Uni Scientific Writing Notes

Notes for the Anleitung zum wissenschaftlichen Arbeiten (scientific writing) course at HdM Stuttgart

Felicitas Pojtinger

2021-11-19

Introduction

Contributing

These study materials are heavily based on professor Charzinski's "Anleitung zum wissenschaftlichen Arbeiten" lecture at HdM Stuttgart.

Found an error or have a suggestion? Please open an issue on GitHub (github.com/pojntfx/uni-sciwriting-notes):

Figure 1: QR code to source repository

If you like the study materials, a GitHub star is always appreciated. :)

License

Figure 2: AGPL-3.0 license badge

SPDX-License-Identifier: AGPL-3.0

Organization

Organization

- Primarily based on the inverted classroom principle
- Sent files should not contain metadata on person-specific info (make pseudonymous)
- Paper must be sent in by 2022-01-09
- Notes must be sent in by 2022-02-27
- Paper may be in German or English

Overview

Overview

1. What is the scientific method?
2. Formulating scientific questions
3. Designing experiments
4. Analyzing experiments
5. Planing scientific papers
6. Researching topics and staying up-to-date
    - 6.1 Finding papers → Sci-Hub
    - 6.2 Analyzing papers
    - 6.3 Referencing papers
7. Writing a scientific paper
8. LaTeX

What is the Scientific Method?

Writing Style

- Structure should not follow the timeline research, but the semantic structure of the discovery
- No rhetorical questions
- No judgmental formulations
- Sentences should be able to stand on their own; reference people and things by their name, not implicit references
- Do not use the present tense when referring to past events, even if it is popular in journalism
- Do not use metaphors which are highly imprecise, even if they are common among technical people
- “I” should not be used in texts
- Summaries should be about the effect of the research on the subject, not the author’s view on the subject
- The “motivation” at the start of the paper should not be the personal factors, but prior pointers

Typical Criteria

- Complexity of the theme
- Amount of personal research
- Quality of the content
- Depth of research
- Selection of sources
- Implementation of prior knowledge
- Structure of the paper
- Visual style (used fonts, formatting etc.)
- Quote style (standardized quotes)

Formulating Scientific Questions

Logic and Conclusion

- Argumentation
- Logical conclusions
- Proofs (i.e. mathematical proofs)
- Experiments and their design, execution and analysis
- New analysis is always based on existing knowledge
- There are different levels of formalism: Argumentation, validation, predicate-based proofs
- Referencing ideas can be done in an “informal” way (whitepapers etc.), but they must not be the base of any claims!

The Purpose of Writing

- Communication is the primary purpose of scientific writing
- But scientific writing is also a means of analysis
  - Formulating thesis helps to grasp the connections between arguments
  - Clear formulation makes it much harder to avoid critical questions
  - Gaps in analysis and open questions become obvious and lead to new research opportunities
  - Writing leads to a deeper internal understanding
- Even if scientific writing is limited to Uni, research methods are always required

The Scientific Thought Model

1. Outlook
2. Own research
    2.1 Discussion
    2.2 Proofs, research, experiments, studies
    2.3 Hypothesis, underlying idea
3. Summary of the current state of research/technology (“related work”)
4. Sources (own and external)

Quality Assurance

- New ideas should be able to be based on existing works
- Peer reviews try to check the quality of scientific works and ensures that existing work can serve as a solid base
- Own share of own work must be made obvious

Scientific Questions

- Formulation a concrete question is required in order to reduce the scope of topics
- The question doesn't have to be clear in the beginning of the writing process, but must be at the end
- The focus is always on the question, not the means: "Does the raft algorithm work reliably?" for example would not include/require an implementation of the raft algorithm, so always make the implementation a requirement of the question!
- The scientific question is not the title of the paper
- Just like the goals of the research need to be clearly defined, the "non-goals" need to be too!

