands)

Gijsbert Kruithof (National Archives of the Netherlands)

Although there are many rdf-vocabularies for person names (e.g. schema.org, foaf, DBpedia and Wikidata), their limited granularity makes them not very suitable for Arts & Humanities research. While historians often find it useful to correctly identify patronymics, nobility scholars have a special interest in noble titles, and onomastics researchers wish to distinguish between all elements of a person name, the standard vocabularies generally only allow for identifying given and family names, and sometimes also honorary and academic titles. This poster presents the Person Name Vocabulary (PNV) [1], an rdf vocabulary that accommodates different levels of data granularity and allows for easy alignment of name elements, including idiosyncratic ones such as family name prefixes and patronymics, with standard vocabularies (schema.org, foaf, DBpedia, Wikidata, and A2A, an XML-vocabulary used by many archival institutions in the Netherlands, as well as CIDOC CRM, a conceptual ontology that is the international standard used by heritage institutions), thus guaranteeing optimal data interoperability. PNV is primarily aimed at, but not limited to, modelling (historical) person names from the Dutch language area.

Since people are central to many (research) datasets, and the same persons may appear in multiple datasets, good findability, re-usability and interoperability of persons' data is beneficial to data users in general and digital humanities researchers in particular. The PNV allows data providers to make their persons' data findable,

re-usable and interoperable, while maintaining the level of detail in which the persons' names are described. The vocabulary's flexibility makes that data users can, for example, pick and choose which name elements to include in an automated process of entity linkage, and in which order. It is also instrumental in tailoring the way in which names are displayed and sorted in user interfaces, allowing institutes from the Netherlands for example to sort 'Jan de Vries' under 'Vries', while acknowledging that the family name prefix 'de' is an integral part of the family name.

The poster will introduce the vocabulary and its predefined mapping with various other vocabularies, and showcase the benefits of implementing it by exploring a use case (which includes name variants — an important, but often difficult, part of name modelling).

Institutes / projects that have already adopted PNV in their rdf representations of person observations, or plan to do so in the near future, include the Dutch National Archives (*nadere toegangen*, i.e. archive indices), the Huygens Institute for the History of the Netherlands (Biography Portal of the Netherlands and other collections), CLARIAH, Historisch Leiden in Kaart (various institutions), Golden Agents (various institutions), and Bossche protocollen (Huygens ING / Erfgoed 's-Hertogenbosch).

[1] <https://w3id.org/pnv#>