

Linked Death—Representing, Publishing, and Using Second World War Death Records as Linked Open Data

Mikko Koho, Eero Hyvönen, Erkki Heino, Jouni Tuominen, Petri Leskinen, and Eetu Mäkelä

Semantic Computing Research Group (SeCo),
Aalto University, Espoo, Finland
<http://seco.cs.aalto.fi/>, firstname.lastname@aalto.fi

Abstract. War history of the Second World War (WW2), humankind’s largest disaster, is of great interest to both laymen and researchers. Most of us have ancestors and relatives who participated in the war, and in the worst case got killed. Researchers are eager to find out what actually happened then, and even more importantly why, so that future wars could perhaps be prevented. The darkest data of war history are casualty records—from such data we could perhaps learn most about the war. This paper presents a model and system for representing death records as linked data, so that 1) citizens could find out more easily what happened to their relatives during WW2 and 2) digital humanities (DH) researchers could (re)use the data easily for research.

1 Introduction

Lots of information about the WW2 is available on the web¹. However, this information is typically meant for human consumption only. The underlying *data* is not available in machine-readable, i.e., “semantic” form for Digital Humanities research and use [5,3] and for end-user applications to utilize. By making war data more accessible our understanding of the reality of the war improves, which not only advances understanding of the past but also hopefully promotes peace in the future [8].

For the case of the First World War, the situation has started to change, with several projects publishing linked data on the web, such as Europeana Collections 1914–1918², 1914–1918 Online³, WW1 Discovery⁴, Out of the Trenches⁵, CENDARI⁶, Muninn⁷, and WWI LOD [12]. A few works have used the linked

¹ <http://ww2db.com>, <http://www.world-war-2.info>, [different Wikipedias](#), etc.

² <http://www.europeana-collections-1914-1918.eu>

³ <http://www.1914-1918-online.net>

⁴ <http://ww1.discovery.ac.uk>

⁵ <http://www.canadiana.ca/en/pcdhn-lod/>

⁶ <http://www.cendari.eu/research/first-world-war-studies/>

⁷ <http://blog.muninn-project.org>

data approach to WW2 data as well, such as [2,1], the Open Memory Project⁸, and WarSampo [8].

This paper discusses the publication and use of casualty (death) records as linked data, as one part of the larger WarSampo system. Here, a dataset of some 95,000 deaths in military action in the Finnish army is concerned. We first present the data, its modeling, the Linked Open Data (LOD) service, and interlinking the data with other WarSampo datasets. After this, two use case applications are presented: 1) analyzing the data for digital humanities research and 2) reassembling the biographical war history of individual soldiers and military units. The latter use case serves, e.g., laymen in trying to figure out what happened to their relatives in WW2. The WarSampo system⁹ was published on Nov 27, 2015 and has had tens of thousands of end users indicating a large public interest in such applications.

2 Dataset, Data Model, and Data Service

Information about all known Finnish casualties of WW2 has been gathered in a relational database at the National Archives. This database contains 94,696 records of people that fought on the Finnish side, and died in 1939–1945 in the Winter War, the Continuation War, or in the Lapland War, or died of injuries obtained in those wars.

For use in the WarSampo project, the casualty database was first converted to CSV format, which was then converted to RDF format. Because the objective was to develop interactive applications directly on top of the large RDF dataset, it was important to keep the amount of RDF triples as low as possible without losing information and still linking the death records to ontological concepts. Thus, a simple data model was created for representing the data as linked data.

The data model is based on the CIDOC Conceptual Reference Model (CRM) vocabulary, which is designed for information exchange and integration of various cultural heritage information [4]. Each death record is represented as an instance of the Document class (`crm:E31.Document`) of CIDOC CRM.

A metadata schema was created that defines the properties used to describe each casualty with the information from the original database. The schema consists of OWL properties which have `crm:E31.Document` as the domain. A list of the properties and their `rdfs:range` constraints are shown in Table 1. The namespace prefixes used in this paper are:

`:` <http://ldf.fi/schema/narc-menehtyneet1939-45/>
`crm:` <http://www.cidoc-crm.org/cidoc-crm/>
`skos:` <http://www.w3.org/2004/02/skos/core#>
`wat:` http://ldf.fi/warsa/actors/actor_types/
`wrank:` <http://ldf.fi/warsa/actors/ranks/>

⁸ http://www.bygle.net/wp-content/uploads/2015/04/Open-Memory-Project_3-1.pdf

⁹ Including a semantic portal in use at <http://sotasampo.fi> and the underlying LOD SPARQL service at <http://www.ldf.fi/dataset/warsa/>.

