

information for
seminars, theses,
and research

Seminar Natural Language Generation (NLG) — Part 4

Basics of Scientific Presentation

Henning Wachsmuth

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



Outline

- **Literature research**

Types, quality, reading, acquisition, and organization



<https://pixabay.com>

- **Oral presentations**

Content, structure, style, talking, and timing



<https://commons.wikimedia.org>

- **Written presentations**

Content, structure, style, citations, and plagiarism



<https://pixabay.com>



Literature research

Doing literature research

■ Literature research

- Fundamental task in science
- Time-intensive and tedious — but necessary
- Often, the first task to be done



■ Goals in general

- Obtain all information relevant to the scope of a given problem
- Obtain background information
- Obtain evidence for your or others' claims

... and similar reasons

■ Additional goals in science

- Find alternative approaches or perspectives
- Find out if your approach to a problem is new
- Don't reinvent the wheel

You are rarely the first to work on a problem. If you are, what does that tell you?

Selecting literature

■ Types of literature

1. **Books.** Theory, basics, approved techniques
2. **Journal articles.** Completed research lines
3. **Conference papers.** State-of-the-art research
In NLP/CS, major publication type
4. **Workshop papers.** New ideas, ongoing research
5. **Preprints.** Papers inofficially published on free web platforms
6. **Conference/Online tutorials.** Easy access to basics and techniques
7. **Popular science magazines.** Public communication of research lines
8. **Other websites.** Anything



<https://pixabay.com>

■ What types to use

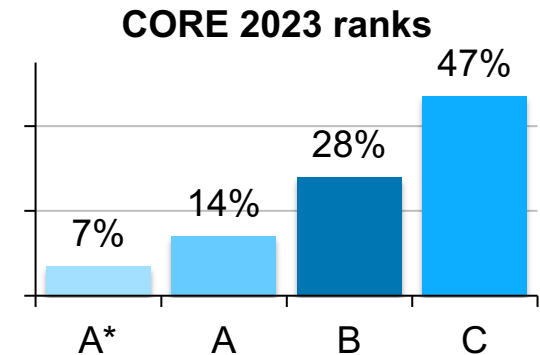
- Generally, literature should be peer-reviewed (mostly holds for Type 1–4)
- Rule of thumb: **books > journals > conferences > workshops**
> tutorials > magazines > websites **preprints?**
- But, for example: top conferences > average journals

The symbol > stands for “preferred” over here

Assessing quality of literature

■ Conference rankings

- Best is A*/A; B still reasonable; unranked doubtful
No ranking has complete coverage, though
- A well-reputed ranking is CORE core.edu.au/conference-portal



■ Journal rankings

- Partitioned into quartiles Q1–Q4 (Q1 best, Q2 still reasonable)
- Journals also have impact factors derived from citations (reasonable as of 4+)
- A good place to check Journals is Scimago www.scimagojr.com

■ Number of citations

- Indicator of importance, rather for *relative* comparisons within a topic
Notice: Newer papers naturally tend to have fewer citations
- A good resource for citation numbers is Google Scholar scholar.google.de

■ Disclaimer

- Good and bad research appears at all places
- Often, only reading helps

Reading and finding literature

- **Reading papers efficiently**

1. Read abstract, introduction, and conclusion
2. Look at figures and tables
3. Decide whether worth reading everything
4. Read goal-driven

Specify questions to be answered during reading



<https://pixabay.com>

- **Finding the next paper**

- Follow promising references at the end of a paper
- Find promising papers citing a paper
- Learn to identify the best search terms

Rule of thumb: As specific as possible, but as abstract as needed

- **Getting started in the seminar**

1. First read the literature that we provide
2. Then find further literature

Acquiring literature

■ Obtaining papers

- Many papers freely available online
- Others might be free from a university network
- If neither, maybe your advisors can help



<https://pixabay.com>

■ Important sources

- **ACL Anthology** for most core NLP papers aclweb.org/anthology
- **ACM Digital Library** for many important computer science papers dl.acm.org
- **DLBP** for any literature related to computer science dblp.dagstuhl.de
- **Google Scholar** for any scientific literature scholar.google.de

... along with general web search

■ Accessing books

- Check whether available in the library (physically or digitally)
- Some accessible online, for example, on Google Books books.google.de

