Nordic Spatial Humanities: Ups and Downs in LOD Implementation across Humanities' Digital Spatial Research Infrastructures in the Nordic Countries

Alexandra Petrulevich, Henrik Askjer, Sara Ellis Nilsson, Peder Gammeltoft, Andreas Lecerof, Emily Lethbridge, Øyvind Liland Gjesdal

Abstract

The article constitutes a report of a LOD application attempt undertaken within the humanities' spatial research (SRI)/spatial data infrastructure (SDI) sector. The case study is carried out on the geocoded data, mostly place-names and place-name attestations, of the four chosen Nordic SRIs: Icelandic Saga Map, Mapping Saints, Norse World and Norwegian Place-names. Ontologically, the case study aims at the implementation of Linked Art Data Model across the four resources. Methodologically, the SRIs data went through cleaning, transformation and augmentation stages. The results section demonstrates that the outcomes of the LOD implementation and the test querying have been uneven which is explained partially by the differences between the SRIs as well as by time constrains. In the discussion, the article reflects on possible alternative methodology, technological challenges such as scalability as well as its contribution to the field.

Keywords: Spatial Research Infrastructures, Linked Open Data, Nordic Spatial Humanities, Linked Art Data Model, geocoded data, place-names, Sampo-UI

Acknowledgements

This work was supported by NordForsk under Grant 345105. The authors would like to express their gratitude for the input of Johan Åhlfeldt, formerly of the University of Gothenburg, Pétur Húni Björnsson, University of Copenhagen, formelt of the The Árni Magnússon Institute for Icelandic Studies, Kristina Neumüller and Björn Karlsson, Institute for Language and Folklore, Sweden, as well as the participants of the Nordic Spatial Humanities workshops.

Introduction

Spatial Research Infrastructure (SRI) studies can at present be viewed as a rapidly developing and maturing interdisciplinary sub-field of applied research in digital (spatial) humanities (cf. Petrulevich & Skovgaard Boeck 2023a, also Bowker et al. 2010 on information infrastructure studies, Golub & Liu on information and knowledge organisation). The field originates in the so-called spatiotemporal turn in the SSH sector and the more mainstream use of geocoded humanities data and a variety of GIS-supported spatial analyses associated with it (Bodenhamer et al. 2010, Gregory & Geddes 2014, Tally 2014, Murrieta-Flores & Howell 2017, Dunn 2019 to name just a few). The field's development converges with the expansion of its empirical base: a steadily growing number of spatiotemporal infrastructural resources in academia and the cultural heritage sector. These infrastructural outputs—databases, interactive maps and other visualization means, APIs as well as associated data and metadata

principles, analysis scripts, websites, named-entity recognition tools, etc.—are meant to support relevant communities, i.e. researchers, professionals, and the interested general public alike, through making basic research and large amounts of spatiotemporal data and metadata openly and freely available for further analyses and reuse. The key goal of the field is to establish patterns of sustainable and balanced SRI design and implementation which can reconcile the agendas of any specific project on the one hand, and the Open Science and FAIR principles of data management on the other. The methodological path towards the goal goes through examination of the already available infrastructural outputs as well as through development, testing and implementation of technological solutions for interoperability, reusability, and sustainability of the already available or planned SRIs.

Interest organizations, funders, and policy makers offer multiple definitions of (spatial) research infrastructures in the humanities. Here, we follow the broad definition introduced in Petrulevich & Skovgaard Boeck 2023b (p. 4): "an SRI can be very broadly defined as a digital resource used for research on location-based data in spatial humanities and other fields."[1] In addition, it should be mentioned that there are two mainstream terms in use within the spatial humanities community, *spatial research infrastructures* (SRIs) and *spatial data infrastructures* (SDIs). Essentially, the terms are synonymous and can be used interchangeably. The former concept emphasizes the research focus of the outputs, while the latter reinforces affinity with various spatiotemporal data repositories in industry, government, and academia (cf. Petrulevich & Skovgaard Boeck 2023b: 4f.). In many contexts, as in "SRI/SDI studies" or "SRI/SDI design", both are used.

In Scandinavia, the field of SRI/SDI studies has recently reached critical mass through the accumulation of dozens of infrastructural outputs in different branches of the humanities and the cultural heritage sector. The underlying humanities materials are heterogenous and include literary sources, manuscripts and other artefacts, place-names, archaeological finds, maps, art, sound files, and more. This heterogeneity is one of the main reasons for why the currently available Nordic SRIs in the humanities differ from other regions with regard to approaches to data modelling, metadata, and re-use of data. Another major force driving these resources apart includes the somewhat different agendas and envisioned target groups of the research projects responsible for the design and development of the SRIs/SDIs in question. Of course, there are multiple common denominators as well: the spatiotemporal focus, the Nordic scope and the ambition to develop novel ways to aggregate and interrogate available data—to name just a few. The general picture of the SRI landscape in the Nordics can be summarized in the following way: novel, bottom-up research is encouraged within short-term spatiotemporal infrastructure projects. However, in the long run, these infrastructures are not sustainable.[2]

[1] The Petrulevich & Skovgaard Boeck 2023 volume includes several definitions of and offers a more nuanced discussion of the concept that e.g. includes interpersonal aspects.

[2] In the past couple of years, major steps have been taken by national Research Councils in the Nordics to ensure there are sustainable infrastructural solutions in the Humanities, cf. e.g. the Huminfra infrastructure in Sweden. At the EU level, there are a number of tailored humanities research infrastructures; however, these focus on either specific aspects, such as teaching and training, or making data available, e.g. DARIAH-EU, OPERAS-EU, or on specific disciplines or sectors such as archaeology or the cultural heritage sector, e.g. ARIADNEplus, Iperion HS. These efforts are crucial and much needed for the development of the related fields and research infrastructure studies in general. The process of establishing a best practice is however always reciprocal; this means that the bottom-up dialogue ideally should precede any prescriptive top-down recommendation. The article showcases possible results of such a dialogue.

The Nordic Spatial Humanities project (2022–2024; NordForsk under grant agreement 345105) was designed as a response to the pressing need for focused dialogue among researchers working on geocoded humanities data in the Nordic countries. One of the aims was to establish a formal collaborative initiative to map out common principles regarding ontologies and metadata. We anticipated that laying the foundations for projects to enrich or to build on long-term infrastructure would result in greater sustainability and data re-use over time. Additionally, we aimed at promoting the use of linked open data (LOD) technology in SRIs at the Nordic level, in line with FAIR principles. To attain these ambitious aims, the project organized three workshops to provide a discussion and cooperation platform for major stakeholders in the field of spatial humanities in Scandinavia, as well as to carry out a case study to link four Nordic SRIs[3] presented in detail in the following section. The workshops were attended by researchers and research engineers at universities and governmental agencies such as archives, museums and libraries, representatives of other research and teaching hubs at the national, Nordic and European level, e.g. DARIAH-EU as well as from industry, e.g. Google and Open Geospatial Consortium. Furthermore, the workshops facilitated dialogue with major international SRI initiatives such as the World Historical Gazetteer project in charge of developing the linked places format and the Australian Time Layered Cultural Map project and its associated data repositories.

The present article is based on the results of the Nordic Spatial Humanities project's case study conducted in Bergen and Reykjavík in September 2022 and May 2023, respectively. The main objective of the article is to outline common data and metadata principles of the four participating SRIs—the Icelandic Saga Map (ISM), Mapping Saints (MS), Norse World (NW) and The Norwegian Language Collections' Norske stedsnavn/Norske stadnamn, i.e. Norwegian Place-names (NPN)—and to develop a technological solution firmly rooted in the already available methodology to link these resources in order to run queries against and retrieve results from all the four SRIs. The four SRIs were chosen because these represent partially complementary resources dealing—at least in part—with relatively similar datasets. ISM, MS, and NW deal with project-specific token attestations in medieval Nordic texts,[4] while place-names attestations from multiple text and oral sources, including medieval spellings, are accounted for in NPN. Among these four SRIs, ISM and NW are the closest and could be labelled sister projects because both deal with place-name attestations in medieval literature, Old Norse and East Norse, respectively. The article poses three research questions: 1) what are the common denominators of the participating SRIs in terms of data modelling and metadata? 2) what type of data transformation produces the best results in terms of linking and overall conformity to LOD principles? and 3) what are the main challenges on the way towards full SRI interoperability based on the outcomes of the Nordic Spatial Humanities case study?

