

e- Slovenian Biographical Lexicon

Petra Vide Ogrin

, Slovenian Academy of Sciences and Arts, Library

Jan Jona Javoršek

, Jožef Stefan Institute

Tomaž Erjavec

, Jožef Stefan Institute

Abstract

Slovenian Biographical Lexicon (SBL) was published in 16 volumes (index included) from 1925 to 1991. It comprises 5,044 biographical entries and, since some entries are family names, SBL covers more than 5,100 persons. SBL used bio-bibliographical methods to provide synthetic assessments of work and significance of historical figures on the basis of primary sources. SBL has been out of print for a long time, but the publication has been seen as an important resource for encyclopaedic and reference editions and research in the Slovenian humanities, social sciences and history of the natural sciences. Therefore, Slovenian Academy of Sciences and Arts (SASA) and the Scientific Research Centre of the SASA decided to produce a freely available on-line digital edition of SBL. In the process of digitization, manually corrected OCR has been semi-automatically converted to XML-based Text Encoding Initiative format (TEI P5). Its extensive annotation vocabulary, notably from the biographical and prosopographical modules, has been used – automatically and manually - to mark-up as much data as possible. The resulting XML document has become the data resource of an online digital repository based on Fedora Commons platform, where we implemented an infrastructure of XML processing methods on top of native relationships and a Lucene/SOLR based search engine to produce a full-fledged web application and search engine with browser, metadata and web application interfaces. In the near future, we intend to build a Slovenian Biographical Hub, which will gather a comprehensive biographical information in one place and connect other local and mostly regional biographical lexicons with the existing e-SBL. Slovenian Biographical Hub will enable simple and advanced searching across all lexicons included in the Hub and integrated in the Fedora Commons system.