The debates of the European Parliament as Linked Open Data

Astrid van Aggelen a.e.van.aggelen@gmail.com
Dept. of Computer Science, VU University Amsterdam, NL

Laura Hollink l.hollink@cwi.nl
Center for Mathematics and Computer Science (CWI), Amsterdam, The Netherlands

Max Kemman max.kemman@uni.lu
Department of History, University of Luxembourg, Luxembourg

Martijn Kleppe kleppe@eshcc.eur.nl
Department of History, Erasmus University Rotterdam, The Netherlands

Henri Beunders beunders@eshcc.eur.nl
Department of History, Erasmus University Rotterdam, The Netherlands

Keywords

Linked Open Data   European Parliament   open government data   RDF   data modeling   multilingual data

Abstract

The European Parliament represents the citizens of the member states of the European Union (EU). The accounts of its meetings and related documents are open data, promoting transparency and accountability, and are used as source data by researchers. However, the official portal of these documents provides limited search facilities. This paper presents LinkedEP, a Linked Open Data translation of the verbatim reports of the plenary meetings of the European Parliament. These data are integrated with a database of political affiliations of the Members of Parliament, and enriched with detected topics from the EU's topic hierarchy and links to three other Linked Open Datasets. The results of this work are available through a SPARQL endpoint as well as a user interface with extensive browse and search facilities. It is now possible to combine in one query the time and topic of the debate, the spoken words - in any available translation - and information about the speaker uttering these, such as affiliations to countries, parties and committees. This paper discusses the design and creation of the vocabulary, data and links, as well as known use of the data.