Scientific Writing in a Nutshell

Summer 2024

Dr. Schapranow

Design IT.
Create Knowledge.

Key Facts

Submission no later than <u>Sun Aug 18, 2024</u> containing:

- Contributions, prototypes, models, source code as a zip archive,

- Brief instruction how to make it working, and

- Your final report as PDF version + source, e.g. Latex/Word, figures, etc.

Final report

- Format: IEEE two-column conference, A4 templates available at https://www.ieee.org/conferences/publishing/templates.html

- Length max. 4 (TaC4DH) or 6 (AI4DH) full pages (excl. appendices, references, figures) respectively

Accessing the templates

Microsoft Word
- US letter (DOC, 30 KB) **Updated Jan 2019**
- A4 (DOC, 30 KB) **Updated Jan 2019**

LaTeX Template Instructions (PDF, 63 KB) [Be sure to use the template's **conference** mode.]
- Template (ZIP, 700 KB) **Updated October 2019**
- LaTeX Bibliography Files (ZIP, 309 KB)

Overleaf
- When working in Overleaf, the template is available at https://www.overleaf.com/gallery/tagged/ieee-official
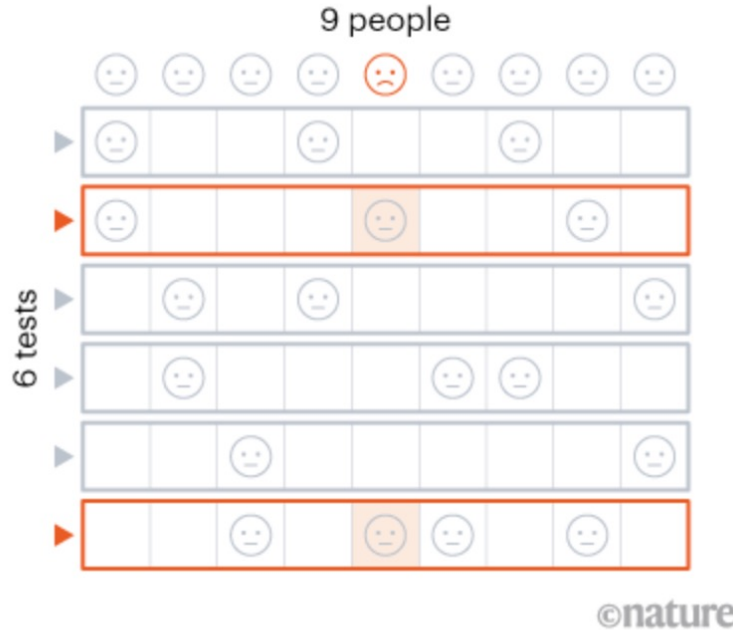
Can Emojis Make Science?

Emojis May Help Science!

Method 4
This method uses only one round of testing. Samples are distributed into a matrix of overlapping groups.

