

# An Ontology and Data Infrastructure for Publishing and Using Biographical Linked Data

Petri Leskinen<sup>1</sup>, Jouni Tuominen<sup>1,2</sup>, Erkki Heino<sup>1,2</sup>, and Eero Hyvönen<sup>1,2</sup>

<sup>1</sup> Semantic Computing Research Group (SeCo), Aalto University, Finland and

<sup>2</sup> HELDIG – Helsinki Centre for Digital Humanities, University of Helsinki, Finland

<http://seco.cs.aalto.fi>, <http://heldig.fi>  
firstname.lastname@aalto.fi

**Abstract.** This paper describes the ontology model and published datasets of a digitized biographical person register. The applied ontology model is designed to represent people via their enduring roles and perduring lifetime events. The model is designed to support 1) prosopographical Digital Humanities research, 2) linking to resources in semantic Cultural Heritage portals, and 3) semantic data validation and enrichment by using SPARQL queries. The linked data approach enables to enrich a person's biography by interlinking it with space and time related biographical events, persons relating by social contacts or family relations, historical events, and personal achievements.

**Keywords:** Semantic Web, Linked Open Data, Actor Ontology, Digital Humanities, Cultural Heritage, Prosopography, Biographical Representation

## 1 Introduction

This resource description paper presents the LOD infrastructure, data model, and datasets used in the Norssit alumni register of short biographies. The data model is designed to support prosopographical research, data aggregation and linking in semantic portals, and easy SPARQL querying. The datasets consist of 10 137 person resources, enriched with graphs of relating career events and family relations, and vocabularies of titles, schools, companies, medals, and hobbies.

The data has been used in creating the Vanhat Norssit Portal<sup>3</sup> allowing the user to search and browse the data about individual persons as well as analyze and visualize data about groups of people in prosopographical research [1,11]. The user can filter the results by making selections on the facets on the left side of the page. For visualizing the data the portal has two views that use Google Chart<sup>4</sup> diagrams. On the first one<sup>5</sup>, the pie charts show the popularity of most common educations, universities and colleges, professions, and employers after the graduation of the alumni. On the second one<sup>6</sup>, years of enrollment and matriculation are shown using histograms, and below these,

<sup>3</sup> <http://www.norssit.fi/semweb>

<sup>4</sup> <https://developers.google.com/chart/>

<sup>5</sup> <http://www.norssit.fi/semweb/#!/visualisointi>

<sup>6</sup> <http://www.norssit.fi/semweb/#!/visualisointi2>

multi-column charts showing the most popular universities and colleges, employers, and occupations by decade. The digitization, lodification, and the Vanhat Norssi Portal is described in more details in [8].

This paper is structured as follows: First, an ontology model for representing people with their life time events and relation roles is introduced. Secondly, the data sets of the use case Norssit alumni with information extraction from textual data is discussed. Then the results of entity linking are evaluated. Finally, the related work, lessons learned, and future work are discussed.

## 2 Person Ontology Model

The ontology model representing people and their biographical information in the use case Norssit alumni is based on the Bio CRM model<sup>7</sup>, which has been developed to facilitate and harmonize the representation of biographies and cultural heritage data on the Semantic Web. Bio CRM is a domain specific extension of CIDOC CRM<sup>8</sup> [3]. It includes structures for basic data of people, personal relations, professions, and events with participants in different roles. Bio CRM makes a distinction between enduring unary roles of actors, their enduring binary relationships, and perduring events, where the participants can take different roles modeled as a role concept hierarchy.<sup>9</sup> Bio CRM provides the general data model for biographical datasets, and the individual datasets concerning different cultures, time periods, or collected by different researchers may introduce extensions for defining additional event and role types.

| Namespace   | Prefix  |
|---|---------|
| <a href="http://ldf.fi/norssit/">http://ldf.fi/norssit/</a>   | :       |
| <a href="http://ldf.fi/schema/bioc/">http://ldf.fi/schema/bioc/</a>                                   | bioc:   |
| <a href="http://purl.org/dc/terms/">http://purl.org/dc/terms/</a>                                     | dct:    |
| <a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/</a>                                   | foaf:   |
| <a href="http://schema.org/">http://schema.org/</a>   | schema: |
| <a href="http://www.w3.org/2004/02/skos/core#">http://www.w3.org/2004/02/skos/core#</a>               | skos:   |
| <a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#</a> | rdf:    |
| <a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#</a>             | rdfs:   |

**Table 1.** The namespaces and prefixes used in the ontology.

The main classes of the person ontology are shown in Figure 1. A principle is that a **foaf:Person** instance only contains the properties that are considered constant, in our case family and first names, places, and dates of birth and death, etc. The person resource is enriched by attaching family relations, achievements, and titles.

