HPS

jonathan.decker@uni-goettingen.de

Jonathan Decker

Scientific Writing

Preparing, Structuring and Formatting Scientific Documents

"*Students shouldn’t go into life without the ability to communicate. Your success in life will be determined largely by...*

- your ability to speak,
- your ability to write, and
- the quality of your ideas,
***in that order**.*"

— Prof. Patrick Winston

Goal: Communicate Your Ideas

- Clearly explain your ideas
- Make your work accessible
- Achieve scientific impact

Image source: uni-goettingen.de

Table of contents

Scientists are Professional Writers

- Inform & Communicate
  - Most scientific communication is written
  - Great writing makes your work more visible
- Writing is Thinking
  - Formalize ideas
  - Identify gaps & errors
  - *Get ideas out of your head*
- Recommendation: Write a lot
  - Personal note-taking
  - Focus on content, not form
  - Practice free writing

Types of Writers

Bottom-up

- Easily starts writing
- Understanding while writing
- Iteratively improve text
- Prone to create hard to follow text
- Creative but chaotic

Top-down

- Plan out writing first
- Understand topic first then write
- Structure first then text
- Prone to writer's block
- Organized but inflexible

Task: Where do you see **yourself**?

Tips for Bottom-up Writers

- Reuse pre-written texts for your drafts
- Create a mind-map to sort your ideas
- Prepare a rough structure to guide your writing

Tips for Top-down Writers

- Start writing even without perfect structure
- Just write without fixing to get to the first draft
- Practice free writing to overcome writer’s block

Free Writing

Task: Perform free writing for 5 min

- Open up any writing app
- Think about any topic
  - Ideas about your seminar topic
  - Your day so far
  - Something that is on your mind a lot
- Just start writing down your thoughts
  - Continue to write for **5 min**
  - Do **not** go back to fix typos or rewrite
- Don’t worry you are not expected to share what you wrote

Role of AI

- AI tools can quickly produce eloquent texts
  - Thought process for author is lost
  - Content is often very general, missing depth
  - Statements are given without sources
- Consider AI for
  - Brainstorming
  - Proof reading / Grammar checking
  - Rewriting
- https://chat-ai.academiccloud.de

CS Paper Structure

Research Paper / Thesis / Project Report

- Abstract
1. Introduction
2. Background
3. Related Work
4. Methodology
5. Results
6. Discussion
7. Conclusion
- References

Survey Paper / Seminar Report

- Abstract
1. Introduction
2. Background / Terminology
3. Aspect 1
4. Aspect 2
5. Aspect 3
6. Discussion
7. Conclusion
- References

Abstract

- Summarizes the entire paper
- Concise, informative, and self-contained
- Typically includes:
  - Context and motivation of the research
  - Key methodologies
  - Main findings
  - Conclusions and implications of the research
- Informs and motivate reader
- Avoids too many details

Introduction

- Sets the stage for the research
- Convinces reader of importance of your work
- Provides background and context
- Outlines the research question or hypothesis
- Mentions the significance and potential impact of the work
- List contributions, e.g. what does this work contribute
- Includes outline of the remaining paper's structure
- Goes from general context to specific problem and then to approach

Background

- Provides necessary information for reader to understand the research
- Includes definitions, theoretical foundations,
- Tailored to the audience’s level of expertise
- Not state-of-the-art but common concepts

Related Work

- Reviews relevant literature
- Places own work in context of existing research
- Highlights gaps in previous studies that the current work aims to fill
- Discusses different methodologies and approaches taken by others
- Critical and analytical, not just descriptive
- Includes the majority of references in case paper or thesis

Methodology

- Describes what you did
- First introduces the research questions as *a priori*
- Goes from general RQs to detailed experiment design
- Includes details on experimental design, data collection, and analysis
- Thorough and transparent to allow for reproducibility
- Discusses assumptions, limitations of experiment design

Results

- Describes what you found
- Uses tables, graphs, and figures
  - Discusses each in detail
- Puts results in context of RQs
  - Were the RQs sufficiently answered?
  - Are there problems with the results?
  - What are implications of the results on the RQs?

Discussion

- Interprets the results, explaining their significance and implications
- Connects findings with existing literature
  - May reference both related work and results
- Discusses the limitations of the study
  - Mentions major hurdles and problems encountered
  - Suggests gaps in research due to limitations or assumptions
- Includes theoretical explanations and practical applications of the research

Conclusion

- Summarizes the main findings and their importance
- Restates the research question and how it was addressed
- Highlights the contribution of the study to the field
- Provides an outlook for further research

Paragraph Burger

- Topic sentence
  - Main idea / Claim
- Body sentences
  - Details, definitions, explanations, evidence, reasoning
- Closing sentence
  - Summarize, conclude and link to next paragraph

Sentence Structure - Topic & Stress Positions

- Topic introduces context
- Stress emphasizes content

Task: Identify the Topic & Stress Positions

- GPUs accelerate AI systems in the compute center.
- In the computer center, GPUs accelerate AI systems.
- GPUs, in the computer center, accelerate AI systems.

Sentence Structure - Topic & Stress Positions

- *Topic introduces context*
- **Stress emphasizes content**

Task: Identify the Topic & Stress Positions

- *GPUs* accelerate AI systems **in the compute center**.
- *In the computer center*, GPUs **accelerate AI systems**.
- *GPUs*, in the computer center, **accelerate AI systems**.