The article consists of five sections. The previous, introductory section contains the background and an overview of the article's main objectives and research questions. These were described in a wider context connected to general SRI studies. The following sections contain detailed descriptions of the participating SRIs and discussions of the methodology and research questions grounded in the Nordic Spatial Humanities case study. In the first, "Participating Nordic SRIs", the four chosen resources and their approaches to data

[3] From the start, the Swedish Digital Place-name Register was included in the case study, but time constraints prevented this group from partaking fully in the development work.
[4] However, MS also includes other source material, such as images and modern, newly digitized indices of medieval art historical material, as well as previously digitized registers for places, e.g. parishes, and archaeological material.

modelling, metadata and LOD are introduced. In the following section, "Towards and beyond CIDOC-CRM", the generalized framework and concrete methodology of the case study are outlined. Third, the case study's results are presented and discussed in the section, "The Nordic Spatial Humanities ups and downs in LOD implementation". Finally, the article's conclusions and contribution to the fields of SRI studies, and more generally to spatial and digital humanities are summarized in "Concluding remarks".

Participating Nordic SRIs

This section presents the four Nordic SRIs chosen for the case study, the Icelandic Saga Map (ISM), Mapping Saints (MS), Norse World (NW), and Norwegian Place-names (NPN). The overviews are structured in the following way: after the outline of aims, materials, scope, and target groups, a more detailed accounts of the resources' data modelling frameworks or principles and metadata are provided. Additionally, the SRIs' LOD "readiness", including LOD implementation and LOD compatibility, is discussed.

Icelandic Saga Map

The *Icelandic Saga Map* (ISM) project is now around a decade old. From its inception, the aim was to provide readers of the sagas with a way of exploring these narratives on a spatial basis, in order to better see and understand how individual saga narratives overlap and intersect geographically in Iceland and beyond (see further Lethbridge 2020, 2023). Since the *Íslendingasögur* are rooted in place and set all around Iceland, the place-name data preserved in these narratives is extensive and can be utilised in many different research contexts (e.g. literary criticism, environmental history, place-name studies, archaeology, folkloristics). The core data of ISM are GPS co-ordinate sets comprising point data linked to place-names that occur in the saga texts. The great majority of these are Icelandic place-names but since the geographical scope of the sagas is wider, place-names around Scandinavia, the British Isles and beyond are also included. In addition, place types are defined, making it possible to conduct analysis of, for example, the distribution of place-names and place types for any single saga or combination of sagas. Texts of sagas belonging to the *Íslendingasögur* were geo-referenced first, from 2015 onwards; later, other saga genres and types of texts have been added (e.g. *Landnámabók*, Icelandic annals, *Færeyinga saga*, *Orkneyinga saga*, *Sturlunga saga* and other contemporary sagas, *Íslendingaþættir*, and a selection of 19th- and 20th-century English-language travel-books that focus on saga-sites). Also in the ISM database but not currently visible on the website are data about Icelandic manuscripts that preserve saga texts (including the manuscripts' dates, contents, and the names of places and individuals that any given manuscript has been associated with), as well as English-language translations of Icelandic place-names, and finally, environmental concepts that build on the CIDOC-CRM ontology and have been mapped onto saga texts as part of work conducted for the NSF-funded *DataArc: Linking Data from Archaeology, History, Sagas and Climate* project (see further Pálsson and Opitz 2019; dataARC).

The ISM platform builds on a Django Python web framework, in line with other digital database resources managed by Árnastofnun (e.g. Árnastofnun/Rannsóknir og gögn). Data is held in a relational (PostgreSQL) database, see Figure 1). The great potential to link the ISM geo-spatial data with other comparable datasets (whether pertaining to manuscripts, written texts, archaeological sites, place-name records, environmental proxy-data) has been clear from the outset.

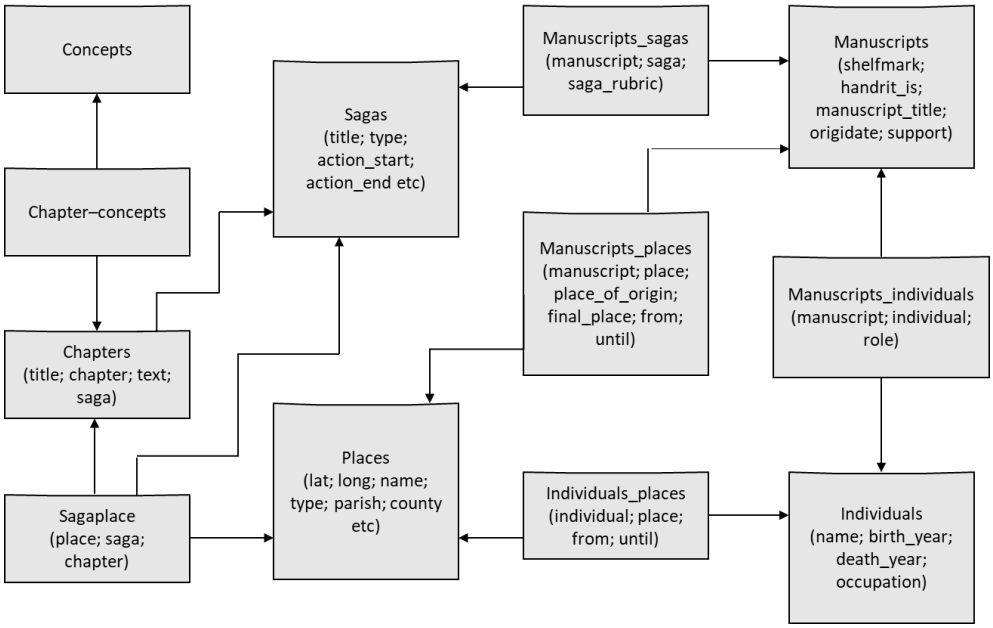

Figure 1. Entity relationship diagramme of ISM backend. Image by Pétur Húni Björnsson and Emily Lethbridge.

Mapping Saints

The project, *Mapping Lived Religion: Medieval Cults of Saints in Sweden and Finland*, is currently developing the online, open-access, research resource, *Mapping Saints* (MS). Set to launch in 2024, the resource was still under development during the duration of the *Nordic Spatial Humanities* project, which presented a challenge as to how it could be used in the tests to link already existing spatial research infrastructures. Although the data was not yet complete, the structure of the relational database and the application of the principles of linked open data had been defined previously in the initial development stages of the project.

MS uses and links to data from previous projects and to national authorities—including spatial data such as parish codes, coordinates of archaeological finds, and remains of buildings. The project also identified and determined coordinates for medieval places mentioned in analogue medieval and early modern sources—all of which are accessible via the interface. The project goal is to map material and textual remains of the cult of saints in Sweden and Finland in the period 1164–1593, i.e. the timescape associated with the medieval ecclesiastical province of Uppsala, from its establishment until the consolidation of the Protestant Reformation in Sweden. This region covered roughly the same area as present-day Sweden, minus the provinces of Blekinge, Skåne, Halland, Bohuslän, Härjedalen, and some parts of Dalarna, but instead including Finland and the Karelian Isthmus (now part of the Russian Federation).

The research platform is modular, which means that in the future we can build on, combine and add functions (Ellis Nilsson et al. 2023). This choice was made to make it possible to add other regions to the map in order to study more of Scandinavia. Indeed, by linking to national repositories for example for coordinates for churches, many of the “unwanted” areas were initially included on the map. Although one person was responsible for the practical side of the digital development, the research platform was a collaborative effort between the project and its partners. It has involved importing, exporting, editing, analyzing, and publishing data.

As for geographical data, it has been necessary for the project to use different datasets in order to compile a more complete dataset for the project, i.e.: 1) places associated with objects (e.g., churches, ruins, and church art), and 2) places in narrative texts (e.g., miracle stories and letters) (Ellis Nilsson et al. 2023, see also Figure 2). Much like with the other projects in the Nordic Spatial Humanities initiative, the development of the portal and its focus have been influenced by the underlining research question. In the case of MS, this focuses on lived religion as expressed via the cult of saints. Due to the varied nature of the source material, it was necessary to create an entity—termed *cult manifestation*—to organize searches for places with a specific, relevant content during a specific time period (Ellis Nilsson et al. 2023, 37–38).