Experiments

- Gathering of data
- Hypothesis
  - Creating the hypothesis
  - Designing the experiment
  - Executing the experiment
  - Testing the hypothesis with the result
  - Further, refined hypothesis ideas
- The hypothesis is often “my idea/solution/architecture works”
  - Experiments support the hypothesis
  - Paper then describes the current technological state, experiments and results
- All dependencies and state required to reproduce the experiment must be notes

Methods of Experiments

1. Design
   - 1.1 Matches the scientific question
   - 1.2 Creativity is required
   - 1.3 Viability in time, budget and with available technology
2. Planning
   - 2.1 Prevention of side effects
   - 2.2 No convenience samples
   - 2.3 No unethical experiments
3. Execution
   - 3.1 With proper process
   - 3.2 Proper documentation, including all unexpected incidents
4. Analysis
   - 4.1 Objective analysis
   - 4.2 No suppression of “unwanted” results
5. Interpretation
   - 5.1 Objective interpretation
   - 5.2 Usage of statistics: Is the result even statistically relevant?

Hypothesis

- Verification using proofs
- Validation based on empirical data
- Multiple supporting hypothesis can build a theory

Experiment Design

- Experiments should produce a result
- Testing in a specific set of parameters
    - Searching for optimal parameter combinations
    - Checking for valid sets
- Sensitivity analysis
    - Checking the hypothesis with parameters
    - Checking if parameters influence results
- Hypothesis tests: Statistically testing the results of experiments

Analysis

- Be neutral
- Always ask question about results, even if they are positive
- Search for additional sources
- Comment on unknown factors, don’t hide them - they are means of finding the next topic to research on!

Working with “Outliers”

- Don’t remove or ignore them
- Test if they are relevant: Do more research - are they statistically relevant?
- If they are not relevant: Classify and document

Comparisons

- The new is not automatically better
  - Comparison with a baseline reference is required
  - Detailed description of the reference system used is required
- Define the used dimensions for the comparison
  - Differences often occur in different dimensions
  - Elaborate why dimensions are being used
- Fair basis: i.e. not using an under powered server
- Also point out that the tool might perform worse under different dimensions (i.e. memory constrained systems)
- Comparison by
  - Comparison the reference solution and the new solution
  - Comparison of the new solution with existing literature

Planning Scientific Papers

Exposé

- Might be required
- Significant research requires planning
  - Assessment of feasibility
  - Usage of time slots
  - Focus on the most important goals or topics
- Short description of the planned research
  - Which problem is the basis of the planned research?
  - Prior, existing research and open questions
  - The main scientific question: Which question is the research going to answer?
  - Goal of the research
  - What theories is the research based on
  - Methods
  - Materials
  - Structure
  - How much time are the individual slots expected to take

Structure

- Based on argumentation or path of discovery
- Balanced
- Not too much hierarchy
- Minimum length of the chapters and sections
- Total average ~50-60 pages
- Per chapter ~3-10 pages
- Typical:
  - Abstract (no section number, in both English and German)
  - Introduction (including overview)
  - Related work
  - Main investigation (multiple sections)
  - Results
  - Summary and conclusions
  - References

Basic Procedure

- Clarification
  - Which questions should be answered?
  - What are the non-goals?
- Creating the project plan
- Getting up to date from a technical perspective
  - Which state is the research based on?
  - Search and analyzing papers
- Own works
  - Sometimes simply structuring the comparison
  - Normally: Experiments!
    - Definition
    - Execution
    - Analysis
- Selecting tools (BibTeX, LaTeX)
- Sketching
  - Creating a structure (i.e. mind maps)
  - Taking note of keywords and images

Planning

- Every project needs planning
- Sketched planning needs to happen early in the project
  - Literature studies are often underestimated
  - Own works
  - Writing (min. four weeks before time is over!)
- More fine-tuned research with more knowledge
- Current state of research must be checked during own research
- Immediate active countermeasures are required
  - Plan must be changed
  - Asses severity of changes

Planning the Main Section

- Structure is central
- Amount of pages per section is required
- Contents per section must be planned: Keywords, sources, images
- Writing takes time; start writing meta before actually starting to write

Planning the Paper for this Module

- Formulating the scientific question
- Creating a structure
- Searching and analyzing literature
- Refining the structure (two layers) including page numbers
- Selecting graphics (with sources)
- Writing
- Checking
- Submitting the paper

Researching Topics and Staying Up-to-Date

Sources

- Web
- Wiki
- Google
- Libraries: Books and articles
- Journals and conferences: Finding journals, special issues, searching for articles
- Use catalogs