<sup>7</sup> <http://seco.cs.aalto.fi/projects/biographies/>

<sup>8</sup> <http://cidoc-crm.org>

<sup>9</sup> <http://ldf.fi/schema/bioc/>

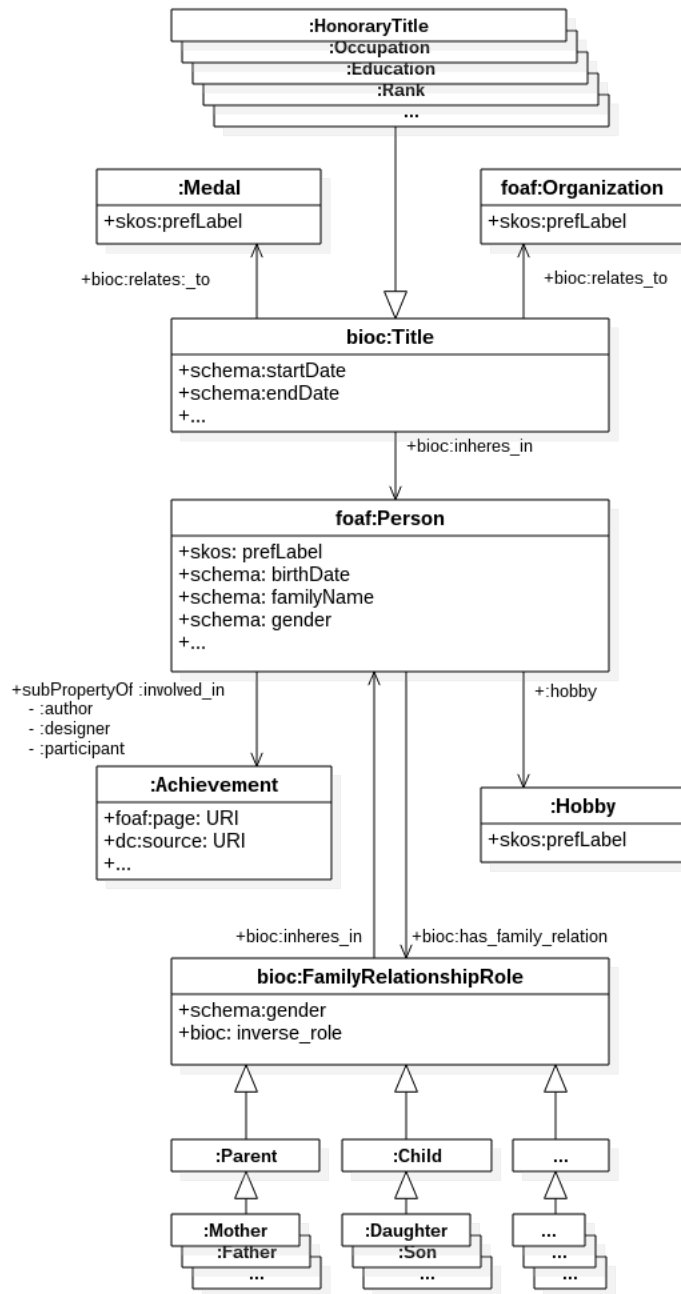


Fig. 1. The ontology schema.

## 2.1 Modeling Family Relations

Family relation is an example of a binary, often even N-nary, relationship connecting two or more people. Social relationships can be modeled in a similar manner. Each family relation is a subclass of **bioc:FamilyRelationshipRole**. The domain specific ontology of family relations can build a hierarchy (e.g. **:Mother** and **:Father** are subclasses of **:Parent**). The RDF example below shows a definition of a class and declaration of a relationship between two people. Notice that the property **bioc:inverse\_role** has two values, depending on the gender of the relative. Genders and inverse relations can be used for data evaluation and reasoning: this SPARQL query<sup>10</sup> fills in the missing inverse relationships in the dataset. A family relation is attached to a person using the property **bioc:has\_family\_relation**, which can have a blank node or a resource as a value.

```
:Mother a
  rdfs:subClassOf      rdfs:Class ;
  rdfs:subClassOf      :Parent ;
  bioc:inverse_role    :Son , :Daughter ;
  schema:gender        schema:Female ;
  skos:prefLabel       "Mother"@en , "äiti"@fi .

:person_1 a
  foaf:Person ;
  schema:gender        schema:Male ;
  bioc:has_family_relation [ a :Mother ;
                           bioc:inheres_in :mother_1 ] .
```

