

# National Linked Open Data Infrastructure for Digital Humanities (LODI4DH)

Eero Hyvönen



[National Library of Finland, photo by Minna Tamper]

Data, the oil of the digital world, is typically interlinked in content, published in different formats and languages, and is distributed in different services across countries. To create intelligent web services based on such heterogeneously distributed data sources, e.g., collection data in different cultural heritage organizations, the data has to be made mutually interoperable and machine understandable. Furthermore, lots of duplicate work can be eliminated if the data is not only interoperable, but also findable, accessible, and re-usable between the data publishers and users. To satisfy these FAIR principles<sup>1</sup>, a shared data infrastructure is needed.

## A National Vision since 2003

Linked Open Data Infrastructure for Digital Humanities in Finland (LODI4DH)<sup>2</sup> is a joint initiative of the Helsinki Centre for Digital Humanities (HELDIG), the University of Helsinki (UH), and Aalto University Department of Computer Science. It aims to create living lab prototypes of centralized national Linked Data

services for open science. The services enable publication and utilization of datasets for data-intensive Digital Humanities (DH) research and application in structured, standardized Semantic Web formats via open interfaces<sup>3</sup>. Major components of LODI4DH include domain ontologies for data linking (historical places and maps, persons, events, keyword concepts, and times), harmonizing metadata models (extending, e.g., the CIDOC CRM ontology), publishing core datasets for re-use (related to cultural heritage and legislation), various Linked Data services for developers, and online learning materials about Linked Data.

LODI4DH is based on a large collaboration network and software infrastructure created during a long line of national projects in DH between UH and Aalto University since 2003. These project collaborations between some 50 organizations have resulted in several in-use infrastructure prototypes and services, such as the ONKI ontology service, deployed as the Finto ontology service<sup>4</sup> by the National Library of Finland, and the “7-star” Linked Data Finland model and service LDF.fi<sup>5</sup>. This line of research started with the national FinnONTO project series (2003-2012)<sup>6</sup> on creating a national ontology infrastructure in Finland, and continued on with, e.g., the projects Linked Data Finland (2012-2014)<sup>7</sup> and Linked Open Data Science Service by the SeCo group at Aalto University and the University of Helsinki.

### **Ontology and Data Services in Use**

ONKI/Finto and Linked Data Finland services already have had a wide user base demonstrating the need for the LODI4DH infrastructure. Applications based on them have also made their way from academic research into real use. For example, the system “BookSampo- Finnish Fiction Literature on the Semantic Web”<sup>8</sup>, based originally on LDF.fi and maintained now by the Finnish public libraries (Kirjastot.fi), had 2 million visitors in 2018, the semantic portal “WarSampo –Finnish World War II on the Semantic Web”<sup>9</sup> had 230 000 distinct users in 2018, and there are tens of thousands of monthly users in the Finto service. Many museums in Finland, e.g., Espoo City Museum, AKSELI Consortium of eight museums, and the new national KOOKOS cataloging system, make use of the ONKI ontologies and Finto.

In addition to the Finnish projects, there are several joint research projects with foreign universities, such as University of Oxford, Stanford University, University of Colorado Boulder, and University of Pennsylvania, where the Finnish Linked Data services for DH have been used, such as Reassembling the Republic of Letters<sup>10</sup> dealing with Early Modern correspondence data and Mapping Manuscript Migrations<sup>11</sup> on medieval pre-modern manuscript data.

LODI4DH aims at harnessing all this work into sustainable services, and integrating the work as a component into the EU ERIC DARIAH infrastructure in the future. LODI4DH infrastructure is open source, publishes open data, and is free of charge for everyone to use. LODI4DH focuses on DH research infrastructures but the underlying Linked Data and Semantic Web technology can and has been utilized in other fields of research, too, extending substantially the utilization potential of the infrastructure.

**Eero Hyvönen<sup>xii</sup>** is director of the Helsinki Centre for Digital Humanities (HELDIG) at the University of Helsinki and professor of semantic media technology at the Aalto University. He has been directing the research on Semantic Web infrastructures and applications of the Semantic Computing Research Group (SeCo) at the University of Helsinki, HELDIG, and Aalto University since 2002. Eero Hyvönen has published over 400 research articles and books and received several national and international awards. He acts in the editorial boards of *Semantic Web - Interoperability, Usability, Applicability, Semantic Computing*, *International Journal of Metadata, Semantics, and Ontologies*, and *International Journal on Semantic Web and Information Systems*, and has co-chaired and acted in the program committees of tens of major international conferences.

---

<sup>1</sup> <https://www.go-fair.org/fair-principles/>

<sup>2</sup> <https://seco.cs.aalto.fi/projects/lodi4dh/>

<sup>3</sup> This work is coordinated internationally by the World Wide Web consortium W3C:

<https://www.w3.org/standards/semanticweb/>

<sup>4</sup> <http://finto.fi/en/>

<sup>5</sup> <http://ldf.fi>

<sup>6</sup> <https://seco.cs.aalto.fi/projects/finnonto/>

<sup>7</sup> <https://seco.cs.aalto.fi/projects/ldf/>

<sup>8</sup> Available online at <http://kirjasampo.fi>

<sup>9</sup> Available online at <http://sotasampo.fi/en>

<sup>10</sup> <http://www.republicofletters.net/>

<sup>11</sup> <http://mappingmanuscriptmigrations.org/>