LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

# Advanced Seminar Media Informatics

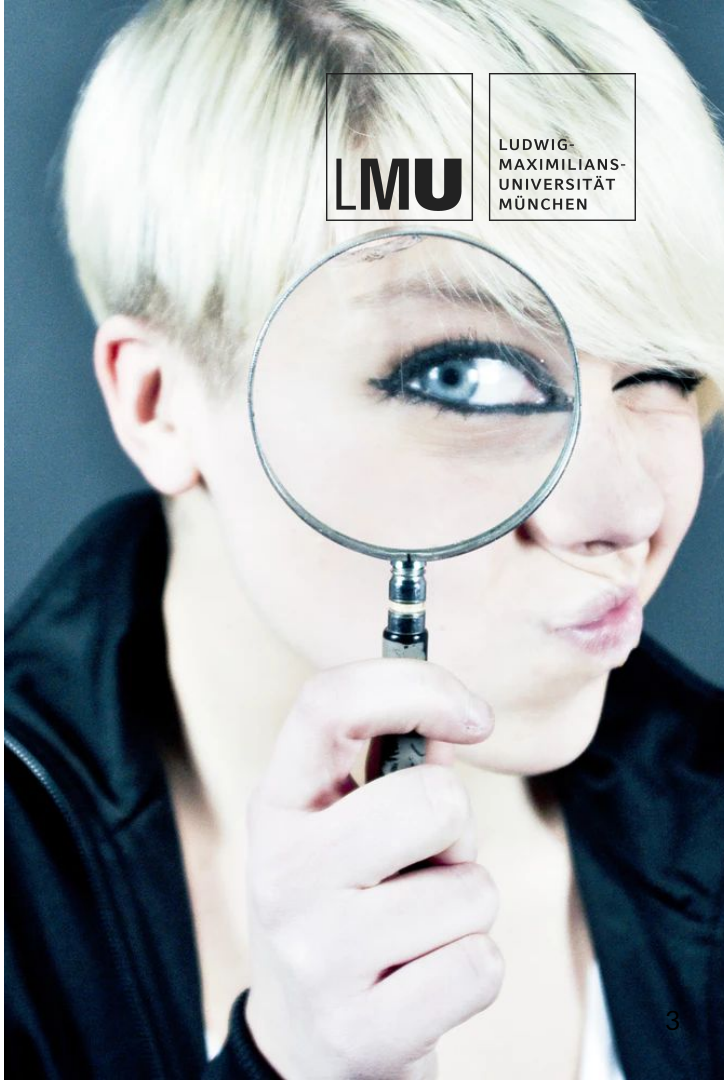
Sarah Theres Völkel | Matthias Schmidmaier | Prof. Dr. Heinrich Hußmann
Summer 2019

# Today

1. 60s pitches + 60s Q&A
2. Why reviewing
3. How to review?
4. Your next steps

# Why reviewing

## Ensuring scientific quality



International Archives of Medicine

Current Archives About

Home / Archives / Vol 8 (2015) / Endocrinology

# Retraction notice on "Chocolate with high Cocoa content as a weight-loss accelerator"

Bild

INFOS ZU BILDPLUS WETTER 23°C SCHWERIN EPAPER KONTAKT ZEITUNGSABO BILD SHOP LOGIN

BILDplus NEWS POLITIK GELD UNTERHALTUNG SPORT FUSSBALL LIFESTYLE RATGEBER REISE AUTO DIGITAL SPIELE REGIO VIDEO

04.06.2019 - 10:00 UHR HOME › RATGEBER › GESUNDHEIT › ERFOLGREICH ABNEHMEN › WER SCHOKOLADE ISST, BLEIBT SCHLANK!

DIESE STUDIE SCHMECKT UNS

# Wer Schokolade isst, bleibt schlank!

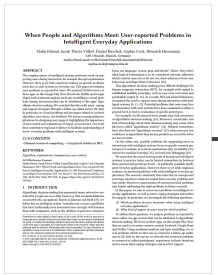
# Why reviewing

- Maintain **standards** of quality
- Provide **credibility** to results
- Make suggestions for **improvement**

# How to review

‘I would strongly recommend to reject this paper!’