## 2.2 Modeling Personal Achievements

A personal achievement refers to any notable activity (e.g. producing a work of art, a design project, receiving a political achievement, or participating in athletes games). Achievements are modeled so that a subproperty of **involved\_in** connects the person to a instance of class **:Achievement**, and furthermore indicates what is the role of the person in a particular achievement. So a single achievement can refer to multiple people, and indicate the roles they participated with, e.g. an actor, author, or director of a movie. Instances of **:Achievement** can contain a description, information of time, place, and provenance, and link to a corresponding LOD resource or to a web page.

```
:norssi_2230 :author :achievement_33 .

:author a
  rdf:Property ;
  skos:prefLabel "teoksia"@fi , "Novels"@en ;
  rdfs:subPropertyOf :involved_in .

:achievement_33 a
  :Achievement ;
  skos:prefLabel "Sinuhe egyptiläinen."@fi ,
  "Sinuhe the Egyptian."@en ;
  dct:source     <https://fi.wikipedia.org/wiki?curid=820> ;
  :wikipedia     <https://en.wikipedia.org/wiki/The_Egyptian> .
```

<sup>10</sup> Example of a SPARQL query: <http://yasgui.org/short/BkRKKXyIZ>

### 2.3 Modeling Career Roles

According to the Bio CRM principle, the occupation or profession of a person is considered a role. Another unary role modeled in a similar manner would be person's nationality. The resource is a subclass of **bioc:Title**. The person involved is attached with a subproperty of **bioc:inheres\_in**, additional information like the company or medal of honour by subproperties of **bioc:relates\_to**.

```
:tekniikan%20tohtori
  a
  rdfs:subClassOf
  skos:altLabel
  skos:prefLabel
    rdfs:Class ;
    :Education ;
    "TkT"@fi, "D.Sc. (Tech.)"@en ;
    "tekniikan tohtori"@fi ,
    "Doctor of Science (Technology)"@en .

:event_21721 a
  bioc:inheres_in
  schema:startDate
  skos:prefLabel
    :tekniikan%20tohtori ;
    :norssi_7686 ;
    "1991"^^xsd:gYear ;
    "TkT 91"@fi .
```

## 3 Norssit Dataset

As a concrete case study, a register 1867–1992 of over 10 000 alumni of the prominent Finnish high school “Norssi” was scanned, OCR'd, and transformed into RDF, then enriched by data linking, published as a linked data service, and finally provided to end users via a faceted search engine and browser for studying lives of historical persons and for prosopographical research. [8]

### 3.1 Information Extraction

The most important data source was the textual description of register entries. In Figure 2 a register entry is depicted, and some of the data fields are picked as examples. Description texts are well-formatted, and always start with person's name (a) and his place and date of birth (b, Jyväskylä, 19th Aug., 1868). Description includes names of person's parents, and his years of enrollment and matriculation (c, *yo 88*). His later university degrees with graduation years are mentioned (d, *FT = Ph.D.*). The career is described as a list of entries in format *Company role years* (e, *toimJ = CEO*). Possible medals of honour are mentioned (f, *VirVR 1 mk = Estonian Cross of Liberty, 1st Class*) as well as military ranks with promotion years (g, *Evl = Colonel Lieutenant*). The description ends with a possible date of death (h, *11th Dec., 1939*) and list of family relations (i, *Veli = Brother*). The data fields were extracted using regular expressions.

### 3.2 Datasets

Currently the Norssit dataset contains ca 892 000 triples defining 131 000 resources. The main graphs are discussed in detail below, with the graph names, amounts of instances and triples, main classes, and properties.

