TECHNISCHE UNIVERSITÄT DRESDEN

ACADEMIC SKILLS IN COMPUTER SCIENCE

Lecture 13: Ethics and Academic Conduct

Markus Krötzsch

Knowledge-Based Systems

TU Dresden, 8th July 2019

Review: Oral presentations

Handling questions

Often people ask questions after talks:

- to learn something
- to test you (esp. in thesis defence)
- to make a public statement (self-promotion, personal views, . . .)

All questions are good for you!

General hints:

- Hear the whole question: do not interrupt (unless asker rambles on forever)
- Answer frankly, but don't forget to contrast with strong points when criticised on a weak point
- Admit if you don't know something (no point hiding)
- Suggest offline discussions when questions get out of hand (but not before!)
- Look at co-authors in your audience – maybe they are eager to answer when you are in a fix?

Handling conflicts

Questions can (rarely) be aggressive attempts to discredit your work

- Really critical or aggressive questions are rare – do not misinterpret an honest query as an attack towards you!
- Don't loose your cool (easier said than done)
- Your advantage: you are the default person everybody listens to
- Your goal in aggressive discussions is usually to get the audience on your side, not the person who asked!

Marginalise and cut off weak attackers: for example "I have to disagree with you on this point – our experimental approach is certainly valid [ideally also briefly include technical reasons] – but this topic is a bit technical for giving you all the details now. I invite you to discuss this over a coffee later on, if you are interested in hearing my perspective." (only works if the attack was weak – and not in your thesis defence!)

Research Conduct

Death of a researcher

> “At June 13 2019, just before the ISCA 2019 conference in Phoenix, a doctoral candidate, whose paper was supposed to be published in this conference, hanged himself in the campus building of University of Florida.
>
> As more and more information comes out, we find that there should be a lot of hidden stories behind his suicide. The most important one may related to his struggling against an academic misconduct. . . .

From a story published on Medium.com, 29 June 2019

Good Practice, Research Ethics, and Academic Misconduct

Good conduct in research is of utmost importance:

- Academic misconduct produces misinformation and injustice
- Societies and organisations often rely on research for significant decisions
- Science and academia as a whole must earn and defend its public credibility
- Tolerating bad research practices weakens an academic institution/area/ecosystem for a long time

Good Practice, Research Ethics, and Academic Misconduct

Good conduct in research is of utmost importance:

- Academic misconduct produces misinformation and injustice
- Societies and organisations often rely on research for significant decisions
- Science and academia as a whole must earn and defend its public credibility
- Tolerating bad research practices weakens an academic institution/area/ecosystem for a long time

Good academic conduct is a concern on many levels:

- Europe, e.g., The European Code of Conduct for Research Integrity (ALLEA – All European Academies)
- Germany, e.g., White Paper “Safeguarding Good Scientific Practice” (DFG – Deutsche Forschungsgemeinschaft)
- Professional societies, e.g., ACM Code of Ethics and Professional Conduct
- Universities, e.g., Good Scientific Practice at TU Dresden

Good Research Practices

ALLEA lays out the following fundamental principles of good research:

- **Reliability** in ensuring the quality of research, reflected in the design, the methodology, the analysis and the use of resources.
- **Honesty** in developing, undertaking, reviewing, reporting and communicating research in a transparent, fair, full and unbiased way.
- **Respect** for colleagues, research participants, society, ecosystems, cultural heritage and the environment.
- **Accountability** for the research from idea to publication, for its management and organisation, for training, supervision and mentoring, and for its wider impacts.

[The European Code of Conduct for Research Integrity, Revised Edition, ALLEA 2017]

Good Research Practices – simplified

ALLEA lays out the following fundamental principles of good research:

- **Reliability** – take all necessary care!
- **Honesty** – don't lie!
- **Respect** – don't cheat or harm others!
- **Accountability** – take responsibility!

[The European Code of Conduct for Research Integrity, Revised Edition, ALLEA 2017]

Academic misconduct

Major forms of misconduct include:

- **Fabrication** (making up fake results)
- **Falsification** (manipulating or misrepresenting findings)
- **Plagiarism** (stealing material or ideas)
- **Exploiting positions of power**
  over humans (bullying, abuse, exploitation of labour, encouraging misconduct)
  over information (breach of confidence, use of confidential information)

Other examples of misconduct

Many further examples of academic misconduct can be given:

- Manipulating authorship or denigrating the role of other researchers in publications
- Re-publishing substantive parts of one's own earlier publications, including translations, without duly acknowledging or citing the original ('self-plagiarism')
- Citing selectively to enhance own findings or to please editors, reviewers or colleagues
- Expanding unnecessarily the bibliography of a study
- Allowing funders/sponsors to jeopardise independence in the research process or reporting of results so as to introduce or promulgate bias
- Misrepresenting research achievements
- Exaggerating the importance and practical applicability of findings

[The European Code of Conduct for Research Integrity, Revised Edition, ALLEA 2017]

Examples of indirect and “meta” misconduct

Academic misconduct also extends beyond own active research:

- Withholding research results
- Accusing a researcher of misconduct or other violations in a malicious way
- Delaying or inappropriately hampering the work of other researchers
- Misusing seniority to encourage violations of research integrity.
- Ignoring putative violations of research integrity by others or covering up inappropriate responses to misconduct or other violations by institutions.
- Establishing or supporting journals that undermine the quality control of research (‘predatory journals’).