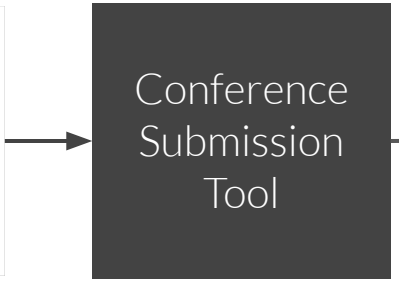
# Review Process

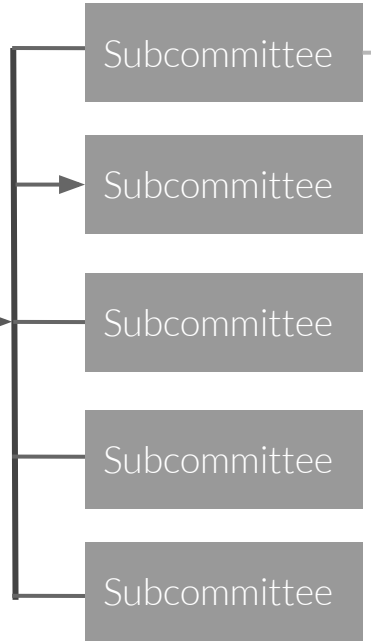

When People and Algorithms Meet: User-reported Problems in Intelligent Everyday Applications

Malin Eiband, Sarah Theres Völkel, Daniel Buschek, Sophia Cook, Heinrich Hussmann

ABSTRACT

1 INTRODUCTION


Conference Submission Tool


Subcommittee

Subcommittee

Subcommittee

Subcommittee

Subcommittee

AC

AC

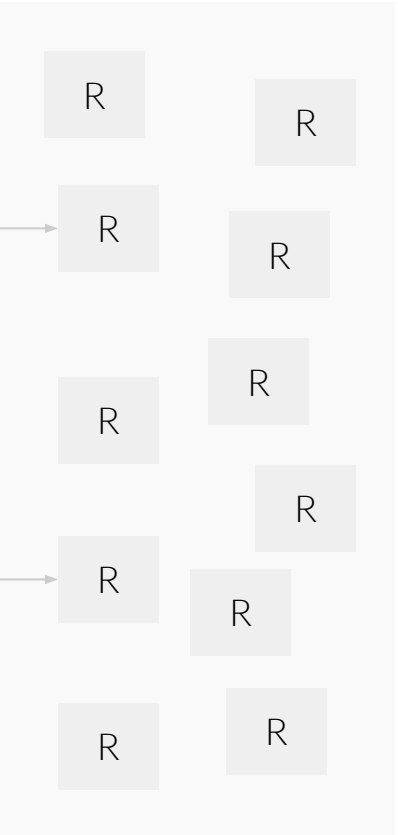
AC

AC

AC

AC

AC


R

R

R

R

R

R

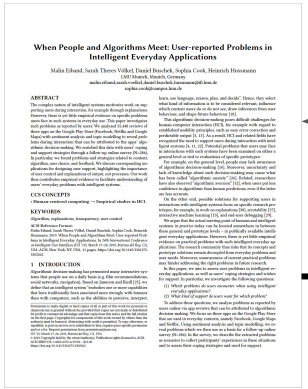
R

R

R

R

# Peer review



When People and Algorithms Meet: User-reported Problems in Intelligent Everyday Applications

Malin Eiband, Sarah Theres Völkel, Daniel Buschek, Sophia Cook, Heinrich Hussmann

LMU Munich, Munich, Germany

malin.eiband,sarah.voelkel,daniel.buschek,hussmann@ifi.lmu.de

sophia.cook@campus.lmu.de

ABSTRACT

The complex nature of intelligent systems motivates work on supporting users during interaction, for example through explanations, control, and feedback. However, current research lacks knowledge of users' actual experiences in everyday use. This paper investigates such problems as reported by users. We analyzed 31,448 reviews of three popular apps that integrate intelligent components. Inspired by Hagras' call for combined analysis and topic modelling to reveal problems during interaction that can be attributed to the apps' algorithmic decision-making. We translated this data into users' coping and support strategies through a follow-up online survey (N=286). Finally, we describe the extracted problems, user strategies, current algorithm, user choices, and feedback. We discuss corresponding implications for designing user support, highlighting the importance of user control and explainability in critical and pervasive. Our work thus contributes empirical evidence to facilitate understanding of users' everyday problems with intelligent systems.