Table 1. Casualty metadata schema of all properties used for describing the death records.

Property description	Property name	Range
mother tongue	:aaidinkieli	:Aaidinkieli
occupation	:ammatti	xsd:string
principal abode	:asuinkunta	
first names	:etunimet	xsd:string
date of becoming wounded	:haavoittumisaika	xsd:date
municipality of becoming wounded	:haavoittumiskunta	
place of becoming wounded	:haavoittumispaikka	xsd:string
burial place	:hautapaikka	xsd:string
burial graveyard	:hautausmaa	:Hautausmaa
military unit	:joukko_osasto	xsd:string
military unit code	:joukko_osastokoodi	xsd:string
known military unit	:osasto	wat:MilitaryUnit
citizenship	:kansalaisuus	:Kansalaisuus
nationality at time of death	:kansallisuus	:Kansallisuus
date of becoming missing	:katoamisaika	xsd:date
municipality of becoming missing	:katoamiskunta	
place of becoming missing	:katoamispaikka	xsd:string
place of domicile	:kotikunta	
date of death	:kuolinaika	xsd:date
municipality of death	:kuolinkunta	
place of death	:kuolinpaikka	xsd:string
number of children	:lasten_lukumaeaerae	xsd:integer
perishing class	:menehtymisluokka	:Menehtymisluokka
marital status	:siviilisaaety	:Siviilisaaety
military rank	:sotilasarvo	wrank:Rank
last name	:sukunimi	xsd:string
gender	:sukupuoli	:Sukupuoli
municipality of birth	:synnyinkunta	
date of birth	:syntymaeaika	xsd:date
full name	skos:prefLabel	rdfs:Literal
WarSampo person instance	crm:P70_documents	crm:E21_Person

The default namespace corresponds to the casualty schema namespace. RDF Schema (RDFS), Web Ontology Language (OWL) and XML Schema namespaces are omitted.

In Table 1 there are a total of 31 properties that are used for describing the casualties. The properties are used only when there is a value for the property. Municipalities are currently linked to three distinct datasets, which is why their range is not defined. The place properties, which give a more specific place for the described events, are literals representing the place names of the original

data. Original text representations of military units are also preserved and a new property `:osasto` is added for linking to WarSampo military units.

The Simple Knowledge Organization System (SKOS)¹⁰ was used to define vocabularies to present the information found in the original database in RDF. The created SKOS vocabularies for describing the death records in the casualty dataset are listed in table 2.

Table 2. SKOS vocabularies for describing the death records.

Vocabulary	Number of concepts
citizenships	10
genders	3
graveyards	802
marital statuses	5
mother tongues	11
municipalities	632
nationalities	11
perishing classes	7

A graveyard vocabulary was created to describe graveyards around Finland, and is also linked to ontologies of Finnish municipalities. These municipalities include current municipalities as well as historical municipalities, as some graveyards are located outside current Finnish borders, and often only the historical municipality of the graveyard is known.

The dataset is published on the Linked Data Finland (LDF) [7] platform, where it is openly available¹¹ for use via a SPARQL endpoint, with the Creative Commons Attribution 4.0 license¹². The SPARQL endpoint¹³ serves all WarSampo data, and has distinct graphs for each separate dataset and a default graph which contains all data. A Fuseki¹⁴ SPARQL Server is used for storing and serving the linked data. The used URIs are dereferenceable and provide information about resources for both human and machine users.

3 Interlinking with WarSampo datasets

The RDF dataset has been enriched by linking it to other parts of WarSampo like military ranks, military units, information about people found in other sources, and municipalities of wartime Finland.