Sentence Structure - Text Flow

- Read your text aloud to hear if it *flows*
- Word repetition is fine if it improves clarity
  - Container build workflows typically include the collection of all dependencies into an **image**. This **image** can then be uploaded to a registry and deployed to a given system.
- Ensure pronouns or synonyms do not cost clarity
  - Various systems have become available to deploy AI models for production use cases. **These** have been developed further thanks to increased interest from the industry.

Sentence Structure - Linking

- Link phrases and sentences
- Use words such as **However**, **Furthermore**, **Therefore**, . . .
  - Recent advancements in AI technologies have encouraged the developments of increasingly larger AI models to serve complex use cases. Large AI models have become more common and the overall energy consumption of the AI sector has increased. More energy efficient methods are direly needed.

Task: Reformulate to improve text flow

Sentence Structure - Linking

- Link phrases and sentences
- Use words such as **However**, **Furthermore**, **Therefore**, . . .
  - Recent advancements in AI technologies have encouraged the developments of increasingly larger AI models to serve complex use cases. Large AI models have become more common and the overall energy consumption of the AI sector has increased. More energy efficient methods are direly needed.

Task: Reformulate to improve text flow

- Consider this
  - Recent advancements in AI technologies have encouraged the developments of increasingly larger AI models to serve complex use cases. **However,** as these large AI models have become more common, the overall energy consumption of the AI sector has increased and has **therefore** created a dire need for more energy efficient methods.

Drafting vs Revising

- Drafting
  - Write initial text
  - Fill out outline
- Revising
  - Iterate on text
  - Improve content, structure, language, style

Do **NOT** mix these phases!

Revising

Content $\Rightarrow$ Structure $\Rightarrow$ Language

- Content
  - Is the main idea clear?
  - Are all details relevant?
- Structure
  - Is the text structure/storyline easy to follow?
  - Does the structure persist through paragraphs and sentences?
- Language
  - Is the language clear and precise?
  - Are grammar and spelling correct?

Feedback

- Ask for feedback
  - More than one for better quality improvements
- Be precise in your request
  - Content
  - Structure
  - Topic sentences
  - Text flow, language, style

"*Perfection is achieved, not when there is nothing more to add,*

but when there is nothing left to take away."

— Saint-Exupéry, *Airman's Odyssey*

Goal of Scientific Language

- Readability
- Understandability
- Clarity of thought
- Precision and accuracy

Active vs Passive Voice

- Prefer active voice
  - More appealing to the reader

- The GPU processed the request.
- The request was processed by the GPU.
- The researchers discovered a security vulnerability.
- A security vulnerability was discovered by the researchers.

Tense

- Use present tense for eternal truths
  - The algorithm has a complexity of $O(n^2)$.
- Use past tense for ideas and outcomes
  - The concept was verified by the experiment.
- Avoid
  - The limitations will be discussed in the following section.

Choice of Words

- Use short, direct words
- Choose words that are unambiguous
- Prefer formal to colloquial language
- Avoid words such as
  - *Use, make, give* (imprecise)
  - *Like, someone* (colloquial)
  - `Obvious` (Nothing is obvious in science)

Images, Tables and Code

- Use `figure`, `table`, `listing` environments
- Set a `label` and use `ref` or `Cref` to reference in text
  - Always use this for discussing them
- Use high-resolution images
- Ensure tables are readable
  - Use alternating row colors, highlighting
- Use code highlighting
  - Only show relevant code snippets, not entire files

Citation

- Present sources for your claims
- Use `cite` command and manage sources as biblatex file
- Sources should directly follow respective claim
  - The performance of the system under study was found to be close to native performance leading to the conclusion that it is suitable for HPC usage. **[REF]**
  - The performance of the system under study was found to be close to native performance **[REF]** leading to the conclusion that it is suitable for HPC usage.
- Sources and authors can also be directly referenced
  - Author et al. **[REF1]** found a new method for compressing textual data, which complements the protocol presented in **[REF2]** allowing for more stable and efficient data transfer.

I vs We

- Use *we* to distinguish between your work and existing work
- Prefer *we* over *the authors*
- Only use *I* for author's opinion
- Commonly use *we* even for solo author

Sexism

- Avoid unnecessary gender specification
  - Use *they* instead

- Avoid
  - A user may see an error message when **he** makes a mistake
- Use
  - A user may see an error message when **they** make a mistake
  - Users may see an error message when **they** make a mistake

Closing Remarks

- You are now empowered to write high-quality scientific documents
- Hone your craft
  - Practice and get feedback
  - There is much more to it than covered today
- Review these slides as needed
- Check out Zobel, *Writing for Computer Science*

References

Saint-Exupéry, Antoine de. *Airman's Odyssey*. Harcourt Brace Jovanovich, 1984. 460 pp. ISBN: 978-0-15-603733-4. Google Books: nIOZdLHReUMC.

Zobel, Justin. *Writing for Computer Science*. London: Springer, 2014. ISBN: 978-1-4471-6638-2 978-1-4471-6639-9. DOI: 10.1007/978-1-4471-6639-9. URL: https://link.springer.com/10.1007/978-1-4471-6639-9 (visited on 04/09/2024).