CCS CONCEPTS

• Human-centered computing → Empirical studies in HCI;

KEYWORDS

Algorithm, explanations, transparency, user control

ACM Reference Format:

Malin Eiband, Sarah Theres Völkel, Daniel Buschek, Sophia Cook, Heinrich Hussmann. 2019. When People and Algorithms Meet: User-reported Problems in Intelligent Everyday Applications. In 24th International Conference on Intelligent User Interfaces (IUI '19), March 17–20, 2019, Marina del Ray, CA, USA. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3301275.3302262

1 INTRODUCTION

Intelligent systems, including ever-pervasive[d] many interactive systems that people use on a daily basis (e.g. film recommendations, social networks, navigation). Based on Jameson and Riedl [15], we define that an intelligent system "embodies one or more capabilities that have traditionally been associated more strongly with humans than with computers, such as the ability to perceive, interpret, learn, use language, reason, plan, and decide". Hence, they select and present information or services to users depending on input which context users do or do not see, draw inferences from user behaviour, and shape future behaviour [18].

This algorithmic decision-making poses difficult challenges for human-computer interaction (HCI), for example with regard to how users perceive, trust, and control the often complex and unpredictable output [1, 11]. As a result, HCI and related fields have recognized the need to support users during interaction with intelligent systems [1, 18, 32]. Various guidelines for how people can interact/use with such systems have been examined on either a general level or tied to evaluations of specific prototypes.

For example, on the general level, people may lack awareness of algorithmic decision-making [16]. Moreover, uncertainty and the fact that machine learning algorithms rarely achieve 100% have been called "algorithmic anxiety" [26]. Related, researchers have also observed "algorithmic aversion" [12], when users put less confidence in algorithmic than human predictions, even if the latter are inferior.

On the other end, possible solutions for supporting users in interactions with intelligent systems focus on specific research prototypes, for example, in work on explanations [36], accessibility [17], or control [19]. Overall, this leads to a fragmented picture.

We argue that the actual meeting point of human and intelligent systems in practice today can be located somewhere in between these general and prototype levels – in publically available intelligent everyday applications. However, there is yet little empirical knowledge about problems that users face with such real-world applications. The research community thus risks that its concepts and prototype solutions remain decoupled from everyday problems and users' needs. This knowledge gap motivates the research questions we have broken addressing the right problems in future research.

In this paper, we aim to assess user problems in intelligent everyday applications, as well as users' coping strategies and wishes for support. In particular, we investigate the following questions:

(1) *What problems do users encounter* when using intelligent everyday applications?

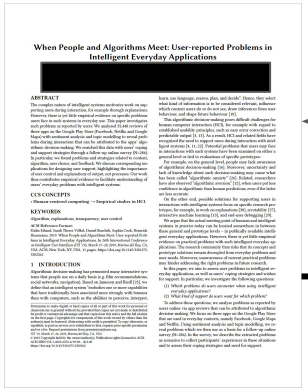
(2) *What kind of support* do users want for which problem?

To address these questions, we analyse problems as reported by users online via app reviews that can be attributed to algorithmic decision-making. We chose reviews of three popular apps that are used in everyday contexts, namely Facebook, Google Maps and Netflix. Using sentiment analysis and topic modelling, we extract problems which we then use as a basis for a follow up online survey (N=286). In the survey, we describe the extracted problems to users and ask them to indicate their coping strategies, wishes and to assess their coping strategies and need for support.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

IUI '19, March 17–20, 2019, Marina del Ray, CA, USA
https://doi.org/10.1145/3301275.3302262

# Peer review


When People and Algorithms Meet: User-reported Problems in Intelligent Everyday Applications

[redacted]

ABSTRACT

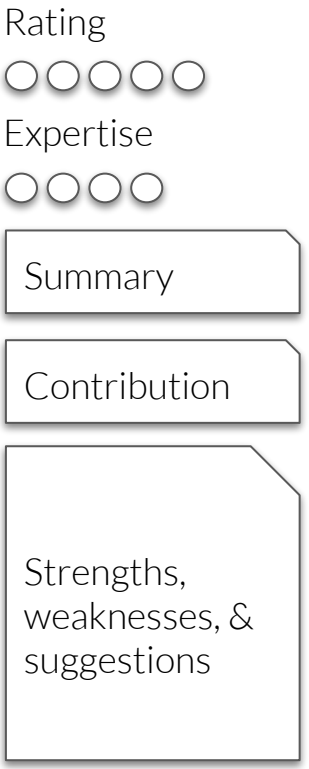
CCS CONCEPTS

KEYWORDS

1 INTRODUCTION

# Elements of a review

- **Reviewer expertise**
  - 1: no knowledge to 4: expert knowledge
- **Short summary of the text**
  - “This paper presents/examines/…”
- **Contribution statement**
  - “This paper contributes … to …”
- **Strengths and weaknesses**
  - Major concerns regarding originality, significance, and validity
  - Smaller concerns (e.g., legibility of tables)
  - Suggestions for improvements
- **Overall rating / recommendation:**
  - 1: definitely reject to 5: definitely accept