A figure displaying the external linking of the death records is shown in Fig. 1. Each casualty is linked to other related WarSampo datasets and to a common

¹⁰ <https://www.w3.org/2009/08/skos-reference/skos.html>

¹¹ <http://www.ldf.fi/dataset/narc-menehtyneet1939-45>

¹² <https://creativecommons.org/licenses/by/4.0/>

¹³ <http://ldf.fi/warsa/sparql>

¹⁴ http://jena.apache.org/documentation/serving_data/

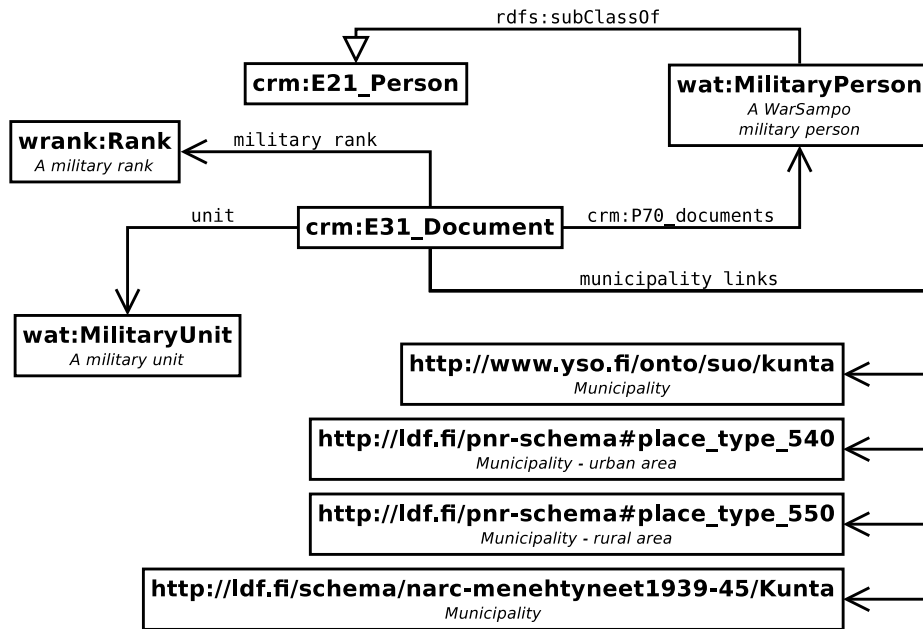


Fig. 1. External links from the death records, which are instances of `crm:E31.Document`. The properties are explained in Table 1. All WarSampo classes are aligned with the CIDOC CRM framework using `rdfs:subClassOf` relations.

WarSampo military rank ontology. The annotated military rank represents the rank of a person at the time of death.

Each death record is linked to a person instance (`crm:E21.Person`) of the WarSampo persons dataset via `crm:P70_documents`. So in our model, the death records are documents about the actual person. New information about people from other sources is not added directly to the casualty dataset, but to the WarSampo persons dataset, in order to maintain the integrity of the casualty dataset as a whole.

There are 3 municipality datasets, which include a dataset of contemporary Finnish municipalities, part of the Finnish Geographic Names Registry¹⁵, and two datasets that complement each other and consist of historical Finnish municipalities. These are the historical wartime municipalities registry [6] of WarSampo, and a municipality ontology based on the casualty dataset. The place properties are currently not linked to places available [6] in the Karelian map names in Finland and Russia, and the Finnish Geographic Names Registry. This is because finding the places to link automatically would result in a low precision due to the abundance of distinct places with identical names.

The death records are programmatically linked with people found in other WarSampo sources, which are gathered in the WarSampo persons dataset. The

¹⁵ <http://www.ldf.fi/dataset/pnr>

linking is implemented using the automatic annotation service ARPA [11], and a fuzzy logic algorithm to calculate a score for the similarity of two people based on each person's name, birth date, death date, and military rank. Source code for all automatic linking is available on GitHub¹⁶.

For each death record, the annotation service first generates a set of candidates with similar names. The candidates are then scored and if the score is above a given threshold, the persons are expected to be the same and are linked. Similarity score is given for string similarity for first names and last name, if they are similar enough, to allow somewhat differing spellings of names. A score is given for matches in military rank, birth date and death date, and subtracted in case they are not matching. In scoring, any of the values are allowed to be missing. In case of multiple matches for a death record, only the best match is used.

