



Linking Knowledge

Linked Open Data for
Knowledge Organization and Visualization

Edited by
Richard P. Smiraglia – Andrea Scharnhorst

Ergon

Linking Knowledge

Edited by

Richard P. Smiraglia
Andrea Scharnhorst

Linking Knowledge

Linked Open Data for Knowledge Organization and Visualization

Edited by

Richard P. Smiraglia
Andrea Scharnhorst

ERGON VERLAG

Cover illustration:
© iStock - your_photo

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>.

© Ergon – ein Verlag in der Nomos Verlagsgesellschaft, Baden-Baden 2021

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in other ways and storage in databanks.

Duplication of this publication or parts thereof is only permitted under the provisions of the German Copyright Law, a copyright fee must always be paid.

Overall responsibility for manufacturing (printing and production) lies with
Nomos Verlagsgesellschaft mbH & Co. KG.

Cover Design: Jan von Hugo

www.ergon-verlag.de

ISBN 978-3-95650-660-4 (Print)

ISBN 978-3-95650-661-1 (ePDF)

Table of Contents

List of Figures	vii
Chapter 1. The Need for Knowledge Organization Introduction to the book <i>Linking Knowledge: Linked Open Data for Knowledge Organization</i> Andrea Scharnhorst and Richard P. Smiraglia	1
Chapter 2. Classifications as Linked Open Data Challenges and Opportunities Rick Szostak, Richard P. Smiraglia, Andrea Scharnhorst, Aida Slavic, Daniel Martínez-Ávila, Tobias Renwick	24
Chapter 3. Knowledge Organization Systems (KOS) in the Semantic Web A Multi-Dimensional Review Marcia Lei Zeng and Philipp Mayr	34
Chapter 4. A Thesaural Interface for the Basic Concepts Classification Tobias Renwick and Rick Szostak	64
Chapter 5. Publishing a Knowledge Organization System as Linked Data The Case of the Universal Decimal Classification Aida Slavic, Ronald Siebes and Andrea Scharnhorst	69
Chapter 6. Modeling and Visualizing Storylines of Historical Interactions Kubler’s Shape of Time and Rembrandt’s Night Watch Charles van den Heuvel and Veruska Zamborlini	99
Chapter 7. Identifying and Classifying the Phenomena of Music Richard P. Smiraglia and Rick Szostak	142
Chapter 8. Graphing Out Communities and Cultures in the Archives Methods and Tools M. Cristina Pattuelli	149
Chapter 9. Digging into the Mensural Music Knowledge Graph Renaissance Polyphony meets Linked Open Data Richard P. Smiraglia, James Bradford Young and Marnix van Berchum	167
Chapter 10. Organizing Scholarly Knowledge leveraging Crowdsourcing, Expert Curation and Automated Techniques Mohamad Yaser Haradeh, Markus Stocker and Sören Auer	181

Chapter 11. Knowledge Spaces Visualizing and Interacting with Dimensionality Charles van den Heuvel and Richard P. Smiraglia	199
Chapter 12. Publishing Linked Open Data Ronald Siebes, Gerard Coen, Kathleen Gregory and Andrea Scharnhorst	218
Contributors	233
Index	235

List of Figures

Chapter 1

Figure 1. Author co-citation among those most cited	17
Figure 2. Author co-citation among contributing authors	18
Figure 3. Most frequently occurring keywords (stress = .017138 R ² = .9602)	19
Figure 4. Most frequently occurring phrases (stress = .18684 R ² = .9670)	19

Chapter 3

Figure 1. The options and actions related with KOS in the LOD dataset production	40
Figure 2. The 5-star LOD Cloud indicates the essential role of LOD KOS vocabularies. (Source: Annotated by the author on the LOD CLOUD 2014-08-30 image http://lod-cloud.net/versions/2014-08-30/lod-cloud_colored.png)	43
Figure 3. Using a template provided to trace data of “Descendants of a Given Parent” for “<costume by function>“ (AAT concept ID 300212133). (Source: http://vocab.getty.edu/queries#Descendants_of_a_Given_Parent)	45
Figure 4. Querying for “<costume by function>“ (AAT concept ID 300212133), receiving and downloading the datasets to make a microthesaurus	46
Figure 5. Data flow diagram of the interlinking procedure in the Swissbib project (Source: Bensmann et al. (2017, 8 Figure 4)	49
Figure 6. Query examples provided by UniProt (upper figure) and the SPARQL query for question #8, automatically “show”ed (lower figure) (Source: http://sparql.uniprot.org/)	52
Figure 7. Templates of TGN-specific queries, provided by Getty Vocabularies LOD service (Source: http://vocab.getty.edu/queries#TGN-Specific_Queries)	53
Figure 8. Using the template provided by the LOD service, a query is submitted (upper figure), resulting a dataset (lower figure) for a specific place type (e.g., caves) in a geographic boundary (Source: http://vocab.getty.edu/queries#TGN-Specific_Queries)	54
Figure 9. Templates of ULAN-specific queries, provided by Getty Vocabularies LOD service (Source: http://vocab.getty.edu/queries#ULAN-Specific_Queries)	55