Research

1. Starting with research
    1.1 Internet (Wikipedia, Library Genesis, Sci-Hub, Scholar, CiteSeerX, arXiv, ResearchGate)
    1.2 Libraries
    1.3 Journals
2. Skimming the first articles
3. Doing more research on interesting literature
    3.1 Finding the primary source
    3.2 Finding papers which have been cited often
4. Finding related authors and researching their latest papers

Skimming Papers

- Don’t start by reading the paper from start to finish
- What did the authors do?
  - New understanding of existing systems
  - New solutions for the issue
  - Explanation of a new research question (with or without a solution)
  - Reviewing existing solutions or ideas
- What is the result of the paper?
- Don’t check only the abstract - skim for keywords too!
- Analyze included graphics
- Checking the title
- Checking figure descriptions
- Don’t check all math unless necessary (which it mostly isn’t)

Reading Papers

- Maintaining a critical view: Many papers over-promise and under-deliver
- Still: Skim the paper first
- Extracting main expressions
- Only read subjects in detail which are interesting for the research topic

Critical Reading

- Be aware of deceptive terminology
- Don’t use “common sense”
- Note implicit and explicit assumptions, approximations: Are they warranted?

Documenting the Reading Process

- Excerpts
  - In sections or with paraphrasing
  - What is the topic? What is being published on it?
- Creating a summary
- Paraphrasing
- Adding comments
- Visualizations: Mind maps, concepts maps or logical formulas

Critiquing Papers

- **Scientifc Standards**: Scientifc questions, methods, literature and other sources
- **Ideology**: Author's bias and own ideology
- **Context**: In reality, norms-values-means
- **Argumentation**: Facts, experiences, norms-values-means, authority

Re-Definitions

- As it is known, ... → I think, ...
- It is obvious ... → I think, ...
- Maybe one could argue, that → I'm not sure what to think
- There is consensus → Some people think
- For obvious reasons → I have no proof
- There is no doubt → I am sure
- It is likely → I have no proof and don't have the time to check
- It is not necessary to take a closer look → I do not want to take a closer look

TODO: Add section on referencing other works

Citation

Bibliography

- Contains all read works
  - Used sources
  - Current state of research
  - Support for argmentations
  - Base for comparisons
- In .bib file
- Can be used for multiple papers

References

- What
  - Bibliographic references
  - Own annotations
    - Excerpts
    - Comments
    - Keywords
    - Opinions
  - Relation to other references
- How
  - Findable
  - Extensible
  - Linkable (in both directions)
  - Useful in bibliography
- Where: List or database

Using References

- Before reading: Taking note of bibliographic data
- While reading: Excerpts, annotation and links between references
- While writing
  - Citing directly (including page number)
  - Automatic creation of references allows automatic import into word processing

Purpose of Citations

- Showing which ideas came from whom and which publication
- Often a requirement due to copyright restrictions (attribution)
- Shows that relevant literature was consulted
- Creates a chain of trust based on trusted sources
- Can allow checking the novelty of a work (what is new, what is referenced?)

Evaluation of Source Quality

- Sources must be verifiable and trusted, so peer-reviewed publications are the best basis
- Wikipedia is a good entry point due to high quality and depth, but citing original sources is often the better choice
- Blogs and popular science publications are useful for citing opinions and events, but should not be used to give an overview of the current state of technology
- Whitepapers should only be used for research specific to the publisher's technology

Primary and Secondary Sources

- **Primary source**: The first publication of an idea by its inventor
- **Secondary source**: Recitation or analysis of an idea
- Reading primary sources allows checkng if secondary sources have maybe misrepresented studies or used out-of-context quotes

Languages

- In non-English publications, using both sources in the native language and English is acceptable
- In English publications, non-English publications should only be cited if no other sources could be found

Quotes

- Short quotes must always be marked using ”
- Longer quotes should be in an own paragraph and have a different style
- Require exact source, including page number
- Have to be 1:1 representation
  - Including punctuation and writing style
  - Mark exclusions and own additions with []
- Should be from primary source
- Quotes are not typically used in informatics papers, except for loosening up the structure or to introduce chapters; in social sciences, they are used more frequently, as they can be a subject to analysis (i.e. in literature analysis)