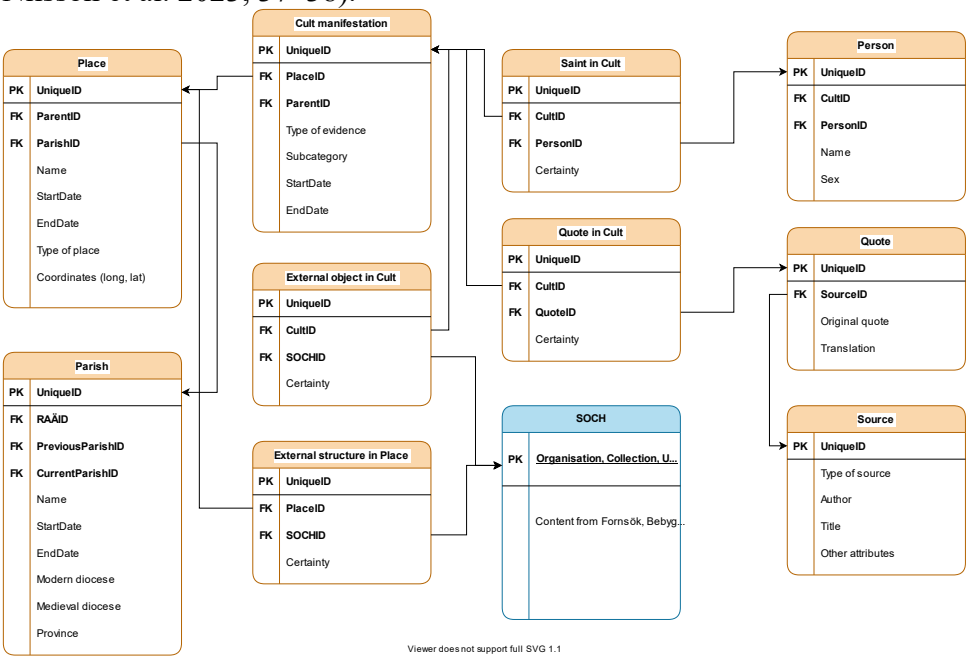

Figure 2. Model of the database for Mapping Saints. PK denotes Primary key and FK denotes Foreign key. (Reproduced with permission from Ellis Nilsson et al. 2023, Fig. 2.1. Image by Johan Åhlfeldt)

In terms of LOD, the project has also applied standards and classification systems from the start and worked within the FAIR principles as much as possible. These include the Getty Research Institute's vocabulary Art & Architecture Thesaurus (AAT) and the art historical classification system for images, *Iconclass*, developed by the Henri van de Waal Foundation. Moreover, Wikidata identifiers—a standard applied in cultural heritage institutions—have been used for a number of entities, e.g. for people and saints. In addition, a number of the digital sources that the MS portal links to have been published or will be published via the aggregator *Swedish Open Cultural Heritage* (SOCH, *K-Samsök*), maintained by the Swedish National Heritage Board. When launched, the resource itself will also link to cultural heritage data via SOCH.

Norse World

NW is described as "an interactive spatial-temporal resource for research into spatiality and worldviews in medieval fictional, non-biblical literature from Sweden and Denmark" (Petrulevich & Skovgaard Boeck 2023c:15). The resource has been developed to facilitate interdisciplinary spatial humanities research on medieval East Norse, i.e. Old Swedish and Old Danish, texts. The material in focus comprises so called foreign place-names and other spatial references, i.e. linguistic expressions referring to localities outside the current,

modern-day borders of Sweden and Denmark (Petrulevich, Backman & Adams 2020). Examples of potential research questions NW could assist researchers in answering are: What places were written about and where? Are some places more popular in certain text types or at certain times? How do place-names link different texts? Is there a shared concept of spatiality?

Described as an SRI, NW comprises above all a database of geocoded attestations of so-called spatial references attested in East Norse manuscripts, excerpted, processed and presented at a manuscript level. By spatial references NW understands place names and non-names, i.e. names of topographical, physical, and cultural features as well as other types of language resources conveying spatial "geocodable" information, e.g. adjectives, adverbs, inhabitant designations, language designations, etc. Additionally, the NW modular system contains a REST-API, an admin user interface and an interactive end user interface including a mapping tool. The SRI in question was conceived 2015 by the project team under the leadership of Jonathan Adams, the PI of the project The Norse Perception of the World: A Mapping and Analysis of Foreign Place-Names in Medieval Swedish and Danish Texts (funded Riksbankens Jubileumsfond under agreement IN16-0093:1). The bulk of the project work was undertaken 2017–2022 by an interdisciplinary team of philologists and developers (Petrulevich & Skovgaard Boeck 2023c:25f.). The work on the NW database that comprises over 6,600 attestations is still ongoing.

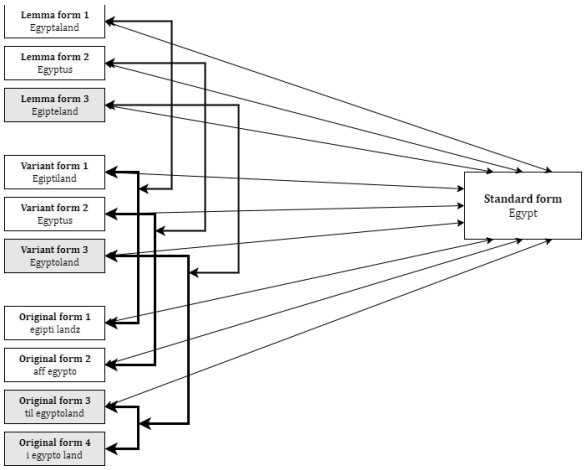

Figure 3. The place-name data model employed by NW. Original, variant and lemma forms in Old Swedish are enclosed in unfilled rectangles, while the corresponding material in Old Danish is placed in filled rectangles. Different line types are used to represent links between the different forms; thick lines link original and variant forms, mid-thick lines link variant and lemma forms, and thin lines link original, variant and lemma forms with the corresponding standard form. Reproduced with permission from Petrulevich 2023, Figure 11.1.

The NW data modelling principles have to be developed at least in part from scratch because the standard gazetteer model typically used in SRIs/SDIs in academia or elsewhere does not have the functionality to account for spatial attestations processed at a manuscript level (cf. Petrulevich 2023). In a typical instantiation of a digital gazetteer model, one set of coordinates, one feature type and one name attribute are used to represent a unique geographical location. Compared to this brief outline, the main NW innovations include the introduction of several layers or normalization, so-called variant forms and lemma forms, in addition to attestations or the so-called original forms that represent excerpts from actual manuscript material, see Figure 3. The lemma form serves as a linguistic umbrella for

multiple attestations of the morphologically the same name, cf. the unique place-name expression IDs developed by the Norwegian Language Collections below. These forms, the original one and the two normalized ones, are then tied to the so-called standard form that has multiple functions in the SRI. In this context, the most important aspect is that the standard form serves as a "link" of the NW resource to any other SRI by "translating" the NW bespoke data model to, essentially, the standard gazetteer structure with coordinates, feature types and name attributes, see Figure 2 (cf. Petrulevich 2023). Additionally, the resource includes multiple metadata on all attestations and standard forms, manuscript material or "sources" (NW/Source) as well as texts or "works" (NW/Work).

The NW SRI did not support LOD from the beginning, for two reasons. Firstly, the resource was conceived in the early days of LOD as a mainstream research infrastructure prerequisite, before the FAIR principles were launched. This implies lack of any relevant previous experience or training on behalf of the project team. Secondly, the need for bespoke data modelling structure suitable for the niched scope of the project impeded early attempts to introduce LOD based on the LOD ontologies known to the team at the time (Backman & Smith 2023: 174–176). The Nordic Spatial Humanities collaboration offered the much-needed opportunity to implement LOD, at least as far as the chosen framework and method allowed, see below. NW also provides a basic REST-API where the attestation material is available in the JSON-LD or GeoJSON formats.

The Norwegian Place-names SRI

Place-name governance is one of the central functions within the Norwegian administration. The Norwegian Language Council (Språkrådet), the Norwegian Mapping Agency (Statens kartverk) and the Norwegian municipalities are all involved in this work. In addition, the place-name experts at Statens kartverk are decentralized at regional mapping centres, so Norwegian place-name standardization is very much in need of digital resources in order to undertake essential work. The Norwegian Place-Names (NPN) aims to be a one-stop go-to-service for experts involved in standardizing Norwegian place-names—at the same time as being an interdisciplinary research infrastructure. Currently, place-name research is not covered as a subject at any of the Norwegian universities, but individual place-name researchers as well as historians and archaeologists are active within the field.

The purpose of the NPN SRI is to make all digital and digitized place-name datasets available online to cater for use in decentralised fashion at all Norwegian universities (and beyond) and in the Norwegian government sector. This is an ambitious and challenging task. For one thing, not all digitizations are similar to each other. Furthermore, it is a time- and resource-consuming exercise to enrich earlier digital datasets and transform them into usable geodata, as well as into a tool for place-name research. At present, all datasets in NPN have been expanded to include unique IDs for each place-name expression (following the principle outlined in Gammeltoft 2023, also cf. lemma forms in NW above) and an ID for administrative location. Both IDs are designed to work across datasets. In addition, most datasets have also been provided with coordinates.