[The European Code of Conduct for Research Integrity, Revised Edition, ALLEA 2017]

$\rightsquigarrow$ conducting research also entails a responsibility for safeguarding good practices

Responsibility, motives, negligence

Judging of misconduct is difficult in practice:

- all authors of a study must be accountable for its contents

But:

- errors are unavoidable in research
- researchers must be free to pursue unorthodox ideas against the mainstream

Responsibility, motives, negligence

Judging of misconduct is difficult in practice:

- all authors of a study must be accountable for its contents

But:

- errors are unavoidable in research
- researchers must be free to pursue unorthodox ideas against the mainstream

Investigations of misconduct involve hard questions:

- Did the accused act with malicious intent? (violation of honesty)
- Did the accused neglect duties for ensuring quality? (violation of reliability)
- In either case: to what extent?

$\rightsquigarrow$ clear-cut decisions are easier in theory than in practice ...

Presentation and analysis of experimental results

A problematic field of honesty and reliability is the presentation of experimental results in research reports:

- Results are often illustrated or analysed in misleading ways
- Statistics is abused to pretend reliability
- Visualisations are adjusted to underline a desired message

$\rightsquigarrow$ the boundary between neglect, incompetence, and misconduct is often fuzzy

Meaningful statistical analysis

Correct use of statistical methods requires prior training!

Most common error: applying a statistical method to data that does not satisfy the minimal requirements for using this method

Meaningful statistical analysis

Correct use of statistical methods requires prior training!

Most common error: applying a statistical method to data that does not satisfy the minimal requirements for using this method

Example: Even a simple measure like the arithmetic mean has particular requirements: it can only be computed on data for which distances are absolute (data with linear units).

Meaningful statistical analysis

Correct use of statistical methods requires prior training!

Most common error: applying a statistical method to data that does not satisfy the minimal requirements for using this method

Example: Even a simple measure like the arithmetic mean has particular requirements: it can only be computed on data for which distances are absolute (data with linear units).

- The difference between $10°C$ and $20°C$ is the same as between $50°C$ and $60°C$, so it makes sense to compute average temperatures in $°C$.
- The difference in aptitude from IQ 90 to 100 is not the same as from 170 to 180, so average IQs are meaningless. Same for many other skill measures (e.g., numerical grades).
  $\rightsquigarrow$ use median instead
- Even medians cannot be used if data has no meaningful order, e.g., if a questionnaire allows people to answer with one of “yes”, “no”, and “not applicable”.
  $\rightsquigarrow$ use mode instead
- Averages are bad measures of central tendency on heavily skewed data (incl. exponential distributions), e.g., the average number of Twitter followers or the average wealth
  $\rightsquigarrow$ at least give median and mode as well
- Averages do not make sense on logarithmic measures

Standard deviation only makes sense if an (arithmetic) mean makes sense!

Meaningful statistical analysis (2)

A common issue is that unjustified assumptions are made in order to apply a preferred analysis.

Examples:

- To use statistics on ordinal data (ordered, but without distance measure), one often assigns arbitrary numbers to values (example: questionnaire answers “agree”, “unsure”, “disagree” are mapped to 1, 0, -1). The result is mostly meaningless.
- Even strictly metric data may have gaps (cases where a no value exists), which sometimes are filled using unjustified assumptions. For example, inputs where an evaluation caused an “out of memory error” cannot be included in an average runtime computation by assuming 10min.
  (Exercise: how about “timeout errors” with timeout 10min?)
- Statistical methods that make predictions for the whole of a population based on a sample have complex and rigorously defined applicability criteria. Especially user studies must take these into account.

Meaningful presentation of results

Diagrammatic visualisation may also misrepresent data in many ways

- Axes in diagrams must be clearly labelled (unit and scale)
  - Non-linear (e.g., log-scale) labelling of axes must be emphasised in the text
  - Truncated y-axes (not starting at zero) may over-emphasize differences
- Trend lines or curves only make sense if both axis are metric
  - In particular: never connect measurements from individual experiments with lines (use bar charts)
- Some chart types (e.g., bubble charts) are hard to read accurately
  - Prefer the simplest, cleanest presentation of your data
  - Don't toy around with 3D effects or shadows

Texas sharpshooter

Given enough data, it is highly likely that some random effects can be shown to be significant even when using rigorous mathematical methods

The Texas sharpshooter: Experiments are conducted and results are analysed to find significant effects. A hypothesis for explaining these effects is fabricated a posteriori and presented as a new theory that is thereby experimentally confirmed.

