Creating a More Transparent Internet

On social media, new forms of communication arise rapidly, many of which are intense, dispersed, and create new communities at a global scale. Such communities can act as distinct information bubbles with their own perspective on the world, and it is difficult for people to find and monitor all these perspectives and relate the different claims made. Within this digital jungle of perspectives on truth, it is difficult to make informed decisions on important things like vaccinations, democracy, and climate change. Understanding and modeling this phenomenon in its full complexity requires an interdisciplinary approach, utilizing the ample data provided by digital communication to offer new insights and opportunities. This book gives a comprehensive view on social media communication, the different forms it takes, the impact and the technology used to mine it, and defines the roadmap to a more transparent Web. This interdisciplinary approach will be of interest to researchers in computational linguistics, linguistics, social sciences, media studies, the Semantic Web, and internet technology.

Piek Vossen is Professor in Computational Linguistics at the Vrije Universiteit Amsterdam. He is Co-founder and President of the Global Wordnet Association and has received the Dutch Spinoza Prize in 2013 for his research. He used this prize for projects on language understanding, including mining perspectives in debates like on vaccination.

Antske Fokkens is Professor in Computational Linguistics at the Vrije Universiteit Amsterdam and Associate Professor at the Applied Geometric Algorithms group at the Technical University Eindhoven. Since 2021, she has acted as University Research Chair on methodological aspects of computational linguistics, in particular when used in an interdisciplinary setting.
STUDIES IN NATURAL LANGUAGE PROCESSING

Series Editor:
Chu-Ren Huang, The Hong Kong Polytechnic University

Associate Series Editor:
Qi Su, Peking University

Editorial Board Members:
Nianwen Xue, Brandeis University
Maarten de Rijke, University of Amsterdam
Lori Levin, Carnegie Mellon University
Alessandro Lenci, Universita degli Studi, Pisa
Francis Bond, Nanyang Technological University

Volumes in the SNLP series provide comprehensive surveys of current research topics and applications in the field of natural language processing (NLP) that shed light on language technology, language cognition, language and society, and linguistics. The increased availability of language corpora and digital media, as well as advances in computer technology and data sciences, has led to important new findings in the field. Widespread applications include voice-activated interfaces, translation, search engine optimization, and affective computing. NLP also has applications in areas such as knowledge engineering, language learning, digital humanities, corpus linguistics, and textual analysis. These volumes will be of interest to researchers and graduate students working in NLP and other fields related to the processing of language and knowledge.

Also in the series
Douglas E. Appelt, Planning English Sentences
Madeleine Bates and Ralph M. Weischedel (eds.), Challenges in Natural Language Processing
Steven Bird, Computational Phonology
Peter Bosch and Rob van der Sandt, Focus
Pierette Bouillon and Federica Busa (eds.), Inheritance, Defaults and the Lexicon
Ronald Cole, Joseph Mariani, Hans Uszkoreit, Giovanni Varile, Annie Zaenen, Antonio Zampolli, and Victor Zue (eds.), Survey of the State of the Art in Human Language Technology
David R. Dowty, Lauri Karttunen, and Arnold M. Zwicky (eds.), Natural Language Parsing
Ralph Grishman, Computational Linguistics
Graeme Hirst, Semantic Interpretation and the Resolution of Ambiguity
András Kornai, Extended Finite State Models of Language
Kathleen R. McKown, Text Generation
Martha Stone Palmer, Semantic Parsing for Finite Domains
Terry Patten, Systemic Text Generation as Problem Solving
Ehud Reiter and Robert Dale, Building Natural Language Generation Systems
Manny Rayner, David Carter, Pierrette Bouillon, Vassilis Digalakis, and Matis Wiren (eds.), The Spoken Language Translator
Michael Rosner and Roderick Johnson (eds.), Computational Lexical Semantics
Richard Sproat, A Computational Theory of Writing Systems
George Anton Kiraz, Computational Nonlinear Morphology
Nicholas Asher and Alex Lascarides, Logics of Conversation
Margaret Masterman (edited by Yorick Wilks), Language, Cohesion and Form
Walter Daelemans and Antal van den Bosch, Memory-Based Language Processing
Chu-Ren Huang, Nicoletta Calzolari, Aldo Gangemi, Alessandro Lenci, Alessandro Oltramari, and Laurent Prévot (eds.), Ontology and the Lexicon: A Natural Language Processing Perspective
Thierry Poibeau and Aline Villavicencio (eds.), Language, Cognition, and Computational Models
Bind Liu, Sentiment Analysis: Mining Opinions, Sentiments, and Emotions, Second Edition
Marcos Zampieri and Preslav Nakov, Similar Languages, Varieties, and Dialects: A Computational Perspective
Creating a More Transparent Internet
The Perspective Web