A focal technology of the Bergen University Library, where the NPN SRI is housed, has until recently been LOD/RDF knowledge graphs (see the Framework section below). In addition, the library's strategy is to work towards open data and use open standards. Thus, the focus early on was to use the Finnish Sampo framework, see below, developed by the Semantic Computing Research Group, Aalto University and the University of Helsinki, Finland, to test

how place-name data performed in a LOD environment. Initial tests were positive and a full prototype based on the Finnish place-name portal NameSampo was developed.

At the time of the Reykjavík workshop, the NPN SRI included two legacy systems in addition to the NPN Sampo portal, Hordanamn, "Place-names of Hordaland", and The National Place-name Database (formerly maintained by the University of Oslo). These were being integrated as datasets into a planned new version of the NPN Sampo portal. Since NPN as a whole contains around three million name entries, we chose to limit the data for the workshop to the Oslo dataset, henceforth called "the NPN dataset". The unique IDs for place-name expressions mentioned earlier were not yet an integral part of the NPN user interface, and they were not included when the NPN dataset was mapped to RDF in preparation for the Reykjavík workshop. Furthermore, the data already added to the NPN Sampo portal conformed to a slightly different ontology than the one envisioned for the workshop. For this reason, it made sense to do the mapping from scratch using the new NPN dataset.

Towards and Beyond CIDOC-CRM: the Framework and the Method

As outlined in the previous section, the four SRIs being studied serve complex and heterogenous geocoded humanities data. Only two of these (MS and NPN) are LOD-compatible. All of the SRIs include idiosyncratic features to a larger or smaller degree. The single common denominator uniting all of the resources is the so-called spatial component or, in more concrete terms, geocoded data. ISM, NW, and NPN can be seen as more similar or compatible, because these model and serve place-name data. In other respects, such as what (other) data types are included, how they are modelled, described, and made available to end users, there is much variation. In order to make these resources interoperable—at least to some degree—the Nordic Spatial Humanities project had to devise a bespoke overall ontological framework as well as define concrete methodological steps with respect to data cleaning, data alignment, and data augmentation.

The Framework

The Nordic Spatial Humanities case study framework has multiple sources of inspiration and thus in a way resembles an ontological patchwork. The main ones described below are the cultural heritage conceptual model CIDOC-CRM and two practical ontological adaptations of it, the Linked Art Data Model and Erlangen CRM, Resource Description Framework (RDF), Simple Knowledge Organization System (SKOS) and Sampo-UI. The framework has been continuously developed throughout the project which implies that some initially chosen solutions were replaced by others. For instance, two ontologies, Erlangen CRM and the 7.1.1 version of the CIDOC-CRM (launched May 2021) were implemented subsequently to reach the same target, i.e. to practically apply the Linked Art Data Model profile of the CIDOC Conceptual Reference Model (see the Method section).

The CIDOC-CRM is described as "a theoretical and practical tool for information integration in the field of cultural heritage" and a "living standard" that is constantly under development by the CIDOC-CRM Special Interest Group (CIDOC CRM). This general ontology has a specific node, E53 Place, dealing with spatial information, last updated in Version 5.1.2, as well as pre-defined relations or properties such as P53, "has former or current location (is former or current location of)", between the said node and other related ones such as E18 Physical Thing, see Figure 4. As becomes evident in Figure 4, the Place component is identified by (the P87 property) E44 Place Appellation that in its turn includes e.g.

coordinates and place-name information, cf. the description of the digital gazetteer model in the previous section.

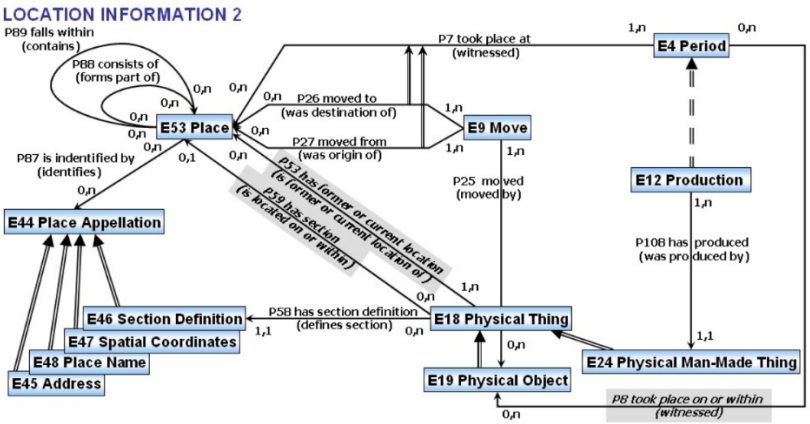

Figure 4. The CIDOC CRM principled for modelling of location. See CIDOC CRM/Place: Image. The image has been cropped.

The Place ontology briefly outlined here represents a theoretical construct that is too general for the concrete task of linking the Nordic Spatial Humanities datasets. For this reason, it was decided to pick a CIDOC-CRM instantiation or profile that was deemed best suited for the task at hand, the Places component of the Linked Art Data Model (Linked Art Data Model/Places). This Places component elaborates on and exemplifies the CIDOC CRM schema, see Figure 5. "Amsterdam" can thus be described as a Place that has a type, "city", a name "Amsterdam", and a description. The place in question can also be related to other places such as the Netherlands. The ambition of the Places component is also to offer GIS integration with external gazetteer services in order to provide coordinates for the Places data.

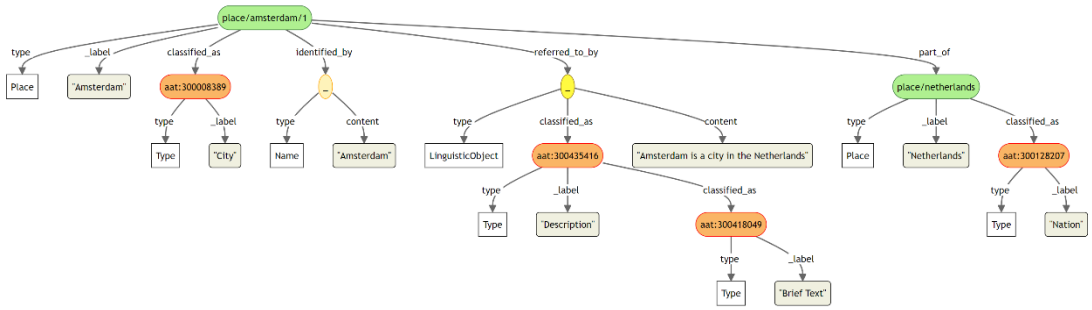

Figure 5. The Linked Art Data Model Place component (Linked Art Data Model/Places).

This initial step of getting accustomed to CIDOC-CRM and its possible practical implementations also included early applications of another CIDOC-CRM ontology, Erlangen CRM, in order to map the Linked Art Model (see above).

Resource Description Framework (RDF) is a standard model to describe interconnected data on the web through three-part statements called triples or edges (W3C/RDF, Backman & Smith 2023:161f.). These statements define relationships between things and thus comprise subjects (so-called source nodes), predicates (so-called edge names) and objects (so-called target nodes; W3C/RDF, Stardog/RDF), cf. "… the *subject*, typically identified by its IRI [Internationalized Resource Identifier] or following on from another statement—is described using attributes or properties—the *predicate*, given using an IRI—that take either string values, an IRI, or a reference to the subject of another such triple: the *object*" (Backman &

Smith 2023:160). The RDF descriptions were developed by the World Wide Web Consortium to facilitate data sharing and interoperability. These formal descriptions are then used to create RDF knowledge graphs defined as sets of RDF triples (Stardog/RDF). In order to build knowledge graphs, one needs to explain the relationships between the nodes in the triples through authoritative vocabulary called the RDF Schema (W3C/RDFS). RDFS provides effective categorization functionality through its two concepts, classes, comprising nodes with similar characteristics, and properties, explaining relationships between subjects or source nodes and objects or target nodes (Stardog/RDF). The CIDOC Conceptual Reference Model described above can be characterized as an "object-oriented metamodel" independent from data definitions and encoding languages such as RDF (e.g. Doerr et al. 2020). Theoretically, the Nordic Spatial Humanities project could have picked a model other than LOD for the implementation. The rationale behind the choice of RDF builds on the ambition of the project to use current best practices for LOD implementation—and thus RDF (cf. Backman & Smith 2023)—as well as on the fact that RDF is the framework deemed closest to CIDOC-CRM (cf. Doerr et al. 2020).