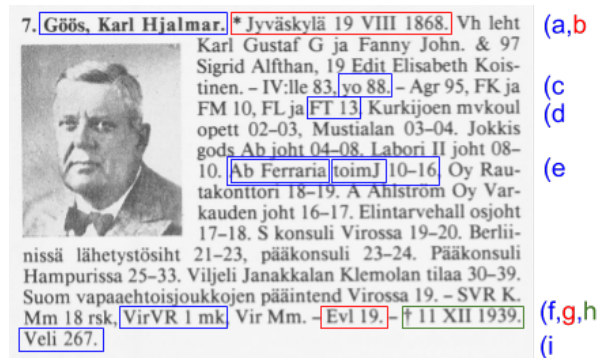


Fig. 2. A biographical entry in the register book with examples of extracted data fields.

## People

Graph: <http://ldf.fi/norssit/people>

Contains: 17 791 instances, 183 972 triples

Core classes: foaf:Person

Properties: schema:familyName, schema:givenName, schema:gender, skos:prefLabel, schema:birthDate, schema:birthPlace

This graph includes data of 10 137 people. A person resource contains biographical data extracted from register descriptions, e.g. given and family names, gender, places and dates of birth and death, years of enrollment, matriculation, or resignation, and provenance data. This graph also includes profile image URIs, family relationship instances, and links to external LOD clouds.

## Career events

Graph: <http://ldf.fi/norssit/events>

Contains: 34 000 instances, 247 237 triples

Core classes: subclasses of bioc:Title

Properties: bioc:inheres\_in, bioc:relates\_to, schema:startDate, schema:endDate

The event graph contains 34 000 career events extracted from register descriptions (e.g. see e in Fig. 2). Each resource contains links to an occupational or educational title, person involved, start and end years, a description, and a possible organization or medal. Altogether 5882 people are enriched with a title.

## Achievements

Graph: <http://ldf.fi/norssit/achievements>

Contains: 3000 instances, 15 578 triples

Core class: :Achievement

Properties: :involved\_in, skos:prefLabel, dct:source, :wikipedia

The achievement graph contains 3000 personal achievements extracted from Wikipedia pages, or BookSampo<sup>11</sup> Linked Data. In the case of Wikipedia, the information was extracted from HTML code under specified subtitles. Each resource provides a description, specified category, links to a Wikipedia page of the achievement, and link to person's Wikipedia page served as an source of information.

### Organizations

Graph: <http://ldf.fi/norssit/organizations>

---

Contains: 2300 instances, 5390 triples

Core classes: **foaf:Organization**, **schema:EducationalOrganization**

Properties: **skos:prefLabel**, **skos:altLabel**, **dct:source**

The organization graph contains the labels and abbreviations of 2401 organizations, e.g. government agencies, companies, colleges, or universities. The labels were collected from text descriptions. An organization is attached to an event using the **bioc:relates\_to** property. Altogether 4805 people are linked to an organization.

### Hobbies

Graph: <http://ldf.fi/norssit/hobbies>

---

Contains: 1760 instances, 3520 triples

Core class: **:Hobby**

Property: **skos:prefLabel**

The vocabulary of hobbies contains labels of 1760 different hobbies (e.g. Music, Sports, or Arts), mentioned in register descriptions. A hobby is attached using the **:hobby** property of Person resource. Altogether 7845 people are related with a hobby.

### Titles

Graph: <http://ldf.fi/norssit/titles>

---

Contains: 350 instances, 1526 triples

Core classes: subclasses of **bioc:Title**

Properties: **skos:prefLabel**, **skos:altLabel**

The titles graph contains classes of 350 educational or occupational titles and military ranks. These are the classes of instances in Career events graph. Altogether 5882 people have a specified title.

### Medals

Graph: <http://ldf.fi/norssit/medals>

---

Contains: 301 instances, 1254 triples

Core class: **:Medal**

Property: **skos:prefLabel**

---

<sup>11</sup> <http://www.kirjasampo.fi>

The medals vocabulary contains 301 different types of honorary medals, the data is extracted from register descriptions (e.g. see f in Fig. 2). A medal is attached to an instance in Career events graph using the **:relates\_to\_medal** property. Altogether 1844 people are mentioned with a medal.

### Bio CRM schema

Graph: <http://ldf.fi/schema/bioc>

Contains: 451 triples

Core classes: subclasses of **bioc:FamilyRelationshipRole**

Properties: **bioc:inheres\_in**, **inverse\_role**, **schema:gender**

This graph contains the definitions of the core classes and properties of the Bio CRM schema. It also includes the domain specific definitions of 45 subclasses of **bioc:FamilyRelationshipRole** (e.g. family members **:Child**, **:Mother**, or **:Father**) and 24 subproperties of **:involved\_in** (e.g. publication, design project, or nomination).