# Typical questions to ask

- How **relevant** is this work?
- How is the **argument flow**?
- What is **new** about this work?
- Which **problem** is this work trying to **solve**?
- Which other works does it **extend**?
- **What** did the authors **do**?
- **How** did they **do** it?
- **How valid** is their approach?
- How **relevant** are the results?

# Criteria for a good paper

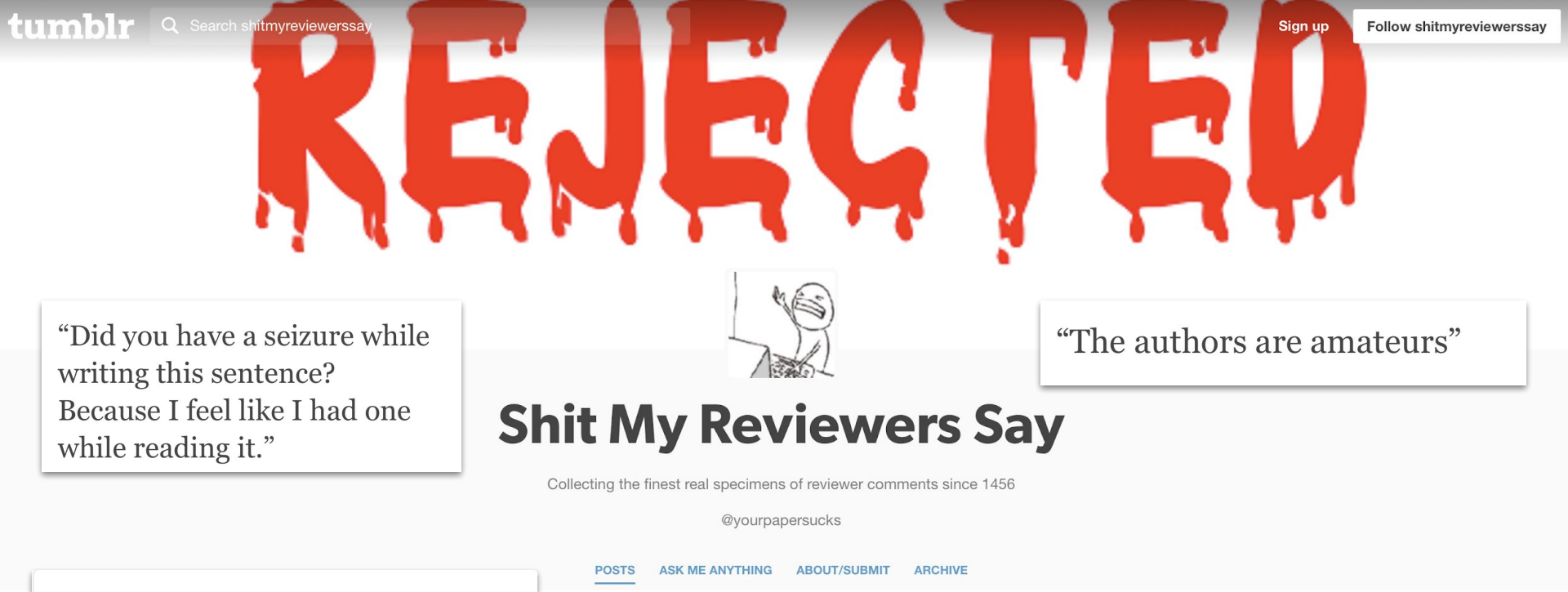
- **Contribution:** Which new insights does the paper contribute to the field of HCI?
- **Benefits:** What can other researchers/practitioners learn from the paper and do with it?
- **Significance:** Is the work useful? Does it have an impact?
- **Novelty/Originality:** Is the work new or are there similar prior publications?
- **Validity:** Are the conclusions justified? Is the methodology sound and analysis valid?
- **Relevance:** How good does the paper match the (HCI) audience?
- **Format:** Legibility, conciseness, clarity

# Critical review

- A review is **NOT** about personal interests or criticism
- The review should focus on content and presentation

**Ethics in scientific communication**

- It is ok to consider a contribution to be superfluous or of no need to the scientific community
- It is not ok to personally judge or insult the authors




"Did you have a seizure while writing this sentence? Because I feel like I had one while reading it."