9 people

6 tests

©nature

Sources: Dorfman, R. *Ann. Math. Statist.* 14, 436-440 (1943); Mutesa, L. *et al.* Preprint at https://arxiv.org/abs/2004.14934 (2020).; Manoj Gopalkrishnan
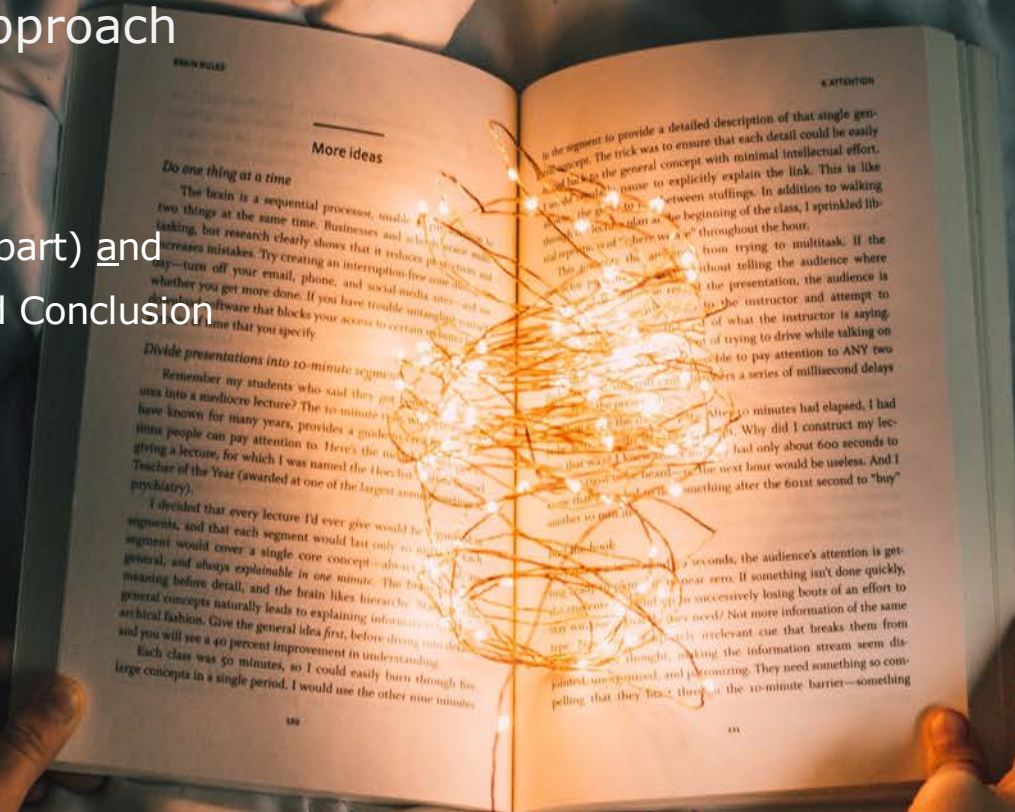
Structure
The Fairytale Approach

1. Introduction
2. Main part
3. Conclusion

Photo by Nong

Structure
The IMRAD Approach

1. Introduction
2. Methods
3. Results (Main part) and
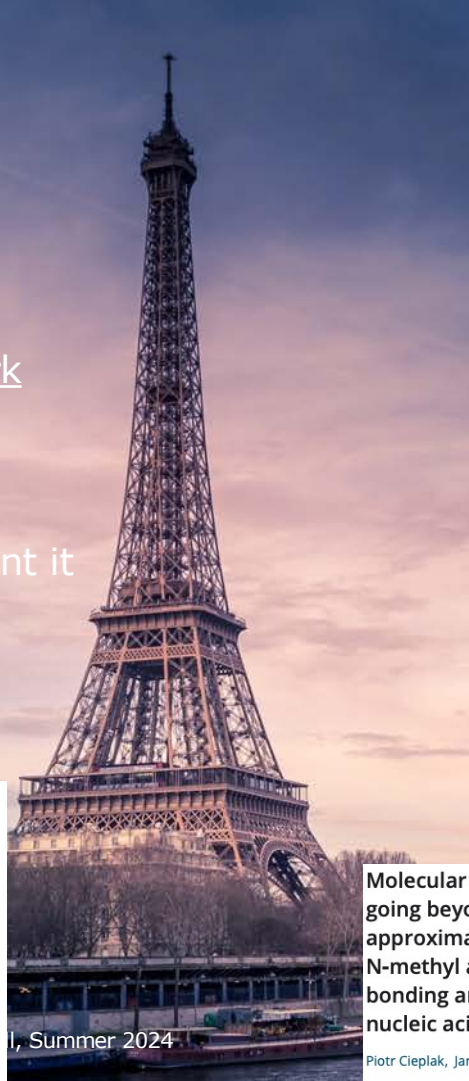4. Discussion and Conclusion

Structure
The Engineering Papers Approach

1. Introduction
2. Related Work
3. Methods
4. Contribution(s)
5. Evaluation
6. Discussion
7. Conclusion and Outlook

Photo by Nong

Title

- Aim: Make reader curious about your work

- Catchy, but precise
- Not too long as you want it to be indexed

mean quantile citation rank (%)

mean quantile title length (%)

N. Engl. J. Med.

Nature

Science

Cell

Lancet

J. Am. Chem. Soc.

Blood

PLoS ONE

Proc. Natl Acad. Sci. USA

Sci. Rep.