As for vocabulary definitions, the Nordic Spatial Humanities project followed the Linked Art Data Model implementing Getty vocabularies (Getty) available through the linked.art profile. However, it turned out these do not meet the needs of the project, because, for instance, a comprehensive list of possible place types does not yet seem to exist (cf. also the description of the CIDOC RDF Schema for recording places and names as in Doerr et al. 2020). The Getty vocabularies were thus complemented by the Nordic Spatial Humanities SRIs' types defined as SKOS concepts. The SKOS principles for definitions based on the SKOS data model operate with two basic units, classes and properties, and include logical interdependencies between these units. SKOS concepts are RDF-compatible with RDF and can be expressed as RDF triples.

The Sampo-UI (Ikkala et al. 2021) is a JavaScript front-end application that can be configured to query and visualize semantic data from SPARQL endpoints. It is based on the principles of the Sampo Model defined by Eero Hyvönen (2023:730f.) as "an informal collection of principles for LOD publishing and designing semantic portals … supported by an ontology and data infrastructure and software tools for user interface design and data publication". The Sampo main principles can be summarized as collaboration support through shared open ontology accounting for multiple perspectives on data and based on the ideas of a clear-cut distinction between the LOD service and the user interface (Hyvönen 2023:731–734). The Sampo-UI includes two different approaches, the "server-side faceted search" or ServerFS sending the queries to a single triplestore and "client-side faceted search" or ClientFS allowing multiple triplestores (Ikkala et al. 2022:75), see Figure 6.

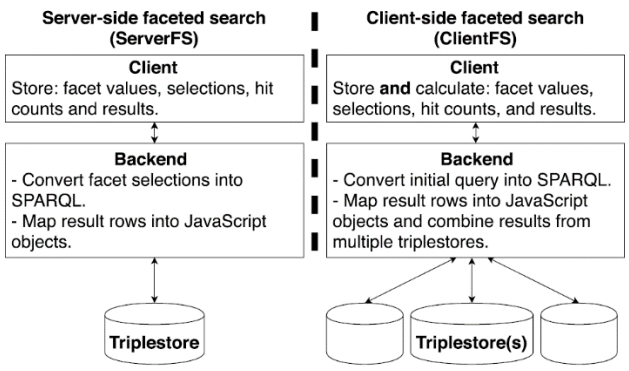

Figure 6. Server-side faceted search and Client-side faceted search in Sampo-UI (reproduced with permission from Ikkala et al. 2021: Fig. 4).

As pointed out above, the former approach is limited to using a single triplestore; moreover, it assembles modular SPARQL queries that are filtered on the server. ServerFS allows for multiple perspectives on the same data: you could for instance set up separate search perspectives for documents, places, people, etc. However, the approach puts more constraints on the way queries are assembled, and it requires a shared ontology. In the latter approach, the faceted search processing is preceded by conversion of result sets into JavaScript objects. This allows combining results from multiple triplestores and no shared ontology is required.

The Method

The general methodological idea of the Nordic Spatial Humanities case study is to clean, transform, and augment the datasets, as outlined in the section "Participating Nordic SRIs", using a single, dedicated tool throughout the process: OpenRefine. OpenRefine has been chosen for these tasks because the tool is open source and has been tested with good results for similar objectives and set-ups previously (e.g. Harlow 2015, Hardesty 2016, Miller & Vielfaure 2022). The resulting data is then expected to be aligned with RDF and, thus, interoperable and LOD compatible. The corresponding methodological steps of data preparation, cleaning, augmentation, and transformation are described below. Additionally, the final section demonstrates the method for setting up an interface for querying the resulting dataset.

The preparation stage included submitting data samples from all the chosen Nordic SRIs for initial check-ups conducted by the Bergen/NPN team. The datasets were submitted several months ahead of the Bergen workshop to make sure data import and parsing could be tested in advance. The data cleaning stage involved employing corresponding functionality in OpenRefine, e.g. the cluster and edit function, to check consistency of metadata, formatting and spelling. All the datasets turned out to be both consistent and clean, i.e. no extensive data cleaning was required for the present case study.

The data augmentation stage included identification of relevant data entities, e.g. Work or Standard form in NW, and assignment of persistent URLs. For this task, we used the OpenRefine functionality to generate version-3 Universally Unique Identifiers (UUIDs) by combining namespace identifiers or local IDs and names. The procedure was deemed the most reasonable in the situation with four different underlying SRIs and corresponding data models at hand. For the data transformation stage, a dedicated plug-in, RDF Transform, was applied to classify data against the main classifier, the Linked Art Model, the chosen application profile of CIDOC-CRM. At this point, the methodological process went through at least two stages: 1) initial implementation of the Erlangen CRM ontology, and 2) the application the 7.1.1 version of the CIDOC-CRM (launched May 2021) including dedicated RDFS vocabulary in order to carry out mappings against the Linked Art Model profile. Additionally, authority vocabulary for the case study, both from CIDOC-CRM and from SKOS, were added to the OpenRefine workplace. This lack of methodological stability of course introduced a risk of inconsistent vocabularies in use across the datasets. In order to ensure consistency, the Erlangen CRM prefixes were replaced by the CIDOC-CRM ones before the different SRI models were imported into a triplestore during the final workshop in Reykjavík.

The querying preparation stage preceding the Reykjavík workshop included setting up a cloud Stardog knowledge graph database to give shared access to the data as well as to enable querying. In the workshop, the team set up a Nordic Spatial Humanities demo website based on a modified version of the Sampo-UI. The modification allowed us to use a Stardog endpoint instead of Fuseki for the querying procedure. This allowed us to send out invites for user access to the participants for both administrative tasks such as uploads, as well as more advanced GUI tools for visualisations and querying. Finally, the team created two search queries per SRI, one geographical search and one text search.

The Nordic Spatial Humanities ups and downs in LOD implementation

The Nordic Spatial Humanities case study produced a functioning RDF output based on three of the four SRIs, ISM, NW and NPN, which is the major upside aspect of the project. However, the case study also showed that the implementation of the chosen framework and the envisaged method was not as straightforward as planned. Additionally, it turned out to be difficult to elicit meaningful results across the four datasets based on the pre-formulated search queries.

The data augmentation step was tailored towards the four SRIs at hand and thus followed their respective more or less idiosyncratic data modelling structures. For NW, the unique URLs were created for four different types of entities, Attestations (NW/Attestation), Standard forms (NW/Standard form), Works (NW/Work) and Sources (NW/Source). Originally, the NW Attestation entity was enriched with a sophisticated linguistic component comprising two normalisation steps, cf. original forms, variant forms and lemma forms in the NW description above, which posited a problem, because the rest of the SRIs in fact lack any corresponding units. For this reason, the classification of the variant and the lemma forms of the NW Attestation entity was managed in an ad hoc way: empty but unique nodes for lemma and variant forms were created. The NW Standard form entity could be processed as the Locality entity of the rest of the SRIs, cf. discussion of the NW data modelling and its relation to the conventional gazetteer model.

The NW Attestation entities were classified as rdf:type “E33_Linguistic Object” within the CIDOC-CRM framework and defined through unique identifiers, see the example extract “til een stadh j grecia som kallas athenis” from the work Själens tröst (eng. Consolation of the Soul) in Figure 7. The NW Standard forms were classified as rdf:type “E53_Place” in the Linked Art Data Model, see the example “Athens” in Figure 7.

```
nw:cc33be71-7fa3-11ed-9ea6-005056be4c1c
        rdf:type        crm:E33_Linguistic_Object ;
        rdfs:seeAlso    <http://webdev.its.uu.se/forn-test/api/getData.php?type=attestation&id=cc33be71-
7fa3-11ed-9ea6-005056be4c1c> ;
        crm:P190_has_symbolic_content  "til een stadh j grecia som kallas athenis" .

nw:6e9bf0c2-8040-11ed-9ea6-005056be4c1c
        rdf:type        crm:E31_Document ;
        rdfs:label      "Sj&auml;lens tr&ouml;st" .

nw:ac316281-803d-11ed-9ea6-005056be4c1c
        rdf:type        crm:E53_Place ;
        rdfs:label      "Athens" ;
        crm:P1_is_identified_by     [ rdf:type         crm:E41_Appellation ;
                                      rdfs:label       "Athens" ;
                                      crm:P2_has_type  nw:08636b0b-33c5-43af-a836-01f425cb7c70
                                    ] ;
```

Figure 7. Examples of NW Attestation and NW Standard form classified as rdf:types "E33_Linguistic Object" and "E53_Place" respectively.

The NPN dataset was modelled differently from the data already added to the NPN Sampo portal due to the aforementioned transition from Erlangen CRM to Linked Art Data Model between the two workshops. The place-name entity in the source data were mapped to crm:E33_E41_Linguistic_Appellations—a subclass of both E33_Linguistic_Object and E41_Appellation, see Figure 8. Since the data was enriched with approximate coordinates from the Norwegian Mapping Authority dataset, we used the property P189i_is_approximated_by, as suggested in the Linked Art documentation at the time. Unlike the other three datasets, the geometry of the coordinates was represented in the well-known text format, using CIDOC CRM alongside GeoSPARQL properties. This was not an issue when performing spatial queries across datasets however, as the GeoSPARQL and WGS84 vocabularies are interchangeable after being indexed by the triplestore.