"The authors are amateurs"

"Publication of this paper will not advance our knowledge in any shape of form, it will just result in other researchers pointing out how bad this study actually is"

"I am personally offended that the authors believed that this study had a reasonable chance of being accepted to a serious scientific journal."

The chapter is missing large portions of related work, especially in the domains of assessing personality from various modalities, correlations with user preferences and recommendation techniques. There are several works from Ferwerda and Tkalcic that deal with various aspects of personality-based personalization. Some examples

Tkalčič, M., Ferwerda, B., Hauger, D., & Schedl, M. (2015). Personality Correlates for Digital Concert Program Notes. In UMAP 2015, Lecture Notes On Computer Science 9146 (Vol. 9146, pp. 364–369). https://doi.org/10.1007/978-3-319-20267-9_32

Skowron, M., Tkalčič, M., Ferwerda, B., & Schedl, M. (2016). Fusing Social Media Cues. In Proceedings of the 25th International Conference Companion on World Wide Web - WWW '16 Companion (pp. 107–108). New York, New York, USA: ACM Press. https://doi.org/10.1145/2872518.2889368

Ferwerda, B., & Tkalcic, M. (2018). You are what you post: What the content of instagram pictures tells about users' personality. In CEUR Workshop Proceedings (Vol. 2068).

Ferwerda, B., Schedl, M., & Tkalcic, M. (2016). Personality Traits and the Relationship with (Non-) Disclosure Behavior on Facebook. In Proceedings of the 25th International Conference Companion on World Wide Web - WWW '16 Companion (pp. 565–568). New York, New York, USA: ACM Press. https://doi.org/10.1145/2872518.2890085

Ferwerda, B., Tkalcic, M., & Schedl, M. (2017). Personality Traits and Music Genres. In Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization - UMAP '17 (pp. 285–288). New York, New York, USA: ACM Press. https://doi.org/10.1145/3079628.3079693

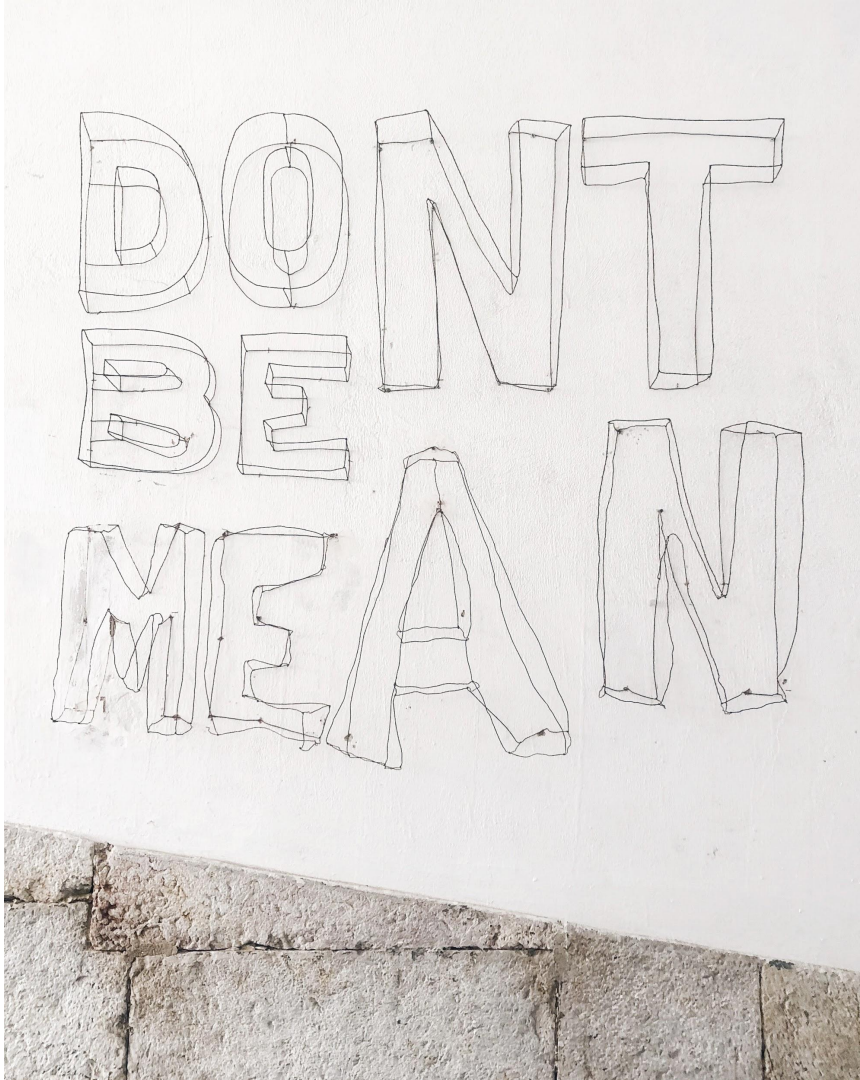
Ferwerda, B., Graus, M., Vall, A., Tkalcic, M., & Schedl, M. (2016). The Influence of Users ' Personality Traits on Satisfaction and Attractiveness of Diversified Recommendation Lists. In M. Tkalcic, B. De Carolis, M. de Gemmis, & A. Košir (Eds.), Proceedings of the 4th Workshop on Emotions and Personality in Personalized Systems co-located with ACM Conference on Recommender Systems (RecSys 2016) (Vol. 1680). Boston, MA.