Figure 10. Using the template (upper figure) provided by the LOD service, a query is submitted to get the dataset for an artist *Wright, Frank Lloyd* and his associative relationships (lower figure)
(Source: http://vocab.getty.edu/queries#ULAN-Specific_Queries) 56

Figure 11. The graphic overview of the group “Activity” of the *Cadastral and Land Administration Thesaurus (CaLAtHe)* (Source: <http://cadastralvocabulary.org/>) 57

Chapter 5

Figure 1. An excerpt from the UDC scheme hierarchy showing captions in English and French 74

Figure 2. Caption of the UDC class 538.9 in six languages and scripts (UDC Summary) 74

Figure 3. Placement and linking of a concept of “Netherlands” in different parts of UDC 76

Figure 4. UDC Summary Linked Data (2011-2019) showing UDC class 311, HTML display and its RDF record 83

Figure 5. Examples of UDC notations from library linked data (catalogue NTNU) 84

Figure 6. UDC Look-up service architecture 86

Figure 7. UDC Look-up service and interpreter 87

Figure 8. Example of a UDC look-up service URI 90

Figure 9. RDF graph representation of class 94 and its caption General History 91

Figure 10. Top level of UDC structure in UDC Summary with corresponding RDF graph representation of class 94 91

Figure 11. RDF graph representation of a complex UDC notation 94(729.885):94(492) . 92

Figure 12. Data elements in UDC LD schema 93

Chapter 6

Figure 1. 3D semantic timeline-visualises development story in time-intervals (longitudinal) and network of relations between storylines (transversal) similar to Kubler’s description of fibers of duration and networks in cross-section (Jensen 2006) 108

Figure 2. *Night Watch* and Derivatives: a) *Night Watch*; b) Etching Claessens 1797 after original; c) Tattoo of *Night Watch* on back Marko Bak during visit to the Rijksmuseum on 18th of May, 2019; and, d) storytelling about the composition of *The Night Watch* by the Rijksmuseum) 111

Figure 3. Storylines of the production and consumption of *The Night Watch* in copies, adaptations and digital reproductions hereof 113

Figure 4. The left-hand side depicts a longitudinal zoom in on *The Night Watch*, while the right-hand side depicts a longitudinal zoom out showing *The Night Watch* among other paintings by Rembrandt 113

Figure 5. Zooming in on immaterial and material aspects of the production of *The Night Watch* and copies or adaptations thereof..... 114

Figure 6. Storylines of Rembrandt’s paintings based on information available in 2019 116

Figure 7. Storylines (longitudinal) on the left-hand side and cross-sections on the right- hand side. The one on top is a snapshot (synchronous cross-section) of Rembrandt’s existing paintings in 2019 whilst the one at the bottom is a kaleidoscope view (asynchronous cross-section) of Rembrandt’s paintings according to information available in 2019, similar to the digital reconstruction hereof for the Virtual exhibition “Discover Rembrandt: His Life and all his Paintings.” 117

Figure 8. Zooming in on the immaterial part of *The Night Watch*, the Militia Group Portrait theme manifests as its content aspect, while the chiaroscuro Feature manifests as its (re)presentation aspect 118

Figure 9. The paintings 1-7 are presented as examples of manifestations of solutions, styles and genres 119

Figure 10. *Endurants* and *Perdurants* can have respectively spatial and temporal extents which are independent of a specific quality structure and can be projected in one or more of them, e.g., someone’s birth date can be projected in both Gregorian and Chinese calendars 124

Figure 11. *Period* and *Duration* are abstract entities which are worth naming. They can be named after a specific event, e.g., the 2nd World War, or may refer to a particular time interval within a calendar, such as the 1960s or the year of the rooster 126

Figure 12. Storylines comprise the participations of an object/entity or of a bundle of them in events through time. A storyline transversal view is a static view or a network, which can be a synchronous view in time (e.g., Figures 2 and 7 top right) or it can be an asynchronous view (e.g., Figure 7 bottom right) as to connect objects that participate in related events at different points in time 128