This scheme also hides the full datasets and experimental setup to give the false impression of a repeatable (non-random) effect.

Other forms of hiding “unpleasant” results are equally dishonest and cases of misconduct.

Plagiarism

Plagiarism

Plagiarism occurs in two main forms:

- Using text and graphics from others without proper attribution
  - Usually entails copyright violation
  - Applies also to slight rephrasing and translation
  - Typical offence of students

  $\rightsquigarrow$ easy to spot and to sanction
- Using ideas and insights from others without proper attribution
  - Source: confidential (e.g., review) or preliminary (e.g., poster publication)
  - Hard to prove (can also be coincidental or unawares)
  - Typical for senior researchers; major concern in some areas of research

  $\rightsquigarrow$ tricky, but claims of priority can be backed by archival pre-publication

"Authors acknowledge important work and intellectual contributions of others, including collaborators, assistants, and funders, who have influenced the reported research in appropriate form, and cite related work correctly. [...] Reviewers and editors respect the rights of authors and applicants, and seek permission to make use of the ideas, data or interpretations presented." – ALLEA, 2017

Self-plagiarism

> “Re-publishing substantive parts of one’s own earlier publications, including translations, without duly acknowledging or citing the original.” – ALLEA, 2017

This concept may scare students and young academics too much:

- Self-plagiarism is to falsely make one research result appear as several findings
- The focus is on reporting the same result (idea) twice on the same publication level

Self-plagiarism

> "Re-publishing substantive parts of one's own earlier publications, including translations, without duly acknowledging or citing the original." – ALLEA, 2017

This concept may scare students and young academics too much:

- Self-plagiarism is to falsely make one research result appear as several findings
- The focus is on reporting the same result (idea) twice on the same publication level

The following is not self-plagiarism:

- Publishing parts of a thesis as a research paper (or vice versa) – never a problem
- In computer science: Publishing about one research activity first at a workshop (early findings), then at a conference (conclusive results), and finally in a journal (extended, revised, completed account), where each cites the earlier appropriately
- Re-presenting own recent results at "informal" venues (e.g., certain workshops) – practices vary from community to community
- Re-using non-essential text pieces (e.g., part of the preliminary definitions or a general part of the introduction) in several publications

Authorship

Authorship

National and international institutions agree that authorship requires a significant contribution to :

- the **design of the research** and **relevant artefacts**,
- **relevant data collection** or **formal analysis**, or
- the **analysis** or **interpretation of the results**

Similar formulations are adopted by DFG, TU Dresden, and ALLEA.

> "All authors agree on the sequence of authorship, acknowledging that authorship itself is based on a significant contribution to the design of the research, relevant data collection, or the analysis or interpretation of the results." – ALLEA, 2017

(This formulation forgets theoretical or design-based research contributions.)

Authorship

National and international institutions agree that authorship requires a significant contribution to :

- the **design of the research** and **relevant artefacts**,
- **relevant data collection** or **formal analysis**, or
- the **analysis** or **interpretation of the results**

Similar formulations are adopted by DFG, TU Dresden, and ALLEA.

> "All authors agree on the sequence of authorship, acknowledging that authorship itself is based on a significant contribution to the design of the research, relevant data collection, or the analysis or interpretation of the results." – ALLEA, 2017

(This formulation forgets theoretical or design-based research contributions.)

Disputes of authorship?

- seek mediation by a responsible ombudsperson
- raise complaint with publication venue

A quote from the DFG guidelines, adopted by TU Dresden

The following contributions on their own are not sufficient to justify authorship:

- merely organisational responsibility for obtaining the funds for the research,
- providing standard investigation material,
- the training of staff in standard methods,
- merely technical work on data collection,
- merely technical support, such as only providing equipment or experimental animals,
- regularly providing datasets only,
- only reading the manuscript without substantial contributions to its content,
- directing an institution or working unit in which the publication originates.

Help of this kind can be acknowledged in footnotes or in the foreword.

"Honorary authorship" is generally not considered to be acceptable under any circumstances. Neither the position of institute director and supervisor nor former supervisor justify designation as co-author.