Purchasing books can make sense when of continuous importance to you

Organizing literature

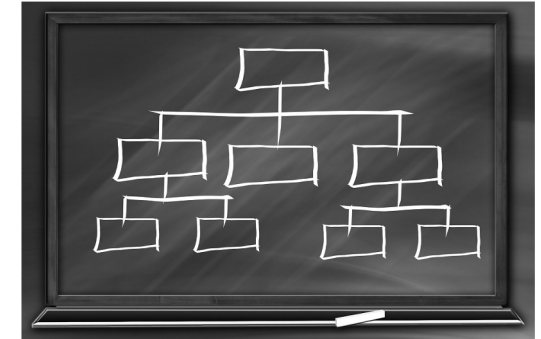
■ Literature organization

- Maintain overview, start from the beginning
- “Extra” effort will pay off

■ Create logical folder structure

- Build your own view of the field
- Logically subdivide topics, but don’t over-engineer

For instance, `./literature/computational-sociolinguistics/social-bias/` — but maybe not deeper



■ Rename all literature consistently

- Simplifies browsing and finding
- We use `<1stauthor><2digityear>-<full-title-lower-case-no-special-chars>.pdf`

For example: `ajjour17-unit-segmentation-of-argumentative-texts.pdf`

■ Organizing meta-information

- Bibliographical information needed when citing literature
- Store bibtex of literature whenever available

Learn more on en.wikipedia.org/wiki/BibTeX; many pages provide bibtex entries

**tools for literature
organization exist**
e.g., Citavi and Zotero



Oral presentations

Content of your talk

■ Scientific presentation is storytelling

- Tell a coherent story with a central theme
- But: Avoid irrelevant information, stay on topic
- Plan what points to make and how to get there
- Don't be complete, be selective
Partly different for articles (and specific talks like lectures)
- Make it exciting, show importance
- Avoid surprise: Clarify why you tell something

■ Science needs to be understood

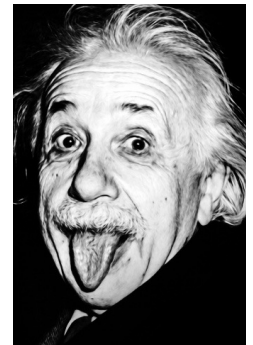
- Adjust complexity to audience
- Leave out formal things, unless needed
Partly different in articles and specific talks like lectures
- Be precise and clear
- Introduce terms, use consistently
- Figures and examples help

"Sometimes **reality**
is too complex.



Jean-Luc Godard

Stories
give it a form."



Albert Einstein

"Everything should be
as **simple** as possible,
but **not simpler.** "

Figures

■ Figures

- Charts, diagrams, graphs, pictures, drawings, ...
- Slides are visual
- **Rule of thumb.** (Almost) No slide without figure

■ What to use figures for

- **Primary.** Replace text; visually explain concepts, ...
- **Secondary.** Support your message with pictures

As often done in this presentation

■ Formatting

- Vector graphics whenever possible
- Others: Optimize sharpness, scale down smartly
Avoid scaling > 100%; 50% is better than 53% — why?
- **Never squeeze or stretch the aspect ratio**
If needed, cut off parts of figures instead
- Check readability of text in figures

“a **picture** is worth
a **1000 words**”



“unsharpness
is the mistake that even
lay persons see”

Herbert Kania

Colors

■ Colors in general

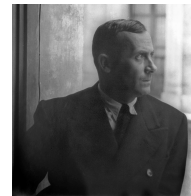
- Presentations are visual, make use of colors
- Fewer colors create a more clear style
- But natural colors have an appeal, too



■ Font colors for important points

- Use colors consistently
- Not too colorful
- I use dark blue here for highlighting

And a cyan-like color for quotes



Joan Miro

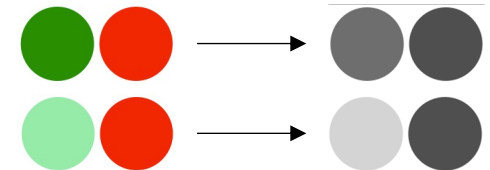
”I try to **apply colors like words** that shape poems, like notes that shape music.“

■ Support your messages

- Use the same color for the same concept
- Can create connections even across slides

■ Color vs. brightness

- Think of color blind people — contrast helps



Tables (and matrices)

■ Tables for what?