### 3.3 Linkage to LOD cloud

The Norssit data is linked to external data clouds shown in Table 2. The linkage was done with string comparison using person's full name with known dates of birth and death. Links were created to Wikipedia, Wikidata, National Biography of Finland<sup>12</sup> and its Swedish complement BLF<sup>13</sup>, BookSampo Linked Data, CultureSampo<sup>14</sup> portal, WarSampo<sup>15</sup> portal, ULAN<sup>16</sup> authority register by The J. Paul Getty Trust, VIAF<sup>17</sup>, and the genealogical data service Geni<sup>18</sup>. For entity linking to databases offering a SPARQL endpoint, the tool SPARQL ARPA<sup>19</sup> was used. In cases where the database provides a REST API, like Wikipedia or Geni.com, a special Python script was used. A Python script was used also in the case of BLF, where the data was only available as a CSV formatted table. [8]

## 4 Evaluation

The results of information extraction, and external data linking are evaluated in this chapter. Evaluation was done by first choosing a random sample at size of  $N = 50$  or  $N = 100$  people, and then manually checking if the data extracted or linkage accomplished was correct.

The results in Table 3 are all literal values of such properties that each person should have. Results indicate whether the information was interpreted correctly or not. For the

<sup>12</sup> <http://www.kansallisbiografia.fi/english>

<sup>13</sup> <http://www.sls.fi/sv/projekt/blf-biografiskt-lexikon-finland>

<sup>14</sup> <http://www.kulttuurisampo.fi>

<sup>15</sup> <http://sotasampo.fi/en/>

<sup>16</sup> <http://www.getty.edu/research/tools/vocabularies/ulan/>

<sup>17</sup> <http://www.viaf.org>

<sup>18</sup> <http://www.geni.com>

<sup>19</sup> <http://seco.cs.aalto.fi/projects/dcert/>



| Data Source        | Links | Description   |
|--------------------|-------|---|
| Geni               | 894   | Family research and family tree data                          |
| Wikipedia          | 609   | <a href="http://fi.wikipedia.org">http://fi.wikipedia.org</a> |
| Wikidata           | 609   | <a href="http://www.wikidata.org">http://www.wikidata.org</a> |
| CultureSampo       | 453   | LOD from museums, archives, libraries, and media              |
| WarSampo           | 352   | Second World War LOD service and portal                       |
| National Biography | 136   | National Biography of Finland                                 |
| VIAF               | 135   | Virtual International Authority Files                         |
| BookSampo          | 90    | Finnish fiction literature on the Semantic Web service        |
| BLF                | 44    | Biografiskt Lexikon för Finland                               |
| ULAN               | 16    | Union List of Artist Names Online                             |

**Table 2.** The data sources linked to the Norssit register.

property of name, the single false result was caused by an error in the OCR process, which caused an erroneous family name in the dataset. In the register book, the dates of birth and death were written in format *dd MM yyyy* with the month in Roman numerals (see b and h in Fig. 2). The two false results were caused by a typical OCR problem of mixing up characters *l*, *1*, and *I*. The year of enrollment was annotated in two digit decade and year format (c in Fig. 2, e.g. 83 for 1883 or 1983), and the century was automatically reasoned based on the person's birth year.

The Norssi high school has had female pupils only after the year 1972; approximately 11 per cent of people in the data set are female. For pupils enrolled 1972 or after, the gender was generated in three steps. First, depending on the known family relations, some people were marked male or female. Next, given names of people remaining without a specified gender were compared to the given names of people with known gender, and people with matching names inherited the corresponding gender. Finally, a list with less than 100 otherwise unidentified rare names was filled manually.

| Description        | Property         | Correct | False | Precision |
|--------------------|------------------|---------|-------|-----------|
| Name               | skos:prefLabel   | 49      | 1     | 0.98      |
| Date of birth      | schema:birthDate | 48      | 2     | 0.96      |
| Year of enrollment | :enrollmentYear  | 50      | 0     | 1.00      |
| Gender             | schema:gender    | 50      | 0     | 1.00      |

**Table 3.** Examples of the precision of the text retrieval.

Unlike in the previous examples, the properties evaluated in Table 4 were not obligatory. So, like for instance for the date of death there are 10 true positive (TP) matches where the information was interpreted correctly, 38 true negative (TN) cases where

the data had no such a date, one false positive (FP) case caused by noise in the OCR-process, and one false negative (FN) where the information was not retrieved. The year of matriculation was an integer extracted from the text just like the year of enrollment. Profile images for each corresponding person (see e.g. Fig. 2) were located from the OCR'd layout in the XML-format.

