

# Natural Language Processing

Matthias Hagen   Martin Potthast   Benno Stein  
Henning Wachsmuth   Matti Wiegmann   Magdalena Wolska

# Contents

- I. Introduction
- II. Corpus Linguistics
- III. Text Models
- IV. Language Models
- V. Words
- VI. Syntax
- VII. Semantics
- VIII. Discourse
- IX. Ethics, Fairness, and Bias
- X. NLP Applications

# Objectives

- ❑ Understand natural language phenomena
- ❑ Understand and explain basic concepts of natural language processing (NLP)
- ❑ Be able to implement and apply algorithms to process natural language
- ❑ Be able to compare and combine approaches to solve NLP tasks
- ❑ Be able to self-educate

# Related Fields

1. Linguistics

[paradigms, models]

2. Statistics

3. Psychology

4. Machine Learning and AI

[methods, algorithms]

5. Data Mining

6. Information Retrieval

7. Knowledge Processing

# Related Fields

9. Search engines
10. Writing assistants
11. Speech processing
12. Machine translation
13. Text generation, dialog systems, and chatbots
14. Optical character recognition, text-to-image, and image-to-text systems

[applications]

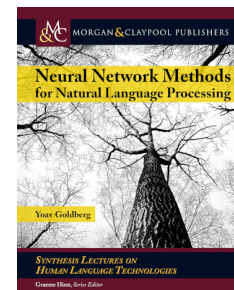
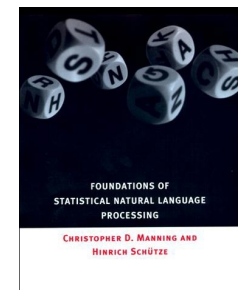
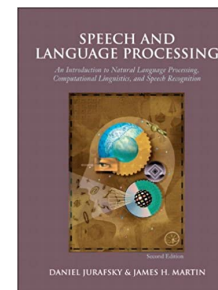
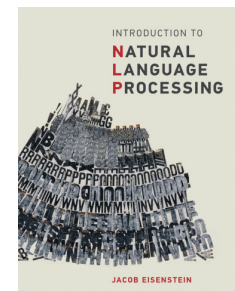
# Literature

## Natural Language Processing:

- ❑ J. Eisenstein.  
*Natural Language Processing*  
MIT Press 2019. [github.com/jacobeisenstein/gt-nlp-class](https://github.com/jacobeisenstein/gt-nlp-class)
- ❑ D. Jurafsky, J. H. Martin.  
*Speech and Language Processing*  
Prentice Hall 2000, 2009, 2025. [web.stanford.edu/~jurafsky/slp3/](http://web.stanford.edu/~jurafsky/slp3/)
- ❑ C. D. Manning, H. Schütze.  
*Foundations of Statistical Natural Language Processing*  
MIT Press 1999. [nlp.stanford.edu/fsnlp/](http://nlp.stanford.edu/fsnlp/)

## Deep Learning for NLP:

- ❑ Y. Goldberg.  
*Neural Network Methods for Natural Language Processing*  
Morgan & Claypool Publishers 2017.



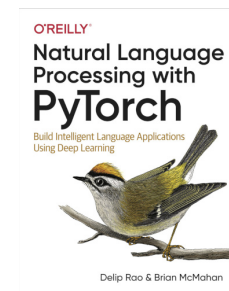
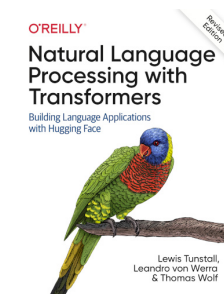
# Literature

## Applied Deep Learning for NLP:

- ❑ D. Rao, B. McMahan.  
*Natural Language Processing with PyTorch*  
O'Reilly 2019. [github.com/delip/PyTorchNLPBook](https://github.com/delip/PyTorchNLPBook)
- ❑ L. Tunstall, L. v. Werra, T. Wolf.  
*Natural Language Processing with Transformers*  
O'Reilly 2022. [github.com/nlp-with-transformers](https://github.com/nlp-with-transformers)

## Text Mining:

- ❑ C. Biemann, G. Heyer, U. Quasthoff.  
*Wissensrohstoff Text: Eine Einführung in das Text Mining*  
[Springer 2022.](https://www.springer.com/9783708921111)



## Remarks:

- ❑ Peer-reviewed research in natural language processing is published at a number of “core” conferences, journals, and monograph series. The [ACL Anthology](#) provides an overview.
  - Important NLP conferences: ACL, EMNLP, NAACL;  
COLING, EACL, AACL, CoNLL, CICLing, IJCNLP, INLG;  
shared tasks at SemEval
  - Important NLP journals: [TACL](#), [Computational Linguistics](#)
  
- ❑ Most important conferences from related fields:
  - Machine learning: NeurIPS, ICLR, ICML, ECML
  - Artificial intelligence: AAI, IJCAI, ECAI
  - Data mining: KDD, SDM, ICDM, PAKDD
  - [Information retrieval](#): SIGIR, WSDM, ECIR; shared tasks at TREC, CLEF, NTCIR, FIRE
  - Web: WebConference, Hypertext



## Remarks:

- Natural language processing research and development is based on many open source projects; a selection:
  - Text annotation: [prodigy](#), [Label Studio](#), [doccano](#)
  - Language processing: [Apache UIMA](#), [spaCy](#), [Stanford NLP: Stanza](#)
  - Algorithm collections: [Apache OpenNLP](#), [NLTK](#), [Stanford NLP](#), [GATE](#)
  - Machine learning: [scikit-learn](#), [Hugging Face](#), [flair](#)