- Presenting numerical results
- Comparing ideas, approaches, or similar
- Listing attribute values of multiple instances
... and similar

■ Table style

- **Amount.** Show only important rows and columns, to keep a table easy to digest
In articles, comprehensiveness may be preferred
- **Alignment.** Text left, numbers right
- **Lines.** Recommended to use only horizontal lines
Except for matrices

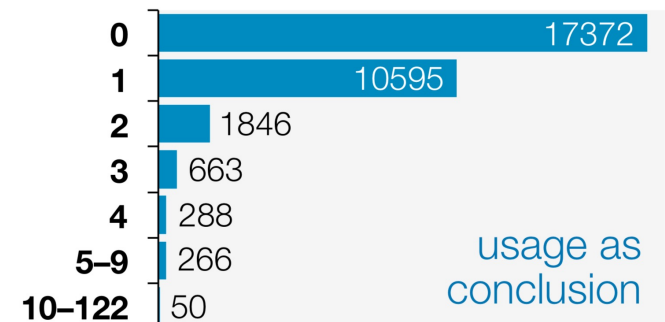
■ Tables vs. charts

- Prefer tables if exact numbers are important
- Prefer charts if differences should be stressed

best results for each ranking approach

#	Dimension	τ	best	worst
1	PageRank	0.28	15	3
2	Number	0.19	6	1
3	Sentiment	0.12	12	4
4	Frequency	0.10	11	9
5	Similarity	0.02	9	10
6	Random	0.00	8	7

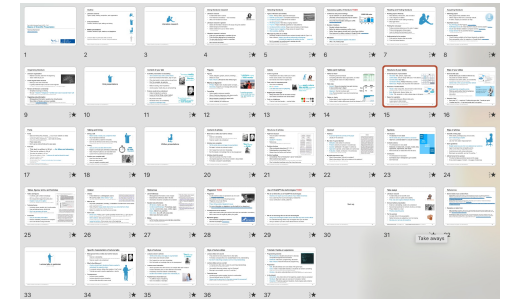
	true	false
true	TP	FP
false	FN	TN



Structure of your slides

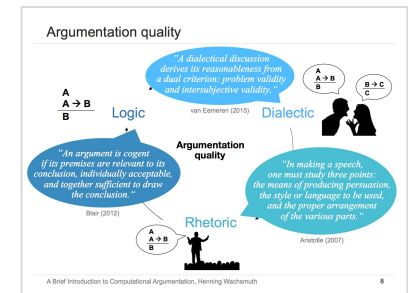
■ Overall structure of presentation

- **Title slide.** Title, authors, maybe date
- **Outline slides.** Only for longer talks (min. ~20 minutes)
- **Content slides.** Your story
- **Conclusion slide(s).** Always! Takeaways, outlook
- **References.** Prepare, but only show when asked for



■ Structure of content slides

- **Header.** Clear unique title, should match content of body
Interesting observation: Titles often not read by the audience
- **Body.** Bullet points, figures, tables, etc.
- **Footer.** Title, presenter, page number/progress, no date



■ Space for separation

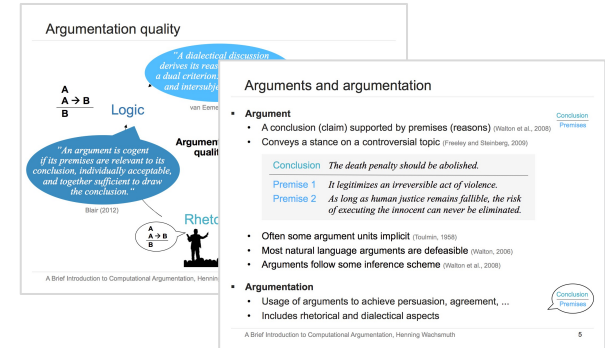
- Leave space between different slide parts
- Leave space to slide borders
Harder to read there + border sometimes clipped

Style of your slides

■ General slide style

- Decide what to put on slide and what to say
- Vary slides to maintain attention
- Animations only when useful, use consistently
- **Clarify what is from you and what from others!**

Also see notes on citations and plagiarism below



■ Text style

- **Avoid grammar and spelling errors**
- Write key points rather than full sentences
- AIA & AUA

May be different in specific talks like lectures

Always introduce acronyms & Avoid unnecessary acronyms

■ Amount of text

- Some say 7x7 — maximum 7 bullet points per slide, 7 words per point
- I'd rather say 3x3 — 3 top-level points with 3 sub-points

Grammar.