Obstructing publication

Good practice requires that:

- all authors agree on their joint authorship before submitting a work for publication
- authors who made significant contributions cannot be omitted from the author list without their consent

$\rightsquigarrow$ authors can block publications of research they contributed to by not explicitly agreeing to be involved

Obstructing publication

Good practice requires that:

- all authors agree on their joint authorship before submitting a work for publication
- authors who made significant contributions cannot be omitted from the author list without their consent

$\rightsquigarrow$ authors can block publications of research they contributed to by not explicitly agreeing to be involved

However, this power must not be used to obstruct legit publications:

- DFG: “Refusals to publish must be justified with verifiable criticism of data, methods or results. Should co-authors suspect an obstructive refusal to give agreement, they must ask ombudspersons [. . .] to mediate.”
- ALLEA names these as examples of misconduct: “Delaying or inappropriately hampering the work of other researchers” and “Withholding research results”

Data management

The requirement for accountability involves the need to preserve relevant data for later inspection:

- DFG: “Primary data as the basis for publications shall be securely stored for ten years in a durable form in the institution of their origin.”
  $\rightsquigarrow$ “securely stored” $\neq$ “openly published” (may not be possible)
- DFG: “The disappearance of primary data from a laboratory is an infraction of basic principles of careful scientific practice and justifies a prima facie assumption of dishonesty or gross negligence.”
- ALLEA: “Researchers, research institutions and organisations ensure access to data is as open as possible, as closed as necessary, and where appropriate in line with the FAIR Principles (Findable, Accessible, Interoperable and Re-usable) for data management.”
- Make it as easy as possible for others to analyse and replicate the findings

Experiments with humans

Experiments with human subjects may require special approval by an ethics board.

- Ethical concerns rarely apply to plain user studies with software
  - Focus on studying the software, not the human
  - Routine computer usage situations
  - Volunteer participants mostly from academic environment
- Warning flags:
  - experimental setup involves deceiving participants (e.g., about true purpose of study)
  - some physical or medical aspect is involved (e.g, software usage under the influence of alcohol)
  - participants involve vulnerable groups of people, such as children
  - experiment may cause significant emotional or cognitive stress
  - confidential data is being collected (e.g,. details on private computer usage)

$\rightsquigarrow$ be considerate; if in doubt, seek competent advice

Supervision, guidance, pressure

Guidelines emphasise the responsibility of academic supervisors:

- Convey standards of good practice and discourage any form of misconduct
- Avoid incentives for fraud by valuing quality over quantity

> "Universities and research institutes shall always give originality and quality precedence before quantity in their criteria for performance evaluation. This applies to academic degrees, to career advancement, appointments and the allocation of resources." – DFG guidelines

$\rightsquigarrow$ quality is much harder to fake

- Second supervisors, additional mentors, and impartial mediators (ombudspeople) should be available to give guidance and to raise concerns

Conflicts of interest

Conflicts arise when reporting own research and when evaluating others and their work

Some typical reasons for conflicts:

1. First-degree relationship, marriage/life partnership/domestic partnership
   - Possibly extended to other close personal ties
2. Personal financial interests or financial interest by persons listed under (1)
   - Includes cases of (conditionally) sponsored research and indirect economic interests (e.g, owning a related start-up company)
3. Current or planned close scientific cooperation
   - Usually includes joint publications (in the past three years or in preparation)
4. Dependent employment relationship or supervisory relationship (e.g. teacher-student relationship up to and including the postdoctoral phase)
   - Doctoral advisor often considered to be conflicted for life
   - Former employment considered void after 3–6 years
5. Close competition or strong personal animosity
6. Possibly: recent mutual reviewing, recent competition, etc.

(see also the DFG Guidelines for Avoiding Conflicts of Interest)

How to handle conflicts

- **Reviewing:** Conflicts of reviewers should be avoided (use other reviewers)
  - In committees that select a single winner: conflicts with one applicant/submission suffice (it is not enough to abstain when discussing this one case, since the discussion on others will affect his or her chances)
- **When being reviewed:** complaints can be raised with responsible people or neutral ombudsperson
- **Publishing:** Potential conflicts must be disclosed (e.g., in publications)

> "All authors disclose any conflicts of interest and financial or other types of support for the research or for the publication of its results." – ALLEA, 2017

Good research practice at TU Dresden

TU Dresden has its own guidelines for good research practice

- Implementation of the DFG guidelines with additional information
- List of university-level contact persons for reporting suspected misconduct or for asking for deliberation in case of conflicts
- The Faculty of Computer Science has appointed own ombudspeople: Prof. Christine Baier and Prof. Horst Reichel
- It is also possible to contact the DFG-supported national commission for good academic practice: see `https://ombudsman-fuer-die-wissenschaft.de/`

Summary

Good research conduct requires reliability, honesty, respect, and accountability

Main forms of academic misconduct are fabrication, falsification, plagiarism, and exploitation

There are clear, undisputed principles of plagiarism and attribution (authorship), but there is a lot of grey area in practice

Discrete, professional advisors are available at faculty, university, and national level

What's next?

- Examinations