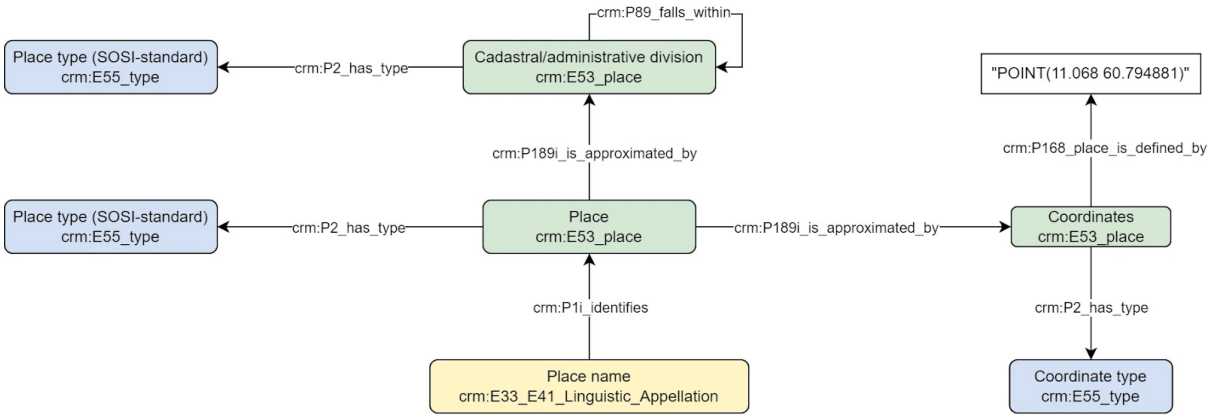

Figure 8. The core model of the NPN dataset after it was mapped to RDF and the Linked Art Data Model.

The data transformation step for NW and ISM involved exporting one RDF file per attestation. This export functionality was developed specifically for the Nordic Spatial Humanities project. The RDF files were then merged into a single file by the Apache Jena command line tool "riot". This single file was then uploaded to the Stardog knowledge graph instance. The NPN dataset was originally extracted from a relational database. It had been flattened and enriched with coordinates in a csv file, which in turn was mapped to RDF using the Python library rdflib.

For the querying, the team was not able to use the Sampo-UI ServerFS functionality because the implementation of the chosen framework, the Linked Art Data Model, was too inconsistent between the partaking SRIs. Instead, the Nordic Spatial Humanities case study opted for the ClientFS approach and thus could add a separate SPARQL query for each participating dataset. The GeoSPARQL queries showed that NW and ISM included information on and attestations of many common localities. However, the only common geographical entity included into the three SRIs, ISM, NW and NPN, turned out to be "Oslo", see Figure 9.

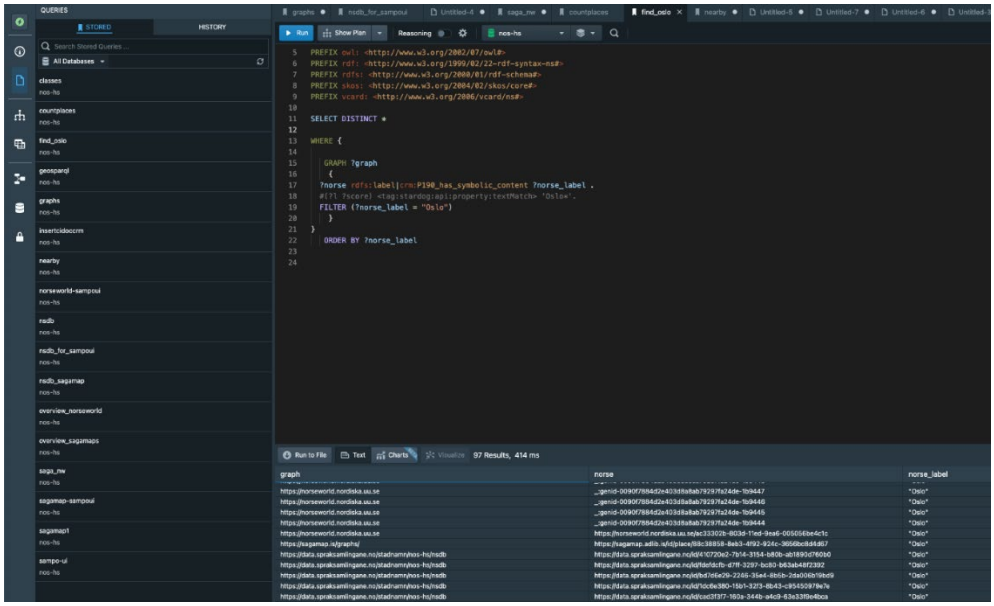

Figure 9. The results of the GeoSPARQL query to establish common localities or places in ISM, NW and NPN.

Thus, the end result was a search portal where the only common denominator between the datasets was the spatial aspect. One could argue that the same result could have been achieved without conversion to semantic data. Instead, we could combine data from multiple REST-APIs and relational databases, the technologies that participating developers are more familiar with. The chosen querying framework, Sampo-UI also suffers from scalability issues, making it more suited for limited research projects than large scale applications targeting a wider audience. Furthermore, developing a spatial humanities infrastructure based on linked data does not imply adopting a particular tech stack. Therefore, showcasing a semantic portal as well as the underlying triplestore technology may not have been the most suitable case for the Nordic Spatial Humanities case study.

What about MS?

The participating SRIs discussed and attempted to link the *Mapping Saints* data with data from the other projects. Initially, it was thought that this would be relatively straightforward, since there were some places in common in the MS dataset, including geodata for places in present-day Sweden, Finland, Norway, and other areas to which Swedish and Finnish people went on pilgrimages in the Middle Ages (e.g. similarities to NW's dataset). These common identifiers meant that, at least for some places, the linking was theoretically possible despite the fact that the MS dataset was still under development. However, the fact that the place data in the MS database was not complete at the time of the attempt made it more difficult to work with and created a challenge in satisfactorily linking it to the other three datasets.

In terms of the method, firstly, the MS place data was extracted in an xml-database (excel file). Exported from the MS relational database which applies a linked open data model, the selection of data included labels connected to medieval places and parishes located in the geographical area encompassed by present-day Sweden. In addition, the data included connections to a person (a saint) and an art historical object (featuring the saint). By extracting the data in this way, most of the connections from the relational database were thereby severed and simplified in the model created for the Bergen workshop. As these

connections are what makes the MS resource platform dynamic and contribute to the worth of the project, this temporary solution would not provide a realistic viable solution for a future linked data model involving MS. Moreover, other than as a demonstration or linking test, there was no attempt to map the place entities onto the same identifiers for MS as the focus for most of the Bergen workshop was on the other datasets. Instead, we first queried the results and then looked for matches.

The actual output produced focused on places associated with objects. Unlike the two other datasets, the model implemented was based on object production (see https://linked.art/model/object/production/). This test was done just after the Bergen workshop ended. It included the possibility of linking the common places and parishes found in all four datasets. It can be assumed that the complete spatial dataset would have further places in common. However, not everything could be connected (see also on NW's linguistic entities above). The data modelling in the three projects was different which affects the time and effort required. The focus of the project affects how well they overlap in terms of datasets and how well they can be linked. Given more time, it is highly probable that the MS dataset would also have been linkable to the other datasets, at least in terms of the place data that the projects have in common.

Discussion and concluding remarks

At the beginning of the article, we posed three research questions: 1) what are the common denominators of the participating SRIs in terms of data modelling and metadata? 2) what type of data transformation gives the best results in terms of linking and overall conformity to LOD principles? and 3) what are the main challenges on the way towards full SRI interoperability based on the outcomes of the Nordic Spatial Humanities case study?