A `crm:P70_documents` relation is then added to the death record that will be linked, that points to the matching person. After this new person instances are created to WarSampo persons dataset for each death record that were not linked to an existing person.

We were able to automatically link 118 death records to people gathered from other data sources in the WarSampo persons dataset. The amount is quite low because the person information from other sources currently contains mostly information about high ranking officers and people who survived the war. The found person links have been manually validated and the discovered links seem to be depicting the same people. Manual validation was also done to person pairs that were close to the score threshold but not linked, and these seem to either not depict the same persons or not have enough information to make an assumption either way. However, as the scoring is manually adjusted to work well for the current persons dataset, when new people from new sources are added to the persons dataset, the scoring may need readjusting.

Military units of casualties are also programmatically linked to military units described in the WarSampo army units dataset, which contains military unit information found in other sources. The linking is implemented using the ARPA service and is based on unit abbreviations found in the casualty dataset, which are matched against manually annotated unit abbreviations in the army units dataset. As the exact abbreviation formats vary somewhat in different sources, multiple different abbreviation formats are generated from the original one for use in the automatic linking. Some 66,700 death records were linked to WarSampo military units, so this accounts for 70% of all the casualties. Currently, military unit information in WarSampo is limited to units of the Winter War. Therefore not all casualties are linked to the military units of WarSampo.

Municipalities in the data are linked automatically based on the labels of the municipalities. As shown in Table 1, there are six properties that relate to municipalities for each death record. The automatic linking leads to 98% of all death records having at least one link to the known wartime municipalities, which is the primary municipality dataset of WarSampo.

¹⁶ <https://github.com/SemanticComputing/Casualty-linking>

4 Use Case 1: Studying Death Records

This use case studies how the data could be used for prosopographical digital humanities research. We present the casualties perspective of the WarSampo portal, which is a tool for interactively analyzing the data in order to find patterns in groups of individuals.

The dataset graph consists of almost 2.4 million RDF triples. Presenting the data in an online service for users to search and browse is not straightforward due to the large size of the dataset. Furthermore, there are lots of links to related data in other WarSampo datasets (people, places, military units, etc.).

Faceted search provides effective support for interactive information-seeking in information systems [13]. A faceted search application was developed for searching and browsing the dataset. The application¹⁷ is part of the WarSampo portal, and provides the casualties perspective as one of the portal’s different perspectives. Faceted search is based on displaying categories for each facet, from which the user can select one, which then narrow down the result set to include only the results that match the user selections.

Fig. 2 shows a screenshot of the faceted search application in the casualties perspective. The data is laid out in a table-like view. Facets are presented on the left of the interface with string search support. The number of hits on each facet is calculated dynamically and shown to the user, and selections leading to an empty result set are hidden.

In Fig. 2, seven facets and the results are shown, where the user has selected “widow” in the marital status facet, focusing the search down to 278 killed widows that are presented in the table with links to further information.

The faceted search is used not only for searching but also as a flexible tool for researching the underlying data. In Fig. 2, the hit counts immediately show distributions of the killed widows along the facet categories. For example, the facet “Number of children” shows that one of the deceased had 10 children and most often (in 88 cases) widows had one child. If we next select the category “one child” on its facet, we can see that two of the deceased are women and 86 are men in the gender facet.

The application is developed in JavaScript as a Rich Internet Application (RIA) on the client side, using the open SPARQL endpoint to fetch data according to user selections. The application is open source¹⁸, and is based on our *SPARQL Faceter* tool [9], which is also open source¹⁹. When the user’s selections of the facets change, an asynchronous SPARQL query is sent from the user’s web browser to the SPARQL endpoint. The SPARQL endpoint returns results of the query to the user’s browser, which does additional processing of the data before displaying the new results to the user. The system works well even with the large casualty dataset, because pagination is used to limit the amount of results that are queried and displayed at a time.

¹⁷ <http://www.sotasampo.fi/casualties/>

¹⁸ <https://github.com/SemanticComputing/WarSampo-death-records>

¹⁹ <https://github.com/SemanticComputing/angular-semantic-faceted-search>

in the hierarchy are displayed on a lower level below the corresponding upper level categories.