The nucleotide sequence of a 3.2 kb segment of mitochondrial maxicircle DNA from *Crithidia fasciculata* containing the gene for cytochrome oxidase subunit III, the N-terminal part of the apocytochrome *b* gene and a possible frameshift gene; further evidence for the use of unusual initiator triplets in trypanosome mitochondria

Paul Sloof ✉, Janny van den Burg, Arthur Voogd, Rob Benne

Molecular mechanical models for organic and biological systems going beyond the atom centered two body additive approximation: aqueous solution free energies of methanol and N-methyl acetamide, nucleic acid base, and amide hydrogen bonding and chloroform/water partition coefficients of the nucleic acid bases

Piotr Cieplak, James Caldwell, Peter Kollman ✉
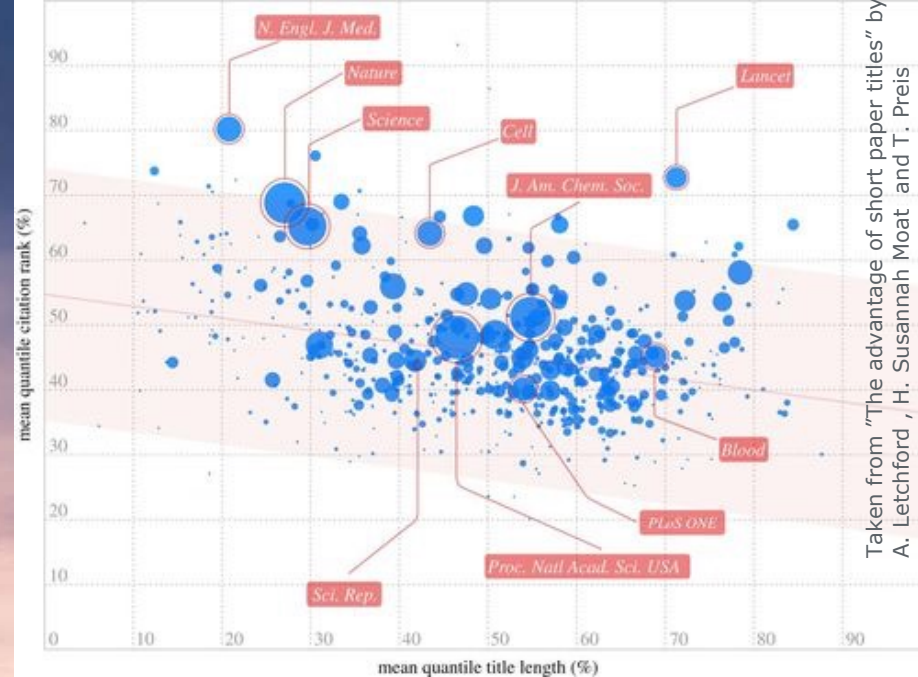
Abstract

- <u>Aim: Make the reader understand the scope of your contribution</u>

- Answer the following four questions:
1. What is the state of the art?
2. What are the limitations of the current state?
3. What is your contribution?
4. How does you contribution improves the state of the art?
- Review your abstract at the end of your work to check whether it is still aligned

Photo by Free Nomad

1. Introduction

- <u>Aim: Brief intro into the topic not an academic lecture</u>

- Guide the reader to the topic in brief

- Name some well-known work

- You might want to place an important figure or graph on the first page to make your work look unique