Firstly, the main common denominator in terms of data structure across the studied SRIs is their spatial component, which includes basic metadata inspired by the digital gazetteer principle, i.e. coordinates, types of locality, and name labels. The four SRIs of the Nordic Spatial Humanities project have many more common features with respect to sources used, area of investigation, temporary perspectives, or the common goal to expose and make available the spatial side of the heterogenous material to a broad interdisciplinary audience. However, two principle differences, that of the scope and that of bespoke data modelling, do in fact outweigh the commonalities. ISM focuses mostly on places in Iceland (although some localities outside of Iceland are mentioned in the chosen source texts), NW on spatial references denoting localities outside of the modern borders of Sweden and Denmark, NPN provides Norwegian place-name data, while MS maps a variety of sources pertaining to cults of saints in medieval Sweden, Finland, Norway, Denmark, and other places in Europe.[5] In this Nordic Spatial Humanities SRI family, ISM and NW are the closest relatives—but this circumstance does not make the LOD implementation task easier or more straightforward because of the differences in conceptual data structuring. The NW prioritizes manuscript variation and excerpts attestations from multiple versions of the same texts, while ISM is primarily interested in spatial aspects of authoritative text editions. The "LOD readiness" factor plays an important role as well. MS and the NPN Sampo portal were designed with LOD implementation in mind and build on established solutions in this area, while ISM, NW,

[5] At the time of the Nordic Spatial Humanities project, however, the dataset primarily contained geocodes for relevant places in present-day Sweden.

and the NPN dataset had to carry out data augmentation of datasets from scratch during and between the workshops.

But does the aforementioned initial LOD implementation in MS and the NPN Sampo portal help in the actual linking task? Yes, on the theoretical side, but unfortunately it did not really help on the practical side. If we had used the NPN SPARQL endpoint instead of the NPN dataset prepared for the workshop, we would still have needed to do a manual mapping to combine it with MS. It is thus important to consider the theoretical versus the practical implications of linked open data, especially when working with partially compatible, incomplete datasets. In hindsight it could have been interesting to combine data from two SPARQL endpoints using a CONSTRUCT-query instead of trying to map the datasets to the same ontology in advance. This would however not eliminate the need to consider the aforementioned differences among the SRIs with respect to data modelling and provided metadata. The differences constitute a considerable challenge—especially combined with such factors as primarily time constraints of the two hands-on workshops as well as, to a much lesser degree, scalability the chosen ontologies and methods (apparent, if the tested set-up is expanded to include many more datasets, see on the Norwegian case below). Secondly, the Nordic Spatial Humanities case study has tried out two major data transformation methods, by mapping the datasets against two different ontologies Erlangen CRM and the 7.1.1 version of the CIDOC-CRM in order to reach the project's ambition to implement the Linked Art Data Model. Of these two, the latter one is deemed the most appropriate and consistent with the Nordic Spatial Humanities framework. However, there exists another, even more advantageous approach, the dedicated Linked Art subset and CIDOC-CRM extensions, of which the project team was not initially aware.

Thirdly, the main challenges on the way towards full SRI interoperability can be divided into conceptual and practical. Conceptually, the project's ambition to practically implement LOD can be questioned—taking into consideration the considerable scope and data modelling differences between the participating SRIs. It would have been more relevant and feasible to limit the case study to data conversion and data exchange between the SRIs. If this was the priority, the data harmonization in collaboration with domain experts and data enrichment with authority data would have been the central components—considerably affecting the outcomes and their (re-)usability. Instead, the Nordic Spatial Humanities case study prioritized the technical aspects of LOD implementation on partially incompatible and incomplete datasets. Practically, the aforementioned issues of time constraints, insufficient familiarity with the CIDOC-CRM ontologies and scalability issues have affected the results. In addition, it would have been advisable to have established the rules for result validation beforehand to ensure that the resulting outcome in fact matched the expected Linked Art Model profile.

The question of scalability has, for instance, resulted in the Norwegian Language Collections at the Bergen University Library abandoning Sampo-UI tech stack and the ambition of keeping all its data in a knowledge graph. Scalability both in terms of computational resources, and the overhead of having to map every new dataset to RDF, became an obstacle to unifying and replacing legacy systems such as the NPN dataset used here, and to publishing parts of the collection not yet available online. The Norwegian Language Collections has instead opted to store the data as JSON and build the new search interface *Stadnamnportalen* on top of an Elasticsearch cluster. This change does not imply that the Norwegian Language Collections has abandoned LOD altogether. All entries in *Stadnamnportalen*, both in each dataset and in an overarching dataset containing the place-

name expression ID entities mentioned above, have persistent UUIDs. This will facilitate mapping parts of the data to RDF, which can be queried in a public SPARQL endpoint.

Finally, the project results and the article's discussion lead to re-visiting the idea of so-called "LOD silos" (cf. Suominen & Hyvönen 2017) describing situations of LOD being implemented on either incompatible or partially incompatible datasets and underlying data models. This article clearly demonstrates that the existence of LOD silos is indeed still an issue at least among the Nordic SRIs studied in this project. The reason for this is however not the lack of skill or ability to "just follow the best practice" but the lack of established best practices and/or satisfactory ontologies for certain niched domains such as place-name standardization or spatial references in written or spoken language. For this reason, the emerging field of (Spatial) Research Infrastructure studies has an important role to play in analysing the current situation, describing the best practices as well as shaping future interoperable outcomes.

References

ARIADNEplus = ARIADNE RI. The data infrastructure serving the archaeological community worldwide. https://www.ariadne-research-infrastructure.eu/ (accessed 2024-05-13).

Árnastofnun/Rannsóknir og gögn = Digital databases maintained by the Árni Magnússon Institute for Icelandic Studies https://www.arnastofnun.is/is/rannsoknir-og-gogn (accessed 2024-05-13).

Backman, Agnieszka & Marcus Smith (2023) Sustainability and best practices for Linked Data heritage resources: some case studies from Sweden. In: Petrulevich & Skovgaard Boeck, eds. Digital Spatial Infrastructures and Worldviews in Pre-Modern Societies. Leeds: Arc Humanities Press, 159–183.

Bodenhamer, David J., John Corrigan, and Trevor M. Harris, eds. (2010) *The Spatial Humanities: GIS and the Future of Humanities Scholarship*. Bloomington: Indiana University Press.

Bowker, G.C., Baker, K., Millerand, F., Ribes, D. (2009). Toward Information Infrastructure Studies: Ways of Knowing in a Networked Environment. In: Hunsinger, J., Klastrup, L., Allen, M. (eds) International Handbook of Internet Research. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-9789-8_5 (accessed 2024-05-13).

CIDOC CRM = The CIDOC Conceptual Reference Model. https://www.cidoc-crm.org/ (accessed 2024-05-13).

CIDOC CRM/Place = https://www.cidoc-crm.org/FunctionalUnits/location-information (accessed 2024-05-13).

DARIAH-EU = DARIAH ERIC, a research infrastructure that enhances and supports digitally enabled research and teaching across the Arts and Humanities. https://www.dariah.eu/about/dariah-in-nutshell/ (accessed 2024-05-13).

dataARC = dataARC digital toolset that enables researchers to discover data about the long-term human ecodynamics of the North Atlantic. https://www.data-arc.org/ (accessed 2024-05-13).

Doerr, Martin, Richard Light, Gerald Hiebel (2020) Implementing the CIDOC Conceptual Reference Model in RDF. Version 1.1. https://www.cidoc-crm.org/sites/default/files/issue%20443%20-%20Implementing%20CIDOC%20CRM%20in%20RDF%20v1.1.pdf (accessed 2024-05-20)

Dunn, Stuart. (2019) *A History of Place in the Digital Age*. Abingdon, Oxon: Routledge.

Ellis Nilsson, Sara, Terese Zachrisson, Anders Fröjmark, Lena Liepe, & Johan Åhlfeldt (2023) Mapping Saints: Creating a Digital Spatial Research Infrastructure to Study Medieval Lived Religion. In: Petrulevich & Skovgaard Boeck, eds. Digital Spatial Infrastructures and Worldviews in Pre-Modern Societies. Leeds: Arc Humanities Press, 33–57.

Erlangen CRM = Erlangen CIDOC CRM Geo Extension (ECRMGeo) https://erlangen-crm.org/ecrmgeo (accessed 2024-05-13).

Gammeltoft, Peder (2023) Place-Name Databases – a Spatio-temporal Mess. In: Petrulevich & Skovgaard Boeck, eds. Digital Spatial Infrastructures and Worldviews in Pre-Modern Societies. Leeds: Arc Humanities Press, 137–157.

Getty = Getty Vocabularies https://www.getty.edu/research/tools/vocabularies/ (accessed 2024-05-13).

Golub, Koraljka & Ying-Hsang Liu (eds.): Information and Knowledge Organisation in Digital Humanities. Global Perspectives. Digital Research in the Arts and Humanities. London: Taylor and Francis.

Gregory, Ian N. and Alistair Geddes, eds. (2014) *Toward Spatial Humanities: Historical GIS and Spatial History*. Bloomington, Indiana: Indiana University Press.