The difference between
knowing your shit and
knowing you're shit.



<https://flickr.com>

Fonts

■ Fonts

- Sans-serif fonts (Arial, Verdana, ...) much more readable on slides
Ambiguity speaks against Arial ("lll") ... but Arial available on all machines
- *Serif* fonts (Times, Garamond, ...) are made for printing
I use them on slides for example texts only
- Prefer simple fonts
- Don't use too narrow fonts just to save space

■ Font size

- This text is written in 26 pt — for titles and stressing
- This text is written in 24 pt
- This text is written in 21 pt
- This text is written in 18 pt — minimum for text that should be read
- This text is written in 16 pt
- This text is written in 14 pt
- This text is written in 12 pt — minimum for extra information that may be skipped
- This text is written in 10 pt
- This text is written in 8 pt
- This text is written in 6 pt — maybe for texts that should on purpose not be readable

Talking and timing

■ Giving a talk

- Match words on slides, but complement them
- No pre-phrased sentences
- Look at audience, speak to everybody
- Don't be *too* formal, but be serious, avoid slang

Occasional jokes may be nice, if you know how to use them



■ Timing

- Use your time, but stick with time limit
- Expect ≥ 2 minutes per (animated) content slide
- Rule of thumb: Audience can read slide twice



■ Practice your complete talk!

- How much time do you need?
- Does your story work?
- Can you explain everything well?





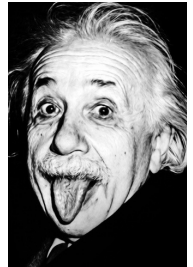
Written presentations

Content of articles

■ Most hints on talks also hold for articles

- Science is storytelling
- Science needs to be understood

Seminar: No scientific novelty expected, rather summarize and discuss



■ Articles more complete

- Tell the whole story, avoid gaps in argumentation
- But again: Include only relevant content
- Don't expect too much prior knowledge
- But: No details on knowledge that can be presupposed

**"Don't make
me think."**



Steve Krug

■ Articles should be sound

- Need to be precise more than in talks
- Use logical arguments, from broad context to deep details
- Formalize concepts if needed/helpful

Structure of articles

■ High-level structure

- Title and author information
- Abstract
- Usually 4–7 sections
- References

... and sometimes appendices (not in seminar article!)

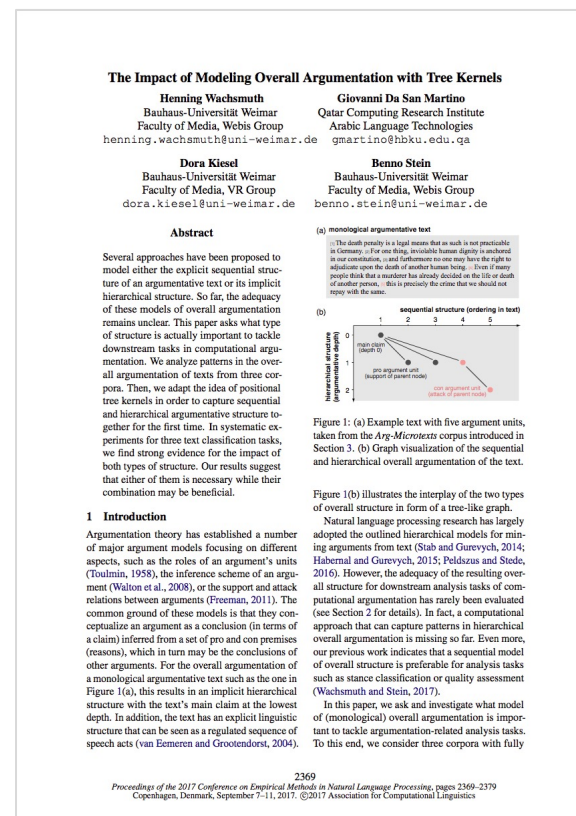
■ Section structure

- Often numbered subsections (2.1, 2.2, ...)
- Subsubsections unnumbered, if any
- Always have text before sub⁺sections