- The last paragraph shares details about the structure of your document, e.g. "<u>The remainder of the work is structured as follows: Sect. 1...</u>"

2. Related Work

- <u>Aim: Show you know other's work and can set your contribution in the context</u>

- Perform literature review, share results, e.g. IEEE, PubMed, ACM, SpringerLink, Google Scholar, Elsevier Scopus, etc.
- Name selected approaches and stress how your work differs, extends, matches, etc.
- You might want to create a structured tabular overview, e.g. selected features, approaches

Table 1: Tools surveyed. Abbreviations: GP–General Purpose, BSP–Biomedical Software Platform. Legend: ●–Fully supported; ◐–Partially supported; ○–Not supported.

Category	Tool	Preparation			DS Selection			Pred. Handling			M. Generation			Eval. & Validation			Interpretation			Presentation			NF		
		F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	NF1	NF2	NF3
GP	Statistical Packages [1, 30, 40]	○	○	○	○	○	○	●	●	●	●	○	○	●	●	○	○	○	○	○	○	○	○	○	●
	ML Toolkits [14, 21, 28, 39, 41]	○	○	○	◐	○	○	●	●	●	●	○	○	●	●	○	●	○	○	○	○	○	○	◐	●
	Scientific Computing [17, 25, 51]	○	○	○	○	○	○	●	●	●	○	◐	○	●	●	○	●	○	○	○	○	○	○	◐	●
	ML Suites [16, 22, 25, 31]	○	○	○	○	○	○	●	●	●	●	●	●	●	●	○	●	○	○	○	○	○	○	◐	●
BSP	tranSMART [49]	◐	○	○	●	●	●	○	○	○	○	○	○	●	○	○	●	○	○	●	●	●	○	○	○
	ATLAS [19]	○	◐	●	●	●	○	○	○	○	◐	○	○	●	○	◐	●	◐	○	◐	●	●	◐	◐	○
	MLBCD [36]	○	○	○	◐	○	○	●	○	◐	●	○	○	●	○	○	●	○	◐	○	●	○	◐	○	○
	ExplICU [11]	●	○	○	●	●	○	○	○	○	○	○	○	●	○	◐	●	○	◐	◐	●	○	◐	○	○
	P1P R package [46]	◐	○	◐	●	○	○	○	○	○	○	○	○	○	○	◐	●	○	○	○	●	○	○	○	○
	MORPHER	●	○	◐	◐	●	○	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●

Photo by Gülfer Ergin

3. Methods

- Aim: Outline how you worked methodically to get your findings

- What particular methods have you used, e.g. design thinking, software engineering, machine learning methods, quantitative measurements, etc.
- Why did you select them, e.g. referenced by others, make your work comparable to other's work, etc.

Photo by Hunter Haley

4. Contribution(s)

- <u>Aim: Share most relevant details of your work</u>

- Describe what you achieved, i.e. results and deliverables
- Highlight your findings, what is unique, what are you proud of, what do you want to share with others
- Share detailed insights, prevent just general sentences
- Clearly distinguish b/w your own work and what you have built upon
- Prevent story-telling, e.g. ~~"First, I did this, but it fails. Then, I tried this. At the end…"~~

Photo by Markus Spiske

4a. Benchmarks and Results

- Aim: Compare your work and support reproducibility

- Describe your specific setup for obtaining results, e.g. computing infrastructure, measurement devices, precision of clocks or measurement devices, what was measured, how often, etc.

- What metric(s) did you measure and for what reason

- Which parameters were kept fixed / varied – hint: keep most parameters fixed and change one parameter at a time

- What results did you acquire

- Please do not place interpretations or assessments here → better move them to your evaluation

5. Evaluation

- <u>Aim: Help the reader to interpret your findings/observations; focus on your observations!</u>

- "I tried to produce identical eggs, <u>but my results differ in size and color</u>."
- "We measured a mean response time of 110 ms with a standard deviation of 15%, <u>which is 10% faster than comparable approaches</u>."
- "We would have expected to double the throughput, <u>but stuck at a factor of 1.5x</u>."