Some of the properties were very sparse in data, so the sample size was increased to  $N = 100$  for the evaluation of hobbies, family relations, careers, and links to LOD cloud. The false results in the family relations were caused by misinterpreting certain words of the Finnish language (e.g. the word *Eno* (Uncle) is also a name of a village). The data of the family relationships was further evaluated with SPARQL queries checking some basic rules, e.g. a parent must be older than the child<sup>20</sup>.

The algorithm for linking to external databases (Wikipedia, Semantic National Biography of Finland [6], WarSampo [5], and Geni.com) was designed to prefer precision on the expense of a lower recall; e.g. to link entities only in cases when assured that they represent the same person. This linkage could not be done based on person's name solely, and required extra information like places and times of birth and death.

| Property                      | TP  | FP | TN | FN | Precision | Recall | F <sub>1</sub> score |
|-------------------------------|-----|----|----|----|-----------|--------|----------------------|
| Day of death                  | 10  | 1  | 38 | 1  | 0.91      | 0.91   | 0.91                 |
| Year of matriculation         | 15  | 2  | 29 | 4  | 0.88      | 0.79   | 0.83                 |
| Profile Image                 | 34  | 0  | 63 | 3  | 1.00      | 0.92   | 0.96                 |
| Hobbies                       | 98  | 0  | 0  | 2  | 1.00      | 0.98   | 0.99                 |
| Family relations              | 98  | 2  | 0  | 0  | 0.98      | 1.00   | 0.99                 |
| Careers                       | 150 | 0  | 0  | 19 | 1.00      | 0.89   | 0.94                 |
| National Biography of Finland | 4   | 0  | 96 | 0  | 1.00      | 1.00   | 1.00                 |
| WarSampo                      | 3   | 1  | 94 | 2  | 0.75      | 0.60   | 0.67                 |
| Wikipedia                     | 5   | 0  | 91 | 4  | 1.00      | 0.56   | 0.71                 |
| Geni.com                      | 12  | 0  | 81 | 7  | 1.00      | 0.63   | 0.77                 |

**Table 4.** Examples of the precision and recall of the dataset linking.

## 5 Discussion

### 5.1 Related Work

Our research group, Semantic Computing Research Group (SeCo) has produced several projects with actor ontologies: The National Biography of Finland, CultureSampo<sup>21</sup>,

<sup>20</sup> Example of a SPARQL query: <http://yasgui.org/short/rJI8CXyUb>

<sup>21</sup> <http://seco.cs.aalto.fi/applications/kulttuurisampo/>

BookSampo<sup>22</sup>, and WarSampo [7] datasets, which are all highly interlinked; and linked to the Norssit project as well. The source material in our project was in a clearly structured format, while *Van de Camp* [2] deals with information extraction from unstructured text. *Szekely et al.* [9] describe linking datasets of Smithsonian American Art Museum with DBpedia and the Getty Vocabularies.

CIDOC CRM includes a mechanism for representing the role of an active participant in an event, modeling it as a property of the property that states the participant (see CIDOC's **P14.1 in the role of**). There is a proposal for encoding CIDOC's properties of properties as RDF<sup>23</sup>, introducing new class for the property and auxiliary properties, which adds complexity to the data model. Simple Event Model (SEM) [4] is a general model for expressing events, with support for three alternative representations for roles, based on using a) **rdf:value**, b) reification, or c) named graphs. Standards for Networking Ancient Prosopographies (SNAP) project<sup>24</sup> has developed an ontology for representing personal relationships.

Bio CRM's approach aims for simplicity and compatibility with CIDOC CRM. The model supports expressing enduring unary roles and binary relationships without the need to model them in the context of an event.

## 5.2 Lessons Learned

In our dataset, some practices were simplified, like modeling of a person's birth with literal values of properties **:birth.place** and **:birth.time** instead of using the CIDOC CRM event **crm:E69.Birth**. These simplifications worked well in this case, and reduced the complexity of the data. Another similar case was modeling people's names as literal values instead of using a resource of the type **crm:E41.Appellation**.

## 5.3 Future Work

We will continue our work on modeling and publishing biographies with data publications dealing with Finnish Biographies<sup>25</sup>, and U.S. Congress Legislators<sup>26</sup>.