■ Section headings

- Conventional: First is "Introduction", last is "Conclusion"
- Other sections partly topic-specific

Other section names exist, but they do not necessarily match seminar articles



Abstract

▪ Abstract

- A concise high-level summary of the paper
- Usually 5–10 sentences

▪ My view of a good abstract

- Motivation and context (1 sentence)
- Problem and why not solved (1–2 sentences)
- Question addressed in the paper (1 sentence)
- Approach in general, some details (2–3 sentences)
- Evaluation, results, conclusion (1–3 sentences)

For seminar articles, may differ a bit (since other people's work presented)

▪ My PhD advisor's view

- What is the problem? Why is it a problem?
- What is the solution? Why is it a solution to the problem?

Notice that this view is NOT in conflict with mine

Abstract

Several approaches have been proposed to model either the explicit sequential structure of an argumentative text or its implicit hierarchical structure. So far, the adequacy of these models of overall argumentation remains unclear. This paper asks what type of structure is actually important to tackle downstream tasks in computational argumentation. We analyze patterns in the overall argumentation of texts from three corpora. Then, we adapt the idea of positional tree kernels in order to capture sequential and hierarchical argumentative structure together for the first time. In systematic experiments for three text classification tasks, we find strong evidence for the impact of both types of structure. Our results suggest that either of them is necessary while their combination may be beneficial.

Sections

■ Introduction

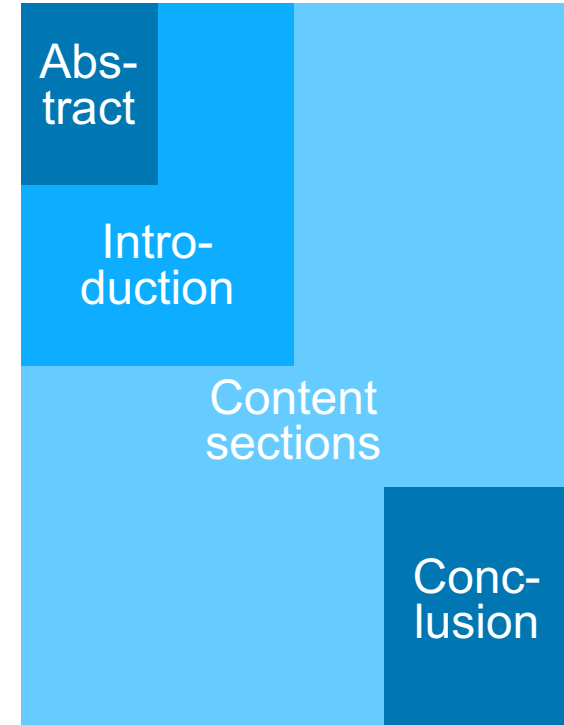
- The abstract in more detail
- Tell the whole story, from context to conclusion
- High-level, understandable for computer scientists

■ Content sections

- The introduction in more detail
- Elaborate on related work, concepts, models, data, approaches, experiments, and results
- More technical, for researchers from the area

■ Conclusion

- The introduction in less detail
- Summarize story in retrospective, give outlook
- Semi-technical



Style of articles

▪ Scientific writing style

- Write clearly, unambiguously, and concisely
- English sentences are short, one statement per sentence
- Again: Avoid grammar and spelling errors

Seminar: Too many of them will negatively affect your grade

- **Don't make things complex**

Common misunderstanding!



<https://pixabay.com>

▪ How to refer to scientific work

- Use “we” form for your own ideas and contributions
- Use in-text citation, “the authors”, or similar to refer to others’ work

Example: “Ajjour et al. (2017) study how... The authors propose to... To this end, they...”