Photo by NordWood Themes
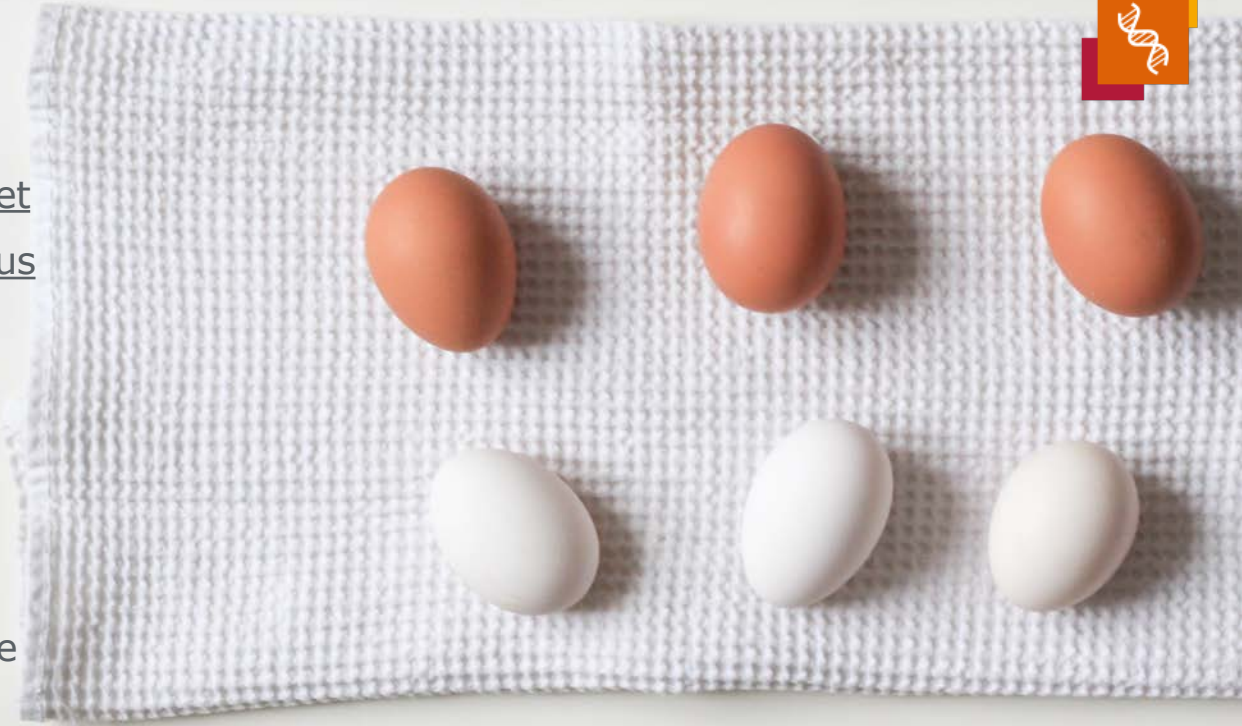
- <u>Aim: Discuss your findings, what did the evaluation show, how to proceed with it</u>

- Did you results meet your expectations, yes/no why?
- You might want to explain why you meet/lack behind certain assumption
- What could be possible explanations for your observations/measurements
- What next steps/changes need to be conducted
  to understand the source for your results
- Do you need to change assumptions, data, parameters, etc.

