

# Contents

## Part I Approaches to Collaboratively Constructed Language Resources

<b>1 Using Games to Create Language Resources: Successes and Limitations of the Approach</b> .....	3
Jon Chamberlain, Karèn Fort, Udo Kruschwitz, Mathieu Lafourcade, and Massimo Poesio	
1.1 Introduction .....	3
1.2 Collaborative Creation and Collective Intelligence .....	4
1.3 Approaches to Creating Language Resources .....	6
1.4 Using Games to Create Language Resources .....	9
1.5 Defining Collaborative Approaches .....	17
1.6 Evaluating the Gaming Approach to Creating Language Resources .....	26
1.7 Conclusions .....	36
References .....	40
<b>2 Senso Comune: A Collaborative Knowledge Resource for Italian</b> ....	45
Alessandro Oltramari, Guido Vetere, Isabella Chiari, Elisabetta Jezek, Fabio Massimo Zanzotto, Malvina Nissim, and Aldo Gangemi	
2.1 Introduction .....	46
2.2 The Model .....	51
2.3 The Acquisition Process .....	54
2.4 The TMEO Methodology .....	57
2.5 Experiments on Noun Word Sense Ontology Tagging .....	60
2.6 Relevance to Natural Language Processing .....	62
2.7 Conclusions and Future Work .....	64
References .....	65

**3 Building Multilingual Language Resources in Web Localisation: A Crowdsourcing Approach** ..... 69  
 Asanka Wasala, Reinhard Schäler, Jim Buckley, Ruwan Weerasinghe, and Chris Exton

3.1 Introduction ..... 70  
 3.2 System Architecture ..... 79  
 3.3 An Illustrative Scenario ..... 85  
 3.4 Prototype Technologies ..... 88  
 3.5 Discussion: Outstanding Challenges ..... 90  
 3.6 Conclusions and Future Work ..... 96  
 References ..... 97

**4 Reciprocal Enrichment Between Basque Wikipedia and Machine Translation** ..... 101  
 Iñaki Alegria, Unai Cabezón, Unai Fernández de Betoño, Gorka Labaka, Aingeru Mayor, Kepa Sarasola, and Arkaitz Zubiaga

4.1 Introduction ..... 102  
 4.2 Background ..... 103  
 4.3 Methodology ..... 106  
 4.4 Results and Discussion ..... 112  
 4.5 Conclusions and Future Work ..... 116  
 References ..... 118

**Part II Mining Knowledge from and Using Collaboratively Constructed Language Resources**

**5 A Survey of NLP Methods and Resources for Analyzing the Collaborative Writing Process in Wikipedia** ..... 121  
 Oliver Ferschte, Johannes Daxenberger, and Iryna Gurevych

5.1 Introduction ..... 121  
 5.2 Revisions in Wikipedia ..... 123  
 5.3 Discussions in Wikipedia ..... 142  
 5.4 Tools and Resources ..... 152  
 5.5 Conclusion ..... 155  
 References ..... 157

**6 ConceptNet 5: A Large Semantic Network for Relational Knowledge** ..... 161  
 Robert Speer and Catherine Havasi

6.1 Introduction ..... 161  
 6.2 Knowledge in ConceptNet 5 ..... 165  
 6.3 Storing and Accessing ConceptNet Data ..... 171  
 6.4 Evaluation ..... 172  
 References ..... 175

**7 An Overview of BabelNet and its API for Multilingual Language Processing** ..... 177  
 Roberto Navigli and Simone Paolo Ponzetto

7.1 Introduction ..... 177

7.2 Knowledge Resources ..... 179

7.3 BabelNet ..... 182

7.4 Statistics on BabelNet ..... 186

7.5 Multilingual NLP in the Fast Lane with the BabelNet API ..... 188

7.6 Related Work ..... 190

7.7 Conclusions ..... 193

References ..... 194

**8 Hierarchical Organization of Collaboratively Constructed Content** ..... 199  
 Jianxing Yu, Zheng-Jun Zha, and Tat-Seng Chua

8.1 Introduction ..... 199

8.2 Related Works ..... 205

8.3 Hierarchical Organization Framework ..... 209

8.4 Evaluations ..... 217

8.5 Application ..... 225

8.6 Conclusions and Future Works ..... 236

References ..... 236

**9 Word Sense Disambiguation Using Wikipedia** ..... 241  
 Bharath Dandala, Rada Mihalcea, and Razvan Bunescu

9.1 Introduction ..... 241

9.2 Wikipedia ..... 243

9.3 Wikipedia as a Sense Tagged Corpus ..... 246

9.4 Word Sense Disambiguation ..... 249

9.5 Experiments and Results ..... 250

9.6 Related Work ..... 255

9.7 Conclusions ..... 258

References ..... 259

**Part III Interconnecting and Managing Collaboratively Constructed Language Resources**

**10 An Open Linguistic Infrastructure for Annotated Corpora** ..... 265  
 Nancy Ide

10.1 Introduction ..... 265

10.2 Requirements for a Collaborative Annotation Effort ..... 267

10.3 ANC-OLI ..... 271

10.4 ANC-OLI in Context ..... 280

10.5 Looking Forward ..... 282

10.6 Conclusion ..... 284

References ..... 284

<b>11</b>	<b>Towards Web-Scale Collaborative Knowledge Extraction</b> .....	287
	Sebastian Hellmann and Sören Auer	
11.1	Introduction .....	287
11.2	Background .....	290
11.3	Collaborative Knowledge Extraction .....	293
11.4	The NLP Interchange Format .....	304
11.5	Interoperability Between Different Layers of Annotations .....	307
11.6	Discussion and Outlook .....	310
	References .....	311
<b>12</b>	<b>Building a Linked Open Data Cloud of Linguistic Resources: Motivations and Developments</b> .....	315
	Christian Chiarcos, Steven Moran, Pablo N. Mendes, Sebastian Nordhoff, and Richard Littauer	
12.1	Background and Motivation .....	316
12.2	Structural Interoperability for Annotated Corpora .....	320
12.3	Structural Interoperability Between Corpora and Lexical-Semantic Resources .....	325
12.4	Structural Interoperability of Linguistic Databases .....	329
12.5	Conceptual Interoperability of Language Resources .....	334
12.6	Towards a Linguistic Linked Open Data Cloud .....	339
12.7	Summary .....	344
	References .....	345
<b>13</b>	<b>Community Efforts Around the ISOcat Data Category Registry</b> .....	349
	Sue Ellen Wright, Menzo Windhouwer, Ineke Schuurman, and Marc Kemps-Snijders	
13.1	Introduction .....	350
13.2	Historical Perspective .....	351
13.3	Community Support in ISOcat .....	357
13.4	Standardization Community Efforts .....	358
13.5	Infrastructure Community Efforts .....	365
13.6	RELcat a Relation Registry .....	371
13.7	Conclusions and Future Work .....	372
	References .....	372
	<b>Index</b> .....	375



<http://www.springer.com/978-3-642-35084-9>

The People's Web Meets NLP  
Collaboratively Constructed Language Resources  
Gurevych, I.; Kim, J. (Eds.)  
2013, XXIV, 378 p., Hardcover  
ISBN: 978-3-642-35084-9