Hardesty, J. L. (2016). Transitioning from XML to RDF: Considerations for an effective move towards Linked Data and the Semantic Web. Information Technology and Libraries, 35(1), 51–64. https://doi.org/10.6017/ital.v35i1.9182

Harlow, C. (2015). Data munging tools in preparation for RDF: Catmandu and LODRefine. The Code4Lib Journal, 30. https://journal.code4lib.org/articles/11013

Huminfra Sweden = Huminfra, Swedish national infrastructure supporting digital and experimental research in the Humanities. https://www.huminfra.se/ (accessed 2024-05-13).

Hyvönen, Eero (2023) Digital humanities on the Semantic Web: Sampo model and portal series. Semantic Web 14 (2023), 729–744. DOI 10.3233/SW-223034

Ikkala, E., E. Hyvönen, H. Rantala and M. Koho, Sampo-UI: A full stack JavaScript framework for developing semantic portal user interfaces, semantic web, *Interoperability, Usability, Applicability* **13**(1) (2022), 69–84. doi:10.3233/SW-210428.

ISM/Icelandic Saga Map. http://sagamap.hi.is/is (accessed 2024-05-13).

IPERION HS = IPERION HS, the European Research Infrastructure on Heritage Science. https://www.iperionhs.eu/ (accessed 2024-05-13).

Lethbridge, Emily (2020) "Digital Mapping and the Narrative Stratigraphy of Iceland." In: Historical Geography, GIScience and Textual Analysis: Landscapes of Time and Place, edited by Charles Travis, Francis Ludlow, and Ferenc Gyuris. Cham: Springer, 19–32.

Lethbridge, Emily (2023) Medieval to Modern: Using Spatial Data from the Digital Projects Icelandic Saga Map and Nafnið.is to Explore the Interaction between Narrative and Place in Iceland. In: Petrulevich & Skovgaard Boeck, eds. Digital Spatial Infrastructures and Worldviews in Pre-Modern Societies. Leeds: Arc Humanities Press, 59–89.

Linked Art Datamodel/Places = Linked Art Datamodel Places https://linked.art/model/place/ (accessed 2024-05-13).

Miller, M., & Vielfaure, N. (2022). OpenRefine: An Approachable Open Tool to Clean Research Data. Bulletin - Association of Canadian Map Libraries and Archives (ACMLA), (170). https://doi.org/10.15353/acmla.n170.4873

MS/Mapping Saints. https://saints.dh.gu.se This still links to the input interface; however, the public interface will be available via the same link from mid-2024. (accessed 2024-05-13).

Murrieta-Flores, Patricia and Naomi Howell. "Towards the Spatial Analysis of Vague and Imaginary Place and Space: Evolving the Spatial Humanities through Medieval Romance." *Journal of Map and Geography Libraries*, 13, no. 1 (January 2, 2017): 29–57. DOI: 10.1080/15420353.2017.1307302.

NameSampo = NameSampo: A Workbench for Toponomastic Research https://nimisampo.fi/en/places/federated-search/table (accessed 2024-05-13).

Nordic Spatial Humanities = The Nordic Spatial Humanities Project. Linking, Building, and Sustaining Humanities Digital Spatial Infrastructures for Research in the Nordic Countries. https://nordic-spatialhumanities.org/ (accessed 2024-05-13).

NW/Attestation = Data and related metadata: attestations https://www3.uu.se/en/research/infrastructure/norseworld/infrastructure/data-and-metadata#anchor-743478 (accessed 2024-05-13).

NW/Norse World = Petrulevich, Alexandra, Agnieszka Backman, Jonathan Adams, Simon Skovgaard Boeck, Jessica Holmlund, Felix Marklund, Jorunn Hartmann, Andreas Lecerof, Rasmus Ljungström, and Alexia-Nadina Hrenyak. Norse World. https://norseworld.nordiska.uu.se/ and https://www3.uu.se/en/research/infrastructure/norseworld/ (both accessed 2024-05-13).

NW/Standard form = Data and related metadata: spatial references and spatial data https://www3.uu.se/en/research/infrastructure/norseworld/infrastructure/data-and-metadata#anchor-753378 (accessed 2024-05-13).

NW/Source = Source and related metadata. https://www3.uu.se/en/research/infrastructure/norseworld/infrastructure/data-and-metadata#anchor-753372 (accessed 2024-05-13).

NW/Work = Work and related metadata. https://www3.uu.se/en/research/infrastructure/norseworld/infrastructure/data-and-metadata#anchor-753376 (accessed 2024-05-13).

The National Place-name Database = a legacy search interface and database maintained by The Norwegian Language Collections, now integrated into Stadnamnportalen https://stadnamnportalen.uib.no/view/nbas (accessed 2024-06-27).

NPN/Norwegian Place-names = NPN Sampo portal, Norske stedsnavn/Norske stadnamn, a spatial research infrastructure with a search interface based on NameSampo, now replaced by Stadnamnportalen. https://stadnamnportalen.uib.no/view/nbas (accessed 2024-06-27), previously https://toponymi.spraksamlingane.no/nb/app

OpenRefine = OpenRefine, a powerful free, open source tool for working with messy data. https://openrefine.org/ (accessed 2024-05-13).

OPERAS-EU = OPERAS, the European Research Infrastructure for the development of open scholarly communication in the social sciences and humanities. https://operas-eu.org/about/ (accessed 2024-05-13).

Pálsson, G. & Opitz, R. (2019) DataArc: a case study using CIDOC CRM as an ontology for transdisciplinary research. In Niccolucci, F. and Richards, J. (eds.) The ARIADNE Impact, Archaeolingua Foundation, Hungary.

Petrulevich, Alexandra (2023) Place-name variation in medieval literature in the digital age. In: Petrulevich & Skovgaard Boeck, eds. Digital Spatial Infrastructures and Worldviews in Pre-Modern Societies. Leeds: Arc Humanities Press, 243–265.

Petrulevich, Alexandra, Agnieszka Backman & Jonathan Adams (2020) “Medieval Macrospace through GIS: The Norse World Project Approach.” The Cartographic Journal 57: 18–27.

Petrulevich, Alexandra & Simon Skovgaard Boeck, eds. (2023a) Digital Spatial Infrastructures and Worldviews in Pre-Modern Societies. Leeds: Arc Humanities Press.

Petrulevich, Alexandra & Simon Skovgaard Boeck (2023b) Introduction. In: Petrulevich & Skovgaard Boeck, eds. Digital Spatial Infrastructures and Worldviews in Pre-Modern Societies. Leeds: Arc Humanities Press, 1–12.

Petrulevich, Alexandra & Simon Skovgaard Boeck (2023c) Norse World from Plan to Action: Building a Digital Gazetteer of East Norse Medieval Literature Step by Step. In: Petrulevich & Skovgaard Boeck, eds. Digital Spatial Infrastructures and Worldviews in Pre-Modern Societies. Leeds: Arc Humanities Press, 15–32.

RDF = Resource Description Framework. https://www.w3.org/RDF/ (accessed 2024-05-13).

RDF Transform = AtesComp/rdf-transform on GitHub. https://github.com/AtesComp/rdf-transform (accessed 2024-05-13).

SCOS = Simple Knowledge Organization System. https://www.w3.org/2009/08/skos-reference/skos.html (accessed 2024-05-13).

SOCH = Swedish Open Cultural Heritage https://www.raa.se/in-english/digital-services/about-soch/ (accessed 2024-05-13).

SPARQL = SPARQL, a standard query language for Linked Open Data https://www.w3.org/TR/sparql11-query/ (accessed 2024-05-13).

Stadnamnportalen = The current Norwegian Place Names (NPN) search interface https://stadnamnportalen.uib.no/ (accessed 2024-06-27).

Stardog = The Stardog Knowledge Graph platform. https://www.stardog.com/ (accessed 2024-05-13).

Stardog/RDF = The Stardog RDF Graph Data Model https://docs.stardog.com/tutorials/rdf-graph-data-model (accessed 2024-05-13).

Suominen, Osma & Nina Hyvönen (2017). From MARC silos to Linked Data silos? From MARC silos to Linked Data silos? *O-Bib. Das Offene Bibliotheksjournal Herausgeber VDB*, *4*(2), 1–13. https://doi.org/10.5282/o-bib/2017H2S1-13

Swedish Digital Place-Name Register (Ortnamnsregistret). https://ortnamnsregistret.isof.se (accessed 2024-05-13).

Tally Jr., Robert T., ed. (2014) *Literary Cartographies: Spatiality, Representation, and Narrative*. Geocriticism and Spatial Literary Studies. New York: Palgrave Macmillan.

W3C/RDF = Resource Description Framework (RDF). https://www.w3.org/RDF/ (accessed 2024-05-13).

W3C/RDFS = RDF Schema (RDFS). https://www.w3.org/wiki/RDFS (accessed 2024-05-13).