Most of the facets are disabled by default, and the user has to click a plus sign on the facet to activate it. Activating facets makes the interface respond more slowly to user selections, as data for each activated facet has to be queried from the SPARQL endpoint based on user selections to show the facet categories. Normally when a user searches or browses the dataset, the facet categories and the results display are updated within a few seconds after the user has made a selection from one of the facets. With selections that have a large result set, and if additionally many facets are enabled, the user may have to wait more than ten seconds.

Another perspective of the WarSampo portal that makes use of the casualty data is the event perspective. The perspective displays wartime events on a timeline and map, as seen in Fig. 3. The casualty data is visualized by a heat map layer on the map, showing an overview of where casualties occurred during different time frames, and also which events happened nearby. The application also displays statistics regarding the casualties during the selected time frame: the total amount of casualties, and the amount per perishing class. People mentioned in the event descriptions are linked to the WarSampo persons dataset, and through this link to the casualties dataset for people that have died in the wars. The application provides hyperlinks to the linked entities shown in their corresponding perspectives. These other perspectives include applications for exploring places, photographs, military units, and magazine articles of the WarSampo system.

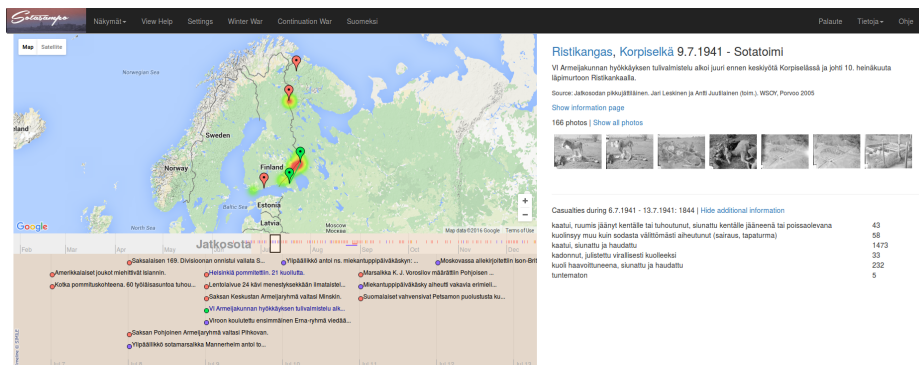


Fig. 3. The WarSampo event perspective with casualties of an 8 day time period visualized as a heat map on top of Google Maps, and casualty statistics in the bottom right corner.

5 Use Case 2: Reassembling Soldier Biographies and Military Unit Histories

This use case studies how we can reassemble soldier biographies and military unit histories based on the information content available in the casualties dataset and linked information in other WarSampo datasets. This use case serves citizens and researchers who are interested in finding information about a person's involvement in the war.

Linking the death records to information about the same people in other sources, events, military units, war diaries, photographs and wartime places provides new information about their activities, involvements in war events, whereabouts and movements during the war. By linking all these pieces of information together, we are able to construct partial biographies of individual soldiers, and the movements and actions of their military units. This allows an individual who is interested in investigating the biography of a relative who took part in the war to look at where the person probably fought, with whom, and when, and in what events his military unit participated. Also, the interlinked dataset, together with applications to effectively use it, provides digital humanities researchers with new perspectives to study the casualties, that would not be possible with a non-linked dataset.

Fig. 4 shows a histogram of the amount of casualties per day of a single military unit, the Second Battalion of the 38th Infantry Regiment, and all of its subunits, which consist of 4 companies. The battalion existed during the Winter War, which was fought from 30 November 1939 to 13 March 1940. The time span in the figure covers the time from first casualty to the last, with the exception of one death that occurred much later in 1940, supposedly due to injuries obtained in the war.