Edited by

PIEK VOSSEN
VU University Amsterdam

ANTSKE FOKKENS
VU University Amsterdam

Section Editors
CAMIEL BEUKEBOOM
VU University Amsterdam

JULIA NOORDEGRAAF
VU University Amsterdam

THOMAS POELL
VU University Amsterdam

IVAR VERMEULEN
VU University Amsterdam
Contents

List of Contributors  page x

1 Introducing the Perspective Web  Piek Vossen and Antske Fokkens  1
  1.1 A Web of Perspectives  1
  1.2 The Urgency of the Perspective Web  2
  1.3 What This Book Is About  3
  1.4 The Structure of This Book  5
References  6

PART I THEORETICAL BACKGROUND
Section Editors: Piek Vossen and Antske Fokkens

2 Perspectives from a Social Psychological and Communication Scientific Perspective  Ivar Vermeulen and Camiel Beukeboom  9
  2.1 Introduction: Perspectives as (Embedded) Belief and Attitude Structures  9
  2.2 Recipient Perspectives: The Role of Perspectives in Processing Incoming Information  12
  2.3 Sender’s Perspectives: How a Sender’s Perspective is Reflected in Communication and Language Use  15
  2.4 Social and Interactive Dynamics  18
  2.5 Conclusion  25
References  26

3 Computational Linguistics for Subjectivity  Preslav Nakov  31
  3.1 Sentiment Analysis  31
  3.2 Other Subjectivity Tasks  36
Contents

3.3 Machine Learning Models 39
3.4 Conclusion and Recommended Readings 43
References 44

PART II SOCIAL IMPACT
Section Editors: Ivar Vermeulen and Camiel Beukeboom
4 Perspectives in a Social Context: The Role of Communication Ivar Vermeulen and Camiel Beukeboom 57
5 The Micro Level: Linguistic Perspective in Written Discourse Kobie van Krieken and José Sanders 60
5.1 Perspective in Language and Discourse 60
5.2 A Text-Linguistic Approach to Perspective in Narrative 62
5.3 Focalization 64
5.4 Demonstration versus Invasion 65
5.5 Quotations 66
5.6 Direct Speech 68
5.7 Indirect Speech 68
5.8 Free Indirect Speech 69
5.9 Multiple Perspectives 70
References 71
5.10 Appendix: I Sympathise with You Guys, Lots of Luck 73
6 The Meso Level: Perspectives in a Social Context Rachel Neo 75
6.1 The Meso Level 75
6.2 Theories of Computer-Mediated Communication and New Media 75
6.3 Traditional Theories of Media Effects 82
6.4 Conclusion 86
References 88
7 The Macro Level: Perspectives Embedded in Society, Culture, and Technology Hong Vu 94
7.1 Introduction 94
7.2 Technologies for Personalized Content 95
7.3 Selective Exposure 98
7.4 Homophily, Echo Chamber, and Filter Bubble 100
7.5 Political Polarization and Audience Fragmentation 104
7.6 Concluding Remarks 105
References 106
PART III MEDIATING PERSPECTIVES

Section Editors: Julia Noordegraaf and Thomas Poell

8 The Mediation of Online Information Julia Noordegraaf and Thomas Poell

References 117

9 The Source and Its Encoding: Reflections on Metadata in Digitized and Born-Digital Media Collections Eric Hoyt

9.1 Introduction 119
9.2 Metadata in Digitized Magazine Collections 121
9.3 Podcasts, RSS Feeds, and the Encoding of Born-Digital Objects 124
9.4 Conclusion 127
References 128

10 Knowledge-Making on Techno-Commercial Platforms: The Example of Facebook Jonas Anderson Schwarz

10.1 Introduction 129
10.2 The Normative Power of Prescribed Social Action 130
10.3 The Corporate Mission 133
10.4 Objectives and Ethics 136
10.5 Conclusion: Perspective-Making on, of, and through Facebook 138
References 140

11 Content, Form, and Reception: Perspectives from Digital Media Data Christina Neumayer

11.1 Introduction 143
11.2 A Brief Note on Perspectives from Political Protest 144
11.3 Content 145
11.4 Form 147
11.5 Reception 149
11.6 Perspectives from Digital Data 151
References 153