- Also, impersonal form can be used well in certain situations

▪ Article format in the seminar

- Provided template predefines layout and its usage
- 8 two-column pages of content, 1–2 pages of references

Tables, figures, terms, and footnotes

■ Tables and figures

- In articles, just number increasingly
Figure 1, 2, ... Table 1, 2, ... (NOT: Figure 2.1, 2.2, ...)
- Explain in text *and* in caption

Rule of thumb: tables/figures should be clear from the caption

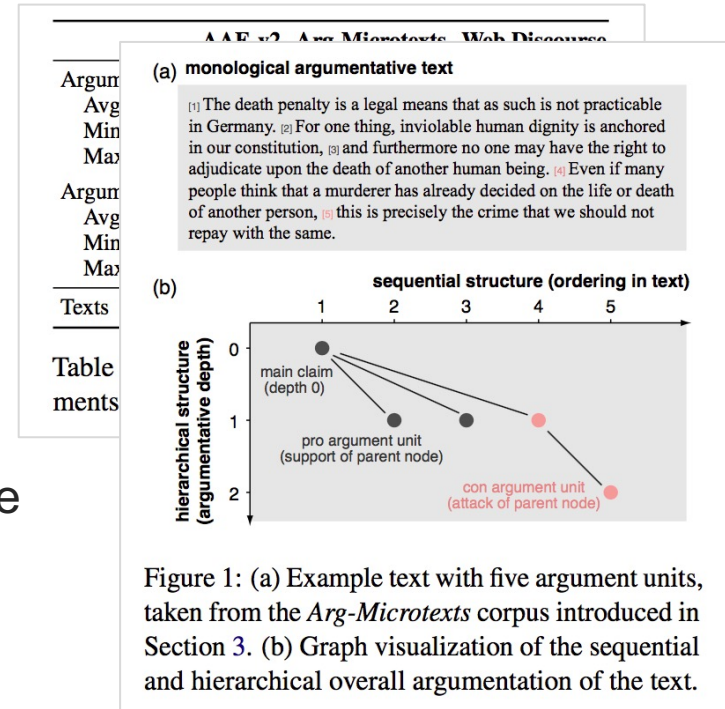
■ Technical terms

- Introduce where needed, don't overformalize
- Use well-defined terms
- Don't use synonyms for terms

Reader is misled to check whether intentional differences exist

■ Footnotes

- Only for secondary information
- Should remain exception (reduces readability)
- In NLP: Don't cite literature using footnotes



alternatives by modeling the stance of each unit towards its parent in the associated tree. This stance can be derived in all corpora.³ All other unit and relation types from the specific models are ignored, since there is no clear mapping between them.

³Alternatively, the stance towards the main claim could be modeled. We decided against this alternative to avoid possibly wrong reinterpretations, e.g., it is unclear whether a unit that attacks its parent always supports a unit attacked by the parent.

Citation

■ Citation

- In-text reference to a bibliographic source
- We use ACL style: **Author names + year**

Other communities use numbers ([1], [2], ...) or acronyms ([ACW17], ...)

mentation, namely, to classify the myside bias and stance of texts. For myside bias, [Stab and Gurevych \(2016\)](#) use features derived from discourse structure, whereas [Faulkner \(2014\)](#) and [Sobhani et al. \(2015\)](#) model arguments to classify stance. [Ong et al. \(2014\)](#) and we ourselves ([Wachsmuth et al., 2016](#)) do similar to assess the quality of persuasive essays and [Reisman, Klebanov et al. \(2016\)](#)

■ What to cite

- Any reuse, paraphrase, summary, or translation of content from any source

Content: Text, figures, and tables

- Rule of thumb: **Always clarify what is from you and what from others**

You also have to cite yourself if you reuse your own sources

- Better one citation too much than one too less

■ How to cite

- **Direct reuse.** Always put in quotes (possibly shorten with [...]), give source

Example: Unit segmentation is "[...] the splitting of a text into argumentative segments" (Ajjour et al., 2017).

- **Other citations.** Give source close-by

Example: As Ajjour et al. (2017) point out, segmentation is the first task of an argument mining pipeline.

- **Large text portions.** Give source once in the beginning

Example: In the following paragraph, we summarize the segmentation approach of Ajjour et al. (2017).

References

▪ List of references

- Bibliographical information at end of paper
- Exactly those references cited in the text
- Information should be complete and consistent

▪ Needed meta-information

- **All literature.** Author, year, title
- **Conferences/Workshops.** Proceedings, pages
- **Journals.** Journal name, issue, number, pages
- **Books.** Edition if any, publisher
- **Web pages.** Give URL with access date

Other meta-information optional

▪ Bibtex

- LaTeX handles references automatically using bibtex

See part on organizing literature above

Aristotle. 2007. *On Rhetoric: A Theory of Civic Discourse* (George A. Kennedy, translator). Clarendon Aristotle series. Oxford University Press.