Demonstrating the value of linking additional data to the death records, we have information of 19 events that are linked to the military unit and its subdivisions. They seem to explain quite well the casualties during the Winter War, as high peaks in casualties mostly occur when the unit is engaged in an

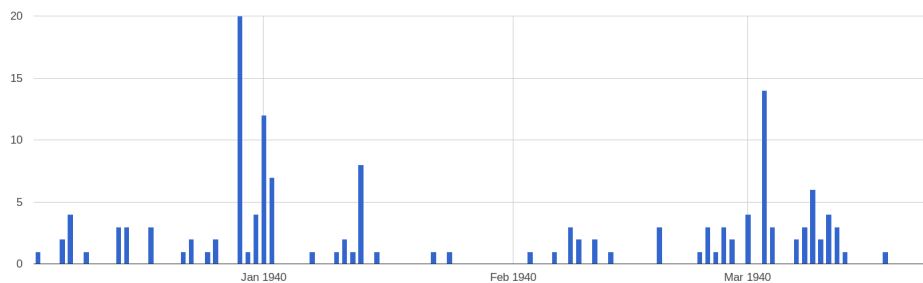


Fig. 4. A histogram visualization which shows casualties of the the Second Battalion of the 38th Infantry Regiment during the Winter War.

assault. However, the second highest peak occurs during a long defensive battle just before the end of the Winter War.

Events of the whole 38th Infantry Regiment during the Winter War are quite well covered in WarSampo. The regiment's second battalion and its subdivisions are known to have participated in the following events:

1. Defensive battle at Lavajärvi, 1939-12-06 – 1939-12-08
2. Battle at Lavajärvi, 1939-12-07 – 1939-12-08
3. Stalling battle at Lavajärvi-Lemetti, 1939-12-08 – 1939-12-10
4. Assault on Karjamökki, 1939-12-14 – 1939-12-14
5. Defense of Syskyjärvi sector, 1939-12-15 – 1939-12-28
6. Assault on Ruhtinaanmäki, 1939-12-29 – 1940-01-03
7. Assault on western Lemetti, 1940-01-06 – 1940-01-07
8. Assault battles at Repomäki, 1940-01-08 – 1940-01-12
9. Assault on Ruunaviita, 1940-01-13 – 1940-01-15
10. Battle at Koivuselkä, 1940-01-17 – 1940-02-06
11. Assault on Kehnovaara, 1940-01-18 – 1940-01-18
12. Assault on hill 63 and its defense, 1940-01-21 – 1940-01-24
13. Occupation of Pukitsanmäki and its defense, 1940-01-23 – 1940-01-26
14. Assault on Pujaski-Borisoff, 1940-01-25 – 1940-01-25
15. Destruction of Soviet elite ski unit at south of western Lemetti, 1940-02-06 – 1940-02-06
16. Capturing an encirclement northeast of Nietjärvi and destruction of a Soviet elite ski battalion, 1940-02-07 – 1940-02-09
17. Capturing 3 encirclements at Konnunkylä (Pujaski, Ahola and between railroad and road), 1940-02-18 – 1940-02-19
18. Capturing an encirclement south of point 26 (about 200 metres east of Koivusilta), 1940-02-21 – 1940-02-21
19. Defensive battle south of Nietjärvi, 1940-02-24 – 1940-03-11

The histogram is created by reading data directly from the WarSampo SPARQL endpoint and visualizing it with YASGUI²⁰ online SPARQL tool. The SPARQL query for retrieving the casualties for this military unit and its subdivisions is the following:

```
PREFIX atypes: <http://ldf.fi/warsa/actors/actor_types/>
PREFIX crm: <http://www.cidoc-crm.org/cidoc-crm/>
PREFIX casualties: <http://ldf.fi/schema/narc-menehtyneet1939-45/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

SELECT ?date (count(?cas) as ?casualties)
WHERE {
  { SELECT ?subunit
    WHERE {
      VALUES ?unit { <http://ldf.fi/warsa/actors/actor_972> } .
      ?unit (^crm:P144_joined_with/crm:P143_joined)+ ?subunit .
      ?subunit a atypes: MilitaryUnit .
    }
  } UNION {
    VALUES ?subunit { <http://ldf.fi/warsa/actors/actor_972> } .
```

²⁰ <http://yasgui.org>

```

}
?cas casualties:osasto ?subunit .
?cas casualties:kuolinaika ?date .
FILTER(?date < "1940-06-01"^^xsd:date)
} GROUP BY ?date ORDER BY ?date

```

All of the information about the Second Battalion of the 38th Infantry Regiment in WarSampo are available through the units perspective of WarSampo²¹. The units perspective visualizes the troop actions both on a map and a timeline, and shows the casualties of the unit as a heat map in the same fashion as in the event perspective. A screenshot of the perspective is shown in Fig. 5.