# Reviewing Rules

The authors have put a lot of **effort** in their submission, and they might be still learning.

Be **accommodating** and **respecting**, and help them improve their work.

That's why we do reviews: to ensure **high quality**



# Overall rating (Smith, 1989)

- Major results - very significant
- Good, solid, interesting work; a definite contribution
- Minor, but positive, contribution to knowledge
- Elegant and technically correct but useless
- Neither elegant nor useful, but not actually wrong
- Wrong and misleading
- The paper is so badly written that a technical evaluation is impossible

# Overall rating (Zobel, 2004)

**Recommending accept**

- Be sure that the paper does not have any serious mistakes
- Convince the AC that the paper has an acceptable standard with respect to originality, significance, and validity
- Make improvements that should be made for the final print version

**Recommending reject**

- Clearly explain the mistakes and if possible, how they could be improved (for a future submission)
- Appreciate the authors' work by indicating which parts are already good and which should be improved / discarded
- Check the paper to a reasonable level of detail

# Some final comments

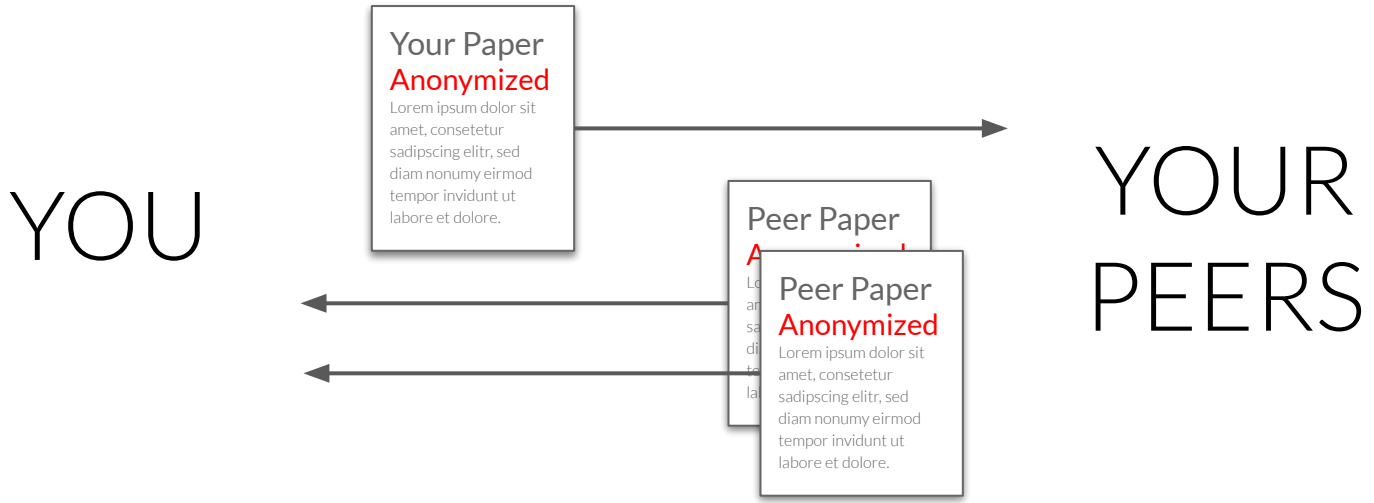
- Also check references, formulas, etc.
- Address the AC in your review by trying to convince her to accept/reject the paper, don't address the authors: "In their paper, the authors present an approach to…"
- Reviewed papers must remain confidential
- Conflicts of interests: reviewers should not review papers of people they worked with