Beata Beigman Klebanov, Christian Stab, Jill Burstein, Yi Song, Binod Gyawali, and Iryna Gurevych. 2016. *Argumentation: Content, structure, and relationship with essay quality*. In *Proceedings of the Third Workshop on Argument Mining (ArgMining2016)*, pages 70–75. Association for Computational Linguistics.

Stefanie Brüninghaus and Kevin D. Ashley. 2003. *Predicting outcomes of case based legal arguments*. In *Proceedings of the 9th International Conference on Artificial Intelligence and Law*, pages 233–242.

Chih-Chung Chang and Chih-Jen Lin. 2011. *LIB-SVM: A library for support vector machines*. *ACM Transactions on Intelligent Systems and Technology*, 2(3):27:1–27:27.

Plagiarism

■ Plagiarism

- To sell another's ideas or expressions as one's own

See en.wikipedia.org/wiki/Plagiarism

- On purpose or due to lack of giving sources
- Plagiarism *not* a trivial offense, can be a crime

- Proper citing avoids all plagiarism issues

- For more information on plagiarism, see the leaflet we provide

https://studip.uni-hannover.de/sendfile.php?type=0&file_id=e70e1567b07afe42da8a95750a0683f4&file_name=upb-plagiarism-leaflet.pdf



■ Consequences in the seminar

- Major cases lead to failing the seminar (and report to examination office)
- Minor cases can still negatively affect your grade

■ My former group...

- Does research on plagiarism detection
- See the tool picapica www.picapica.org



Use of ChatGPT-like technologies

- **We do not forbid the use of ChatGPT-like technologies**
 - However, we require you to be authors of your slides and articles
 - If you used help from such technologies, you need to indicate this
 - If your contribution is small, justify this

H

Dear ChatGPT, how could I start the motivation of my seminar talk on prompt engineering?

- **We do not generally encourage the use of such technologies**
 - Using such technologies entails various risks that may have harmful effects
 - Our intended learning effects may not happen when you use them
- **We encourage learning a responsible use of such technologies**
 - Useful to *support* writing and illustration tasks
 - Not useful to *take over* such tasks, let alone for other tasks like search
 - Ask your advisor, if you are unsure about if/how to use them

Sum up

Take aways

■ Literature research

- Fundamental part of scientific work
- Literature varies in quality and suitability
- Find, read, and organize literature efficiently



■ Oral and written presentation

- Science is storytelling, needs to be understood
- Several best practices for content, structure, and style
- Proper citation is a must
- Practice presenting early



■ For the seminar

- Consider hints in this presentation
- Notice that some are subjective, much is missing
- Develop your own way of presenting



References

▪ Some slides reuse content from:

- **Engels (2010).** Gregor Engels. Einführung in wissenschaftliches Schreiben und Präsentationstechniken. Presentation within the Seminar "Information-Driven Software Engineering". Paderborn, 2010. https://cs.uni-paderborn.de/fileadmin/informatik/fg/dbis/Lehre/ws10_11/PG_IDSE/Dokumente/2010-04-15_Schreiben_Praesentieren.pdf
- **Becker (2012).** Steffen Becker. Scientific Working. Presentation within the Seminar "Model Driven Software Engineering with Eclipse". Paderborn, 2010. www.hni.uni-paderborn.de/fileadmin/Fachgruppen/Softwaretechnik/Lehre/Proseminar_Model_Driven_Software_Engineering/ProSem_MDSD_Guidelines.pdf

▪ Examples are taken from:

- **Ajjour et al. (2017).** Yamen Ajjour, Wei-Fan Chen, Johannes Kiesel, Henning Wachsmuth, and Benno Stein. Unit Segmentation of Argumentative Texts. In Proceedings of the Fourth Workshop on Argument Mining, pages 118–128, 2017. <http://aclweb.org/anthology/W17-5115>
- **Wachsmuth et al. (2017f).** Henning Wachsmuth, Giovanni Da San Martino, Dora Kiesel, and Benno Stein. The Impact of Modeling Overall Argumentation with Tree Kernels. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pages 2369–2379, 2017. <http://aclweb.org/anthology/D17-1252>