For even deeper understanding of the history, links to digitized images of the war diaries of the army units are provided, containing rich primary source descriptions of the events. By following the municipality links to the places perspective of WarSampo one can, e.g., study what kind of war events took place in the person's birth place, see photographs taken at specific locations the troops were located in, or read magazine articles depicting wartime events that took place in some specific place.

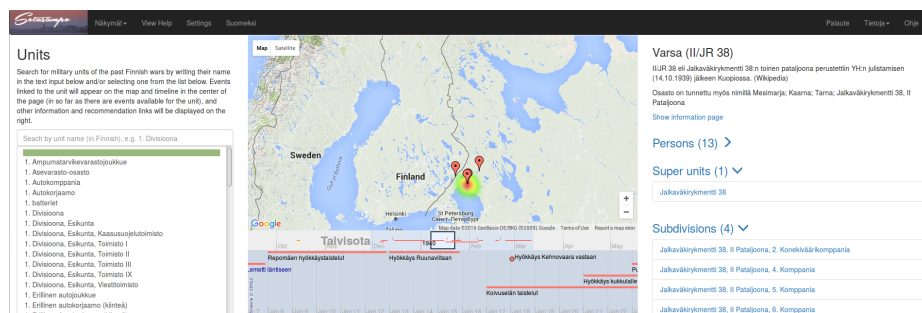


Fig. 5. WarSampo military units perspective displaying information about the Second Battalion of the 38th Infantry Regiment.

Fig. 6 depicts an integrated view of information related to a casualty in the person perspective of the WarSampo portal. On the left side is a search interface for finding people by name, and on the right is information about the currently selected person. Basic information about a person (e.g., name, birth and death dates and places, occupation, marital status, military rank with promotion dates if available) is displayed on the top. After that, thumbnails of the linked photographs involving the person are shown. By clicking the thumbnails the user can explore the higher resolution versions of the photographs and their captions. Below the photographs are war time events of the person, his military units, military ranks, municipalities where he is known to have been and a link to a Wikipedia page about this person.

²¹ http://www.sotasampo.fi/en/units/?uri=http://ldf.fi/warsa/actors/actor_972

For some people, there is also a biography text from the National Biography of Finland shown on the page, and possibly further information like linked magazine articles.

In order to get further context for the person examined, the user can browse the army units the person belonged to during the war, and places related to his life events (e.g., birth and death municipalities on historical and contemporary maps). This way the user is able to track the person's participation in the war by investigating the movements of his army units and the durations of the battles the units fought.

The screenshot shows the WarSampo portal interface. At the top, there is a navigation bar with the logo and menu items like 'Näkyvät', 'Suomenkieli', 'Palauta', 'Tietojä', and 'Ohje'. The main content area is titled 'Persons' and features a search box with instructions: 'Search for known persons from the past Finnish wars by writing their name in the text input below and/or selecting a person from the list below. Information regarding the person and recommended links will appear on the right.' Below the search box is a list of search results, with 'Niilo Juhani Korhonen' selected. The main profile for Niilo Juhani Korhonen (15.12.1915 Nilsä - 26.9.1943) is displayed. It includes his profession as 'opiskelija', marital status as 'naimisissa, 0 children', and service details: 'Served in 9./JR 10' and 'Luutnantti'. A source link is provided: 'Source: Mannerheim-ristin ritarit. http://www.mannerheim-ristinritarit.fi/ritarit'. There are 7 photos shown. The 'Units (2)' section lists 'Jalkaväkirykmentti 10, III Pataljoona, 9. Komppania' and 'Jalkaväkirykmentti 10'. The 'Ranks (1)' section lists 'Luutnantti'. The 'Places (3)' section lists 'Nilsä', 'Kuopio', and 'Kuopio'.

Fig. 6. A screenshot of the person perspective of the WarSampo portal depicting linked information related to a casualty.