Photo by Volodymyr Hryshchenko

7. Conclusion & Outlook

- Conclusion (<u>Aim: Sum up your findings / main contribution</u>)
  - ☐ Help to recall your findings / contributions in a nutshell
  - ☐ Stress why and how your contributions extend the current
    state of the art; why is your work important

- Outlook (<u>Aim: Guide the reader through your planned next steps</u>)
  - ☐ Do you plan certain next steps, experiments
  - ☐ How do you plan to proceed your research
  - ☐ How will you address issues identified during your evaluation

Photo by Jon Eric Marababol

Bibliography

- Recommendation: Use a bibliography management tool to make your life easier
- Always check (even if you use existing bib snippets for publications):
  □ Check for consistency, e.g. upper- vs. lowercase of titles and journal names (use curly brackets in bib files to lock format "{MySQL}")
  □ More than three authors? → First author's name + "et al." (can be achieved by "Schapranow and others" in bib files)
  □ Online references? → Please add URL and "last checked" date – Hint: Check all online references prior to submission at the same date

Photo by j zamora

Style of Writing

- Be precise and formal, e.g. "I or we did X and Y to achieve Z."
- Underline statements by adding appropriate references at the end of the sentence
- Write about facts and not about opinions or your individual challenges – the discussion section is the right place to share your personal opinion
- Consider splitting complex sentences in multiple parts
- Please prevent "already mentioned, as state before, again…"

- Use commas when appropriate, e.g.
  - With conjunctions,
  - With most "w-"pronouns,
  - Must not be placed before _that_

Photo by Taneli Lahtinen

The Perfect Length of the Paper

"I have studied this manuscript very carefully with lemon juice and X-rays and have not detected a single flaw in either design or writing style. I suggest it be published without revision. Clearly it is the most concise manuscript I have ever seen-yet it contains sufficient detail to allow other investigators to replicate Dr. Upper's failure. In comparison with the other manuscripts I get from you containing all that complicated detail, this one was a pleasure to examine. Surely we can find a place for this paper in the Journal-perhaps on the edge of a blank page."

THE UNSUCCESSFUL SELF-TREATMENT OF A CASE OF "WRITER'S BLOCK"[1]

DENNIS UPPER

VETERANS ADMINISTRATION HOSPITAL, BROCKTON, MASSACHUSETTS

———————

REFERENCES

———————

https://www.realclearscience.com/blog/2014/01/shortest_science_papers.html

Formal Aspects

- Check spelling and grammar repeatedly
- If a new line ends with a single word, consider rephrasing the sentence
- Place multiple references together at the end of the line to support readability, e.g. "[1,7-10,15]".
- Check if anything – figures, tables or text – reaches into the paper margin
- Place figures either at top / bottom of a page, prior to referring to them
- Refer to all figures, tables, equations, listings, etc. within the text
- Add an appropriate caption to your figures, tables, equations, listings, etc. to summarize them for the reader
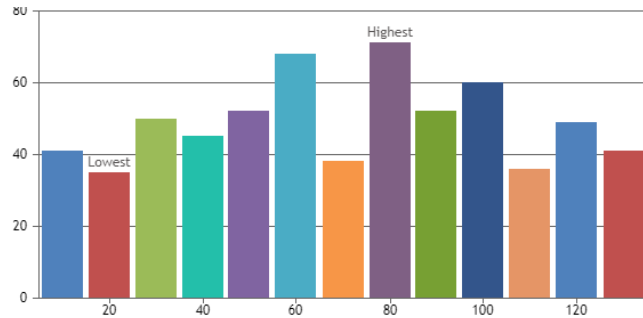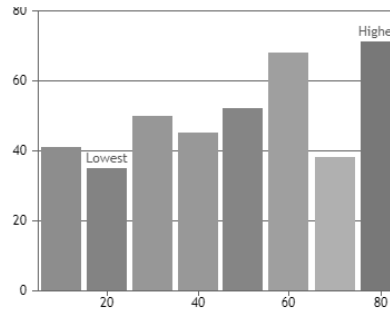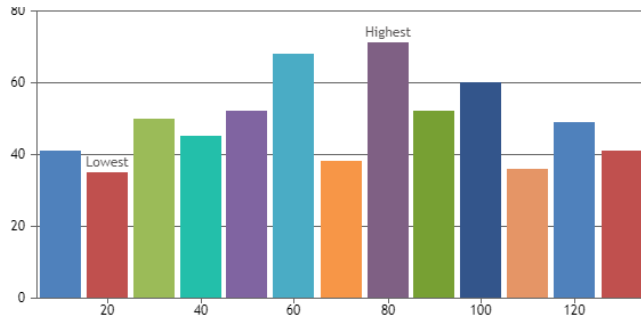
Von GFlund auf Wikimedia.org