12 Quality and Perspectives Davide Ceolin, Julia Noordegraaf, and Lora Aroyo

12.1 Introduction 156
12.2 Information Quality Dimensions 158
12.3 Related Work 160
12.4 A Tool for Information Quality Assessment 161
12.5 Conclusion and Future Work 166
References 167
PART IV MINING AND MODELING PERSPECTIVES
Section Editors: Piek Vossen and Antske Fokkens

13 Mining and Modeling Perspectives  Piek Vossen and Antske Fokkens 171

14 Natural Language Processing Tasks for the Extraction of Perspectives  Chantal van Son, Roser Morante, and Piek Vossen 173
14.1 Introduction 173
14.2 Modeling and Processing Attribution 176
14.3 Factuality Profiling 178
14.4 Natural Language Inference 182
14.5 Micro-Propositions 184
14.6 Conclusions 187
References 188

15 Toward Automatic Discovery of Diverse Perspectives  Sihao Chen, Daniel Khashabi, and Dan Roth 193
15.1 Introduction 193
15.2 Minimal Perspective Discovery: Tasks and Challenges 195
15.3 Related Work 197
15.4 The Perspectrum Dataset 198
15.5 Building Systems Based on Perspectrum 199
15.6 Toward Automated Perspective Discovery 203
15.7 Discussion and Conclusion 204
References 205

16 Formal Representation and Extraction of Perspectives  Aldo Gangemi and Valentina Presutti 208
16.1 Viewpoints and Perspectives: A Landscape 208
16.2 What is a Perspective? 211
16.3 Representing and Extracting Perspectives 213
16.4 Conclusions 225
References 226

17 The User Perspective in Professional Information Search  Suzan Verberne 229
17.1 Introduction 229
17.2 Professional Search 231
17.3 Personalized IR 232
17.4 Explainable Recommendation and Search 234
17.5 Toward Explainable Professional Search 235
# Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.6</td>
<td>Conclusions</td>
<td>238</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>239</td>
</tr>
<tr>
<td>18</td>
<td>Harvesting Perspectives in Social Media</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td>Tommaso Caselli</td>
<td></td>
</tr>
<tr>
<td>18.1</td>
<td>Introduction</td>
<td>244</td>
</tr>
<tr>
<td>18.2</td>
<td>Methods</td>
<td>245</td>
</tr>
<tr>
<td>18.3</td>
<td>Case Studies</td>
<td>249</td>
</tr>
<tr>
<td>18.4</td>
<td>Conclusions</td>
<td>255</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>255</td>
</tr>
<tr>
<td>19</td>
<td>GRaSP: A Model for the Perspective Web</td>
<td>260</td>
</tr>
<tr>
<td></td>
<td>Piek Vossen</td>
<td></td>
</tr>
<tr>
<td>19.1</td>
<td>Introduction</td>
<td>260</td>
</tr>
<tr>
<td>19.2</td>
<td>A Model for the Perspective Web</td>
<td>262</td>
</tr>
<tr>
<td>19.3</td>
<td>The GRaSP Model</td>
<td>263</td>
</tr>
<tr>
<td>19.4</td>
<td>Related Work</td>
<td>269</td>
</tr>
<tr>
<td>19.5</td>
<td>A Future Perspective on the Web</td>
<td>271</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>275</td>
</tr>
</tbody>
</table>
Contributors

Lora Aroyo
Google Research NYC

Camiel Beukeboom
Vrije Universiteit Amsterdam

Tommaso Caselli
University of Groningen

Davide Ceolin
Centrum Wiskunde & Informatica, Amsterdam

Sihao Chen
University of Pennsylvania

Antske Fokkens
Vrije Universiteit Amsterdam

Aldo Gangemi
University of Bologna and ISTC-CNR, Rome

Eric Hoyt
University of Wisconsin-Madison

Daniel Khashabi
Allen Institute for Artificial Intelligence
List of Contributors

Kobie van Krieken
Radboud University Nijmegen

Roser Morante
Vrije Universiteit Amsterdam

Preslav Nakov
Qatar Computing Research Institute, HBKU

Rachel Neo
University of Hawaii at Manoa

Christina Neumayer
University of Copenhagen

Malvina Nissim
University of Groningen

Julia Noordegraaf
University of Amsterdam

Thomas Poell
University of Amsterdam

Valentina Presutti
University of Bologna

Dan Roth
University of Pennsylvania

José Sanders
Radboud University Nijmegen

Jonas Anderson Schwarz
Södertörn University Stockholm

Chantal van Son
Vrije Universiteit Amsterdam
xii

List of Contributors

Suzan Verberne
University of Leiden

Ivar Vermeulen
Vrije Universiteit Amsterdam

Piek Vossen
Vrije Universiteit Amsterdam

Hong Vu
University of Kansas