6 Discussion

War history data is usually scattered in many isolated silos and may exist in totally different formats, e.g., books, paper archives, and databases. In this paper we have examined the benefits and challenges of linking casualty records of war to multiple related datasets, and publishing them as linked data for DH research and applications to use. Two use cases were studied related to supporting DH research and services for the public.

In the future, linking with other WarSampo data will be developed further as new datasets are added to the system. We plan to develop tools for statistical analysis of the data, and collaborate with humanities researchers in studying

how linked data and our tooling can help to solve their research problems. A hierarchical occupation ontology is planned to be used and linked to the death records to provide insight into the social status of each casualty. Whether we could take advantage of existing occupation taxonomies, such as the Historical International Standard Classification of Occupations (HISCO) [10], will be explored.

It would be beneficial for research purposes to develop the casualties perspective to allow exporting the data based on facet selections, for use with other applications and visualization tools.

Acknowledgements

Tomi Ahoranta and Jérémie Dutruit from the National Archives contributed in publishing the original casualty database as LOD.

Our research is partially funded through the Finnish Open Science and Research Initiative by the Finnish Ministry of Education and Culture.

References

1. de Boer, V., van Doornik, J., Buitinck, L., Marx, M., Veken, T.: Linking the kingdom: enriched access to a historiographical text. In: Proc. of the 7th International Conference on Knowledge Capture (KCAP 2013). pp. 17–24. ACM (2013)
2. Collins, T., Mulholland, P., Zdrahal, Z.: Semantic Browsing of Digital Collections. In: Proc. of the 4th International Semantic Web Conference (ISWC 2005). pp. 127–141. Springer-Verlag (2005)
3. Crymble, A., Gibbs, F., Hegel, A., McDaniel, C., Milligan, I., Posner, M., Turkel, W.J. (eds.): *The Programming Historian*. 2 edn. (2015), <http://programminghistorian.org>
4. Doerr, M.: The CIDOC CRM—an Ontological Approach to Semantic Interoperability of Metadata. *AI Magazine* 24(3), 75–92 (2003)
5. Graham, S., Milligan, I., Weingart, S.: Exploring big historical data. The historian’s macroscope. Imperial College Press (2015)
6. Hyvönen, E., Ikkala, E., Tuominen, J.: Linked data brokering service for historical places and maps. In: Adamou, A., Daga, E., Isaksen, L. (eds.) Proc. of the 1st Workshop on Humanities in the Semantic Web (WHiSe). pp. 39–52. No. 1608 in CEUR Workshop Proceedings, Aachen (2016), <http://ceur-ws.org/Vol-1608/#paper-06>
7. Hyvönen, E., Tuominen, J., Alonen, M., Mäkelä, E.: Linked Data Finland: A 7-star Model and Platform for Publishing and Re-using Linked Datasets. In: *The Semantic Web: ESWC 2014 Satellite Events, Revised Selected Papers*. pp. 226–230. Springer-Verlag (2014)
8. 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)

9. Koho, M., Heino, E., Hyvönen, E.: SPARQL Faceter—Client-side Faceted Search Based on SPARQL. In: Troncy, R., Verborgh, R., Nixon, L., Kurz, T., Schlegel, K., Vander Sande, M. (eds.) Joint Proc. of the 4th International Workshop on Linked Media and the 3rd Developers Hackshop. No. 1615, CEUR Workshop Proceedings (2016), <http://ceur-ws.org/Vol-1615/semdevPaper5.pdf>
10. van Leeuwen, M.H.D., Maas, I., Miles, A.: HISCO: Historical international standard classification of occupations. Leuven University Press (2002)
11. Mäkelä, E.: Combining a REST lexical analysis web service with SPARQL for mashup semantic annotation from text. In: The Semantic Web: ESWC 2014 Satellite Events, Revised Selected Papers, pp. 424–428. Springer-Verlag (2014)
12. Mäkelä, E., Törnroos, J., Lindquist, T., Hyvönen, E.: World War I as Linked Open Data (2015), <http://seco.cs.aalto.fi/publications/>, submitted for review
13. Tunkelang, D.: Faceted search. Synthesis lectures on information concepts, retrieval, and services, Morgan & Claypool Publishers (2009)