Reference Style

- Reference should link to an information source
- Using a reference means that the statement of the work inherit the quality properties of the reference, as it is based on it
- Source reference must contain the relevant data to uniquely identify a source
- Different styles are available
    - Chicago style (EU method)
    - Harvard style
    - Legal style (footnotes)

Literature List

- Contains properties for each source
  - Name of authors
  - Title of publication
  - Name, volume, year, edition and page number
  - Publisher, location, date of publication
- BibTeX can generate literature lists for most styles

Writing a Scientific Paper

Diligence

- **Formally**: A general diligence guideline is mandatory in scientific writing
- **Practically**: Spelling, syntax and layout issues
  - Readers get more critical and find more errors
  - Worse marks

Types of Papers

- Survey papers: Overview of a subject
- Scientific protocol: Documentation and interpretation of a experiments
- Research paper
- Thesis (BSc, MSc, PhD)
- Certificate

Types of Documents

- Protocol
- Whitepaper
- Specification
- Offers
- Presentation
- Advertisement
- Functional descriptions
- Manuals
- Press releases
- Patents
- News articles
- Blogposts

Choice of Language

- German
  - Easier as a native speaker
  - Many proofreaders
- English
  - Important for all relevant documents
  - More readers

Tips on Style

- Writing is an exercise
- Structure is a hard requirement
- No suspense
- Use simple and clear styles
- Rather try to impress with content than with complex sentences
- Clearness is important, because it is required in the job, makes reading and writing easier and is polite to the reader
- Use foreign words with care
- Keep the audience in mind

How to Deal with Writer's Block?

- Deadlocks?
- Just start writing *anything*
- Work on structure instead
- Creating a mind map
- Don't trash drafts, refactor them instead

Title Style

- Scientific question != title (title should not be a question)
- Don’t be too general or to precise
- Must contain the main theme

Writing the Abstract

- Short summary of the subject's field and the solution
- Must include the result (should *not* build suspense)
- Should not contain short abbreviations, references, formulas and sentences like "In this paper ..."
- ~250 words
- Should be in English and German
- Current and future relevance of the subject
- Contexts in which the subject has been analyzed

Writing the Overview

- Last section of the introduction (first section)
- Shows the relations and dependencies between the sections
- Should not just reiterate the table of contents

Writing the “Related Work” Section

- Overview of prior and similar work
- Creates the base/foundation of knowledge
- Who researched what?
- Where has the result been published?
- Which problems have not been solved in prior work?
- In which context does the work stand to related work?
- Should exist before starting to write!

Writing the Outlook Section

- Was has been researched?
- What could be improved?
- Short summary of the results
- Meaning of the results
- Which problems could not be solved?
- Judgement of the implementation
- Learned experiences
- New contexts to other research topics

Writing the Acknowledgements Section

- Not a formal requirement, but a social requirement
- Especially relevant if access to internal info or external unis has been provided
- Can be used for other sources or ideas that can’t be formally sourced

Scientific Grammar and Style

- Third person
- Simple past
- Never reference self or other groups/people
- Short sentences and words
- Don’t repeat formulations but do repeat words instead of using synonyms (server, node, VPS etc. - choose one!)
- Use SI units
- Use significant figures
- Use consistent list style, examples, unit structure (Mbit/s instead Mbps, Mbit/sec etc.)
- The first sentence of each paragraph should be the paragraph’s introduction
- Define acronyms
- Simple and reserved
- Should leave no space for interpretation

Embedding Figures

- Always numerated
- Must have an alt text
- Referenced in text by figure number
- Text must *never* flow to the left or right of the figure
- Source can be in alt text (i.e. “(...) using data from [3]”)

Infographics

- Diagrams (ER, UML etc.)
- Code or pseudocode
- Sequential numbering of tables and figures
- Tables must have their titles on their top
- Figures must have their titles below
- Use consistent font sizes for descriptions

Common Mistakes

- Spelling
- Style/Syntax
- It's a project description, not a scientific paper
- Separating defects
- Broken references
- Missing alt texts
- Text in description of graphics too small
- Inconsistent terminology

Last Checks

- Spelling (i.e. LTex for LaTeX)
- Check if all diagrams and graphics
- Check for broken References
- Empty pages
- Do all graphics work in black/white?
- Have all acronyms been introduced before they have been used?
- Always re-check everything after fixing