

Introduction to Natural Language Processing

Organizational Information

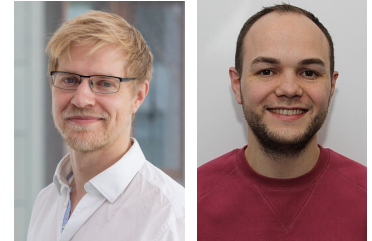
Henning Wachsmuth

<https://ai.uni-hannover.de>

Organizational Information

Bachelor course

- **Lectures.** Henning Wachsmuth
- **Tutorials.** Timon Ziegenbein
- **Languages.** English, Python



Information

- **Web.** <https://www.ai.uni-hannover.de/de/teaching/courses/inlp>
- **Stud.IP.** <https://studip.uni-hannover.de/dispatch.php/course/overview?cid=cfe0060d32c4224fb9d6479131521fce>

Time and location

- **Lectures.** Thursday 13:00–14:30, 3702-031 (Schneiderberg 32)
- **Tutorials.** Wednesday 15:30–17:00, 3408-MZ2 (Appelstr. 9A)
First tutorial introduces Python and the assignment concept

Consultation?

- Set up appointment via e-mail: h.wachsmuth@ai.uni-hannover.de

Organizational Information

Teaching at NLP Group, <https://ai.uni-hannover.de/en/studies/courses>

Courses

- **Introduction to Natural Language Processing (bachelor, summer)**. NLP fundamentals, from rule-based methods to statistical methods
- **Statistical Natural Language Processing (master, winter)**. Core NLP, from statistical methods to neural methods
- **Computational Argumentation (master, summer)**. State-of-the-art NLP, advanced methods in a specific research context

Seminars

- Natural Language Processing (bachelor, winter)
- Natural Language Generation (master, summer)

Labs and projects

- Argumentation Technology (master, summer)
- Ethical Artificial Intelligence (master, winter)

Organizational Information

This Course

Overall goals

- Learn major skills needed to approach fundamental natural language processing (NLP) tasks

Contents

- Several rule-based and statistical NLP techniques
- Several NLP tasks and approaches
- Required basics of linguistics and empirical methods

Competences

- Understanding of theory and practice of NLP
- Design and implementation of NLP methods for given tasks
- Scientific experiments and evaluations on large amounts of text

Organizational Information

Outline of the Course

Introduction

1. Overview

Modeling linguistic knowledge

2. Basics of linguistics
3. NLP using rules
4. NLP using lexicons

Modeling statistical patterns

4. Basics of empirical methods
5. NLP using regular expressions
6. NLP using context-free grammars
7. NLP using language models

Application

9. Practical issues

Organizational Information

Course Elements

Teaching

- **Lectures.** Presentation of course content and organizational info
- **Tutorials.** Presentation of assignments and solutions, Q&A

Assignment sheets (details in first tutorial)

- **Amount.** 6 in total, bi-weekly (~50% written, ~50% programming)
First sheet published on April 15; to be submitted by April 29, 23:59 (UTC+2)
- **Group work.** You need to submit with 2–4 people
- **Bonus.** (a) Min. 60% of all points: exam grade + 1/3, (b) Min. 85%: + 2/3
Example for (b): grade of 2.7 is changed to 2.0; only grades < 5.0 can be improved

Exams

- **Written.** 90 minutes, questions on all lecture parts, English, July 22
- **Registration.** May 15–31, 2024
More details on the exam later

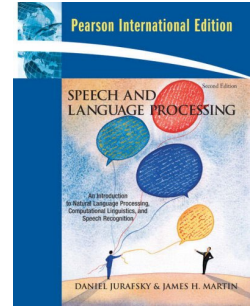
Organizational Information

Textbooks (Not Obligatory)

Speech and Language Processing

(Jurafsky and Martin, 2009)

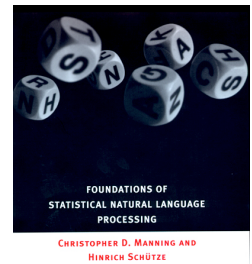
- Oriented towards computational linguistics
- Comprehensive
- **Draft of upcoming third edition freely available**



Foundations of Statistical Natural Language Processing

(Manning and Schütze, 1999)

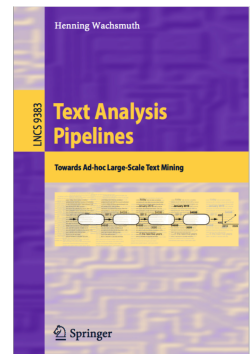
- More oriented towards computer science
- Comprehensive, a bit outdated



Text Analysis Pipelines

(Wachsmuth, 2015)

- Rather oriented towards computer science
- Focused on NLP processes
- Book preprint freely available



References

- **Jurafsky and Martin (2009)**. Daniel Jurafsky and James H. Martin. Speech and Language Processing: An Introduction to Natural Language Processing, Speech Recognition, and Computational Linguistics. 2nd edition, Prentice-Hall, 2009.
Free draft of 3rd edition: <https://web.stanford.edu/jurafsky/slp3/>
- **Manning and Schütze (2009)**. Christopher D. Manning and Hinrich Schütze. Foundations of Statistical Natural Language Processing. MIT Press, 1999.
- **Wachsmuth (2015)**. Henning Wachsmuth: Text Analysis Pipelines — Towards Ad-hoc Large-scale Text Mining. LNCS 9383, Springer, 2015.
Free preprint: https://webis.de/downloads/publications/papers/wachsmuth_2015b.pdf