HISCO<sup>27</sup> [10] is a vocabulary of historical occupations and professions, which has a hierarchical structure. We are extending the HISCO vocabulary with Finnish labels, mostly extracted from Wikidata and WordNet<sup>28</sup>, some manually translated.

**Acknowledgements** Our work is part of the project Semantic Web Publications – Texts as Data Services (Severi)<sup>29</sup>, funded mainly by Tekes. The development of Bio CRM

<sup>22</sup> <http://seco.cs.aalto.fi/applications/kirjasampo/>

<sup>23</sup> <http://www.cidoc-crm.org/roles-in-the-cidoc%E2%80%9900crm-modelling-properties-of-properties>

<sup>24</sup> <https://snapdrgn.net>

<sup>25</sup> <http://www.kansallisbiografia.fi/english>

<sup>26</sup> <https://github.com/unitedstates/congress-legislators>

<sup>27</sup> <http://historyofwork.iisg.nl>

<sup>28</sup> <https://wordnet.princeton.edu>

<sup>29</sup> <http://seco.cs.aalto.fi/projects/severi>

was started in the EU COST project *Reassembling the Republic of Letters*<sup>30</sup>. Thanks to Vanhat Norssit for funding the digitization of the register and opening the data.

## References

1. ter Braake, S., Fokkens, A., Sluijter, R., Declerck, T., Wandl-Vogt, E. (eds.): *BD2015 Biographical Data in a Digital World 2015*. CEUR Workshop Proceedings (2015), <http://ceur-ws.org/Vol-1272/>
2. van de Camp, M.M.: *A Link to the Past*. Ph.D. thesis, Tilburg University (2016), <http://www.taalmonsters.nl/pdf/phd-thesis.pdf>
3. Doerr, M.: The CIDOC CRM – an ontological approach to semantic interoperability of meta-data. *AI Magazine* 24(3), 75–92 (2003)
4. van Hage, W.R., Malaisé, V., Segers, R., Hollink, L., Schreiber, G.: Design and use of the simple event model (SEM). *Web Semantics: Science, Services and Agents on the World Wide Web* 9(2), 128–136 (2011)
5. Heino, E., Tamper, M., Mäkelä, E., Leskinen, P., Ikkala, E., Tuominen, J., Koho, M., Hyvönen, E.: Named entity linking in a complex domain: Case second world war history. In: *Language, Technology and Knowledge*, June 19–20. pp. 120–133. Springer-Verlag (2017)
6. Hyvönen, E., Alonen, M., Ikkala, E., Mäkelä, E.: Life stories as event-based linked data: Case semantic national biography. In: *Proceedings of ISWC 2014 Posters & Demonstrations Track*. pp. 1–4. CEUR Workshop Proceedings (2014), <http://ceur-ws.org/Vol-1272/>
7. Hyvönen, E., Heino, E., Leskinen, P., Ikkala, E., Koho, M., Tamper, M., Tuominen, J., Mäkelä, E.: WarSampo Data Service and Semantic Portal for Publishing Linked Open Data about the Second World War History. In: *The Semantic Web – Latest Advances and New Domains (ESWC 2016)*. pp. 758–773. Springer-Verlag (2016)
8. Hyvönen, E., Leskinen, P., Heino, E., Tuominen, J., Sirola, L.: Reassembling and enriching the life stories in printed biographical registers: Norssi high school alumni on120-133 the semantic web. In: *Language, Technology and Knowledge*, June 19–20. pp. 113–119. Springer-Verlag (2017)
9. Szekely, P., Knoblock, C.A., Yang, F., Zhu, X., Fink, E.E., Allen, R., Goodlander, G.: Connecting the Smithsonian American Art Museum to the Linked Data Cloud. In: *Extended Semantic Web Conference*. pp. 593–607. Springer-Verlag (2013)
10. Van Leeuwen, M.H., Maas, I., Miles, A.: Creating a historical international standard classification of occupations an exercise in multinational interdisciplinary cooperation. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 37(4), 186–197 (2004)
11. Verboven, K., Carlier, M., Dumolyn, J.: A short manual to the art of prosopography. In: *Prosopography Approaches and Applications. A Handbook*, pp. 35–70. Unit for Prosopographical Research (Linacre College) (2007), <http://dx.doi.org/1854/8212>

---

<sup>30</sup> <http://www.republicofletters.net>