Figure 13. Modeling a particular type of storyline, namely of products, their production and consumption, material or immaterial 130

Figure 14. Modeling Kubler’s views of chain of solutions as well as style as longitudinal views over the creation of products that manifest a solution or style 133

Figure 15. Modeling of Kubler’s concepts and complementary interpretations related to storylines and their transversal views 135

Chapter 8

Figure 1. The Linked Jazz ontology (image by Sarah A. Adams) 153

Figure 2. Visualization of the ego-network of Billie Holiday 154

Figure 3. Properties in context on Sam River (Q26) Wikibase page 157

Figure 4. Table view of SPARQL query results 158

Figure 5. Gender view of the Linked Jazz graph (image by Karen Hwang) 160

Figure 6. Map visualization of New Orleans historical districts (top) and HOLC’s predatory loan districts (bottom)(images by Genvieve Milliken) 162

Chapter 9

Figure 1. CMME incipits for Weerbeke *Ave regina celorum a4* 170

Figure 2. Data model of Di4KG CMME project 172

Figure 3. Text portion of authority record for composer name 173

Figure 4. Authority record including term from *LCSH* 175

Figure 5. RDF rendering of authority record for Jacquet of Mantua’s *Magnificat* 176

Figure 6. Model of the Mensural Music Knowledge Graph 177

Figure 7. Opening theme of Beethoven’s *Fifth Symphony* 178

Figure 8. Bibliographic representation of a score of Beethoven’s *Fifth Symphony* 178

Chapter 10

Figure 1. New knowledge, old methods: For centuries, the same method has been used to pass on research knowledge—scientific articles	182
Figure 2. Connecting semantic descriptions of research contributions with various research artifacts using the Open Research Knowledge Graph	183
Figure 3. ORKG layered architecture from data persistence to services	184
Figure 4. RDF inspired base data model used within the ORKG infrastructure	185
Figure 5. Interplay between crowdsourcing and automated approaches in ORKG	187
Figure 6. Abstract annotator for a paper abstract during adding paper information	189
Figure 7. Evaluation dataset imported in the ORKG and visualized as a table	191
Figure 8. QA system prototype to answer question over structured data within the ORKG	192
Figure 9. Comparison of question answering systems based on the tasks performed by these systems (data imported from Diefenbach et al. 2018)	193
Figure 10. Comparison of evaluation results of the question answering systems presented in Figure 9 (depicted values are imported from Diefenbach et al. 2018)	194
Figure 11. Automatic comparison of basic reproduction numbers published in the literature	195
Figure 12. Visualization of R0 values and 95% confidence interval over time, as a possible output of a data science activity that reuses ORKG comparison data	195
Figure 13. ORKG comparison for the material science domain (data imported from Juzeliūnas and Gray 2019)	196

Chapter 11

Figure 1. a. Conversion data and linking classifications to the semantic web in Digging into the Knowledge Graph project (top); b. Pre-classification of unstructured data with AI in Golden Agents project (bottom)	201
Figure 2. Otlet sketches of 3D knowledge spaces: projecting thought on 3D space, dimension expansion and capturing dynamics of 3D knowledge spaces	202
Figure 3. Dahlberg’s ICC matrix (version 2012)	204

Figure 4. Dahlberg matrix and 3D reconstruction of ICC (virtual reconstruction by Chiara Piccoli)	209
Figure 5. Representation of first two layers of ICC in SW model Ontology 4Us	210
Figure 6. Concept and state vectors; subspace and search volumes in multidimensional knowledge space (Meincke and Atherton 1976)	212
Chapter 12	
Figure 1. An example of an RDF triple	220
Figure 2. A screenshot of an RDF representation	220
Figure 3. The original data in tabular form	222
Figure 4. Examples of concepts and instances in our example dataset	222
Figure 5. An example of our data model showing the relationships between concepts and the subject : predicate : object structure (marked in color)	223
Figure 6. An example of modelling data	223
Figure 7. Example of the three parts of a URI and a namespace	224
Figure 8. An example of a URI containing version information	225
Figure 9. Example URIs that use specific terms to represent the relationship between an object in the real world (here, a fossil) and the types of descriptions	225
Figure 10. An example of filtering data	227
Figure 11. An example of a modelled dataset (left), with some potential external concepts for the data to be linked to (right)	229
Figure 12. An example of a modelled dataset (left), linked to some external concepts (right)	230
Figure 13. A map showing a combination of Linked Data from the Dutch census with geographic data; the heatmap shows the total number of female inhabitants	231