Figures, Charts, Listings, …

- "A picture is worth a thousand words"
- Creates an long-lasting association in the reader's mind
- A strong form of communicating findings

Figures, Charts, Listings, …
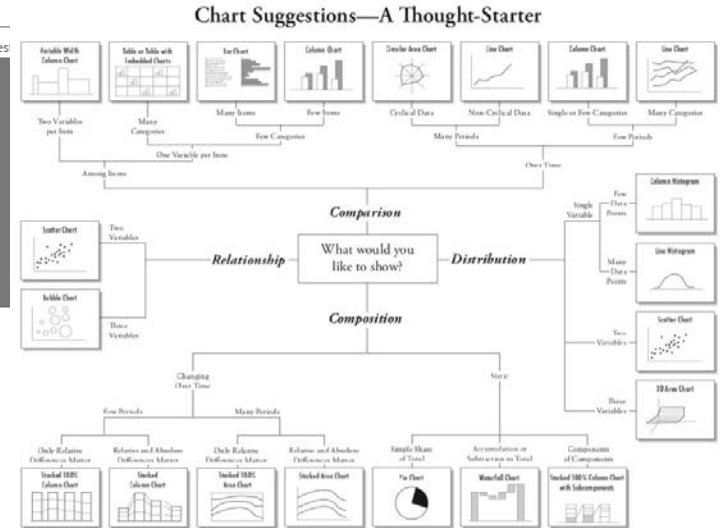

- "A picture is worth a thousand words"

- Creates an long-lasting association in the reader's mind

- A strong form of communicating findings

- Use visual hatchings to make charts readable even in grayscale printing

- Text in figures should be the same size as the remainder of your document

- Resolution should be adequate for printing, i.e. at least 300dpi

LaTeX: Tips and Tricks

- Dos
  - Introduce \label{fig:abc} for anything you want to reference, e.g. section, figures, charts, listings, etc.
  - Use \autoref or \cref instead of <labelname>~\ref
  - For references, use a .bib file and an appropriate bib editor, e.g. BibDesk, Mendeley
  - Check line breaks in latex source because they have a special meaning, e.g. indent

- Don'ts
  - Prevent footnotes, e.g. for URLs; instead add a proper reference
  - Prevent manual formatting, e.g. bold or italics, because it is managed by .sty/.cls
  - Manual line breaks (\\) or new pages (\newpage)
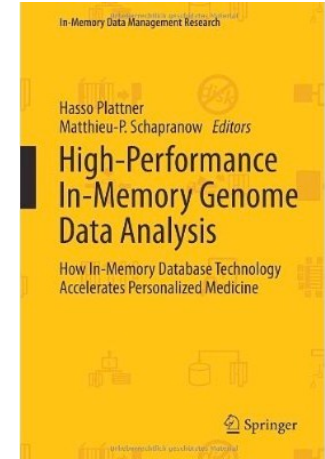
Photo by Scott Rodgerson

Happy writing!

Dr. Matthieu-P. Schapranow
Group Leader & Scientific Manager Digital Health Innovations

Hasso Plattner Institute
Digital Health Cluster
Rudolf-Breitscheid-Str. 187
14482 Potsdam, Germany

we.analyzegenomes.com
@AnalyzeGenomes

In-Memory Data Management Research

Hasso Plattner
Matthieu-P. Schapranow *Editors*

High-Performance In-Memory Genome Data Analysis

How In-Memory Database Technology Accelerates Personalized Medicine

Springer