

# Preface

I hope for this book to serve as a good starting point for students and researchers in Semantic Web (SW) interested in discovering what Natural Language Processing (NLP) has to offer. At a time when Open Data is becoming increasingly popular, there is a pressing demand for tools to help the SW community transform those data into a shareable, normalized format, making all these data accessible as Linked Open Data. But a large portion of the data held by organizations seeking to make their data openly accessible are not stored in tables, but in much less structured forms, that is, textual forms such as reports, notes, memos, and articles. Manually generating structured information from them seems like an insurmountable task. Certainly, NLP can help uncovering the information held in text, thus augmenting the real content of the Semantic Web in a significant and lasting way.

My main goal is not just to foster interest in NLP in the readership, but awareness of how useful it can be to the Semantic Web community. I will not delve very deeply into linguistic principles, but instead focus on how NLP approaches different kinds of problems and provides solutions to them. My aim is also to show how, for the past 40 years, researchers in NLP have been interested in problems closely related to the ones faced by the Semantic Web community. Problems such as ambiguity and linking of knowledge are not specific to one field or the other, but central to both.

This book covers the basics of Natural Language Processing (NLP), with a focus on Natural Language Understanding (NLU). Here, understanding refers to semantic processing, Information Extraction, and knowledge acquisition, which I see as the key links between the SW and NLP communities. Much emphasis will be placed on mining sentences in search of entities and relations. In our quest in NLU, we will encounter challenges for various text analysis tasks, including part-of-speech tagging, parsing, semantic disambiguation, Named Entity Recognition, and Relation Extraction. I will present the standard algorithms associated with these tasks so as to provide an understanding of the fundamental concepts. Furthermore, I chose to emphasize the importance of experimental design and result analysis, and for doing

so, most chapters show small experiments on corpus data with quantitative and qualitative analysis of results.

I assume that readers are familiar with the Semantic Web and are looking to learn about NLP in order to expand their horizons. That being said, the book provides enough information for a reader new to both fields to understand their underlying principles and the challenges they face. Also, the reader should be familiar with algorithms and simple programming principles, as this book will often use algorithms to describe problem-solving approaches.

Since I chose to cover a small number of simple algorithms in details, I do include a *Further Reading* section in most chapters for links to relevant state-of-the-art research in which readers can find more complex algorithms. I believe that understanding the fundamentals within basic algorithms is the best preparation for understanding more complex algorithms. I hope that through this book, important challenges in NLP will become familiar to the reader and that the book will stimulate the reader's interest in investigating them further.

Montreal, Canada

Caroline Barrière



<http://www.springer.com/978-3-319-41335-8>

Natural Language Understanding in a Semantic Web  
Context

Barrière, C.

2016, XVII, 317 p., Hardcover

ISBN: 978-3-319-41335-8