# You will be a reviewer!

# Trade in you paper

→ **submit** anonymized final paper until 18.06.19

→ you will be assigned two peer papers



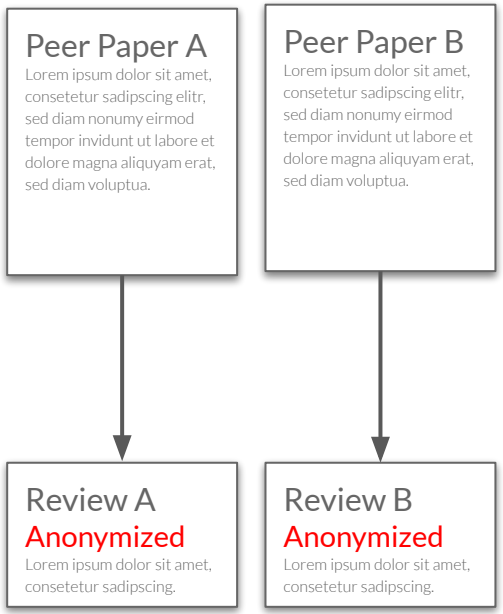
# Write reviews

→ **read** the two assigned papers

→ write a **review** on each paper
guided by these slides and your own intuition

→ **submit** as PDFs until 28.06.19 via Uniworx

YOU



# Don’t worry

Your review won’t influence any grades!

It will only help the authors improve their work.

→ provide constructive feedback

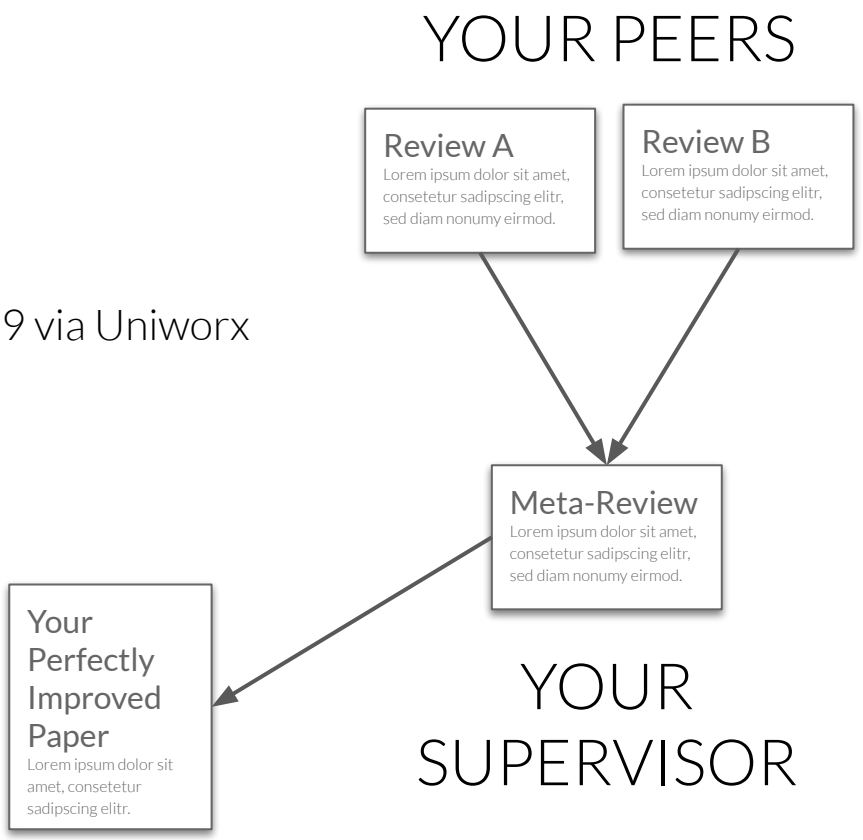
**Be helpful and nice!**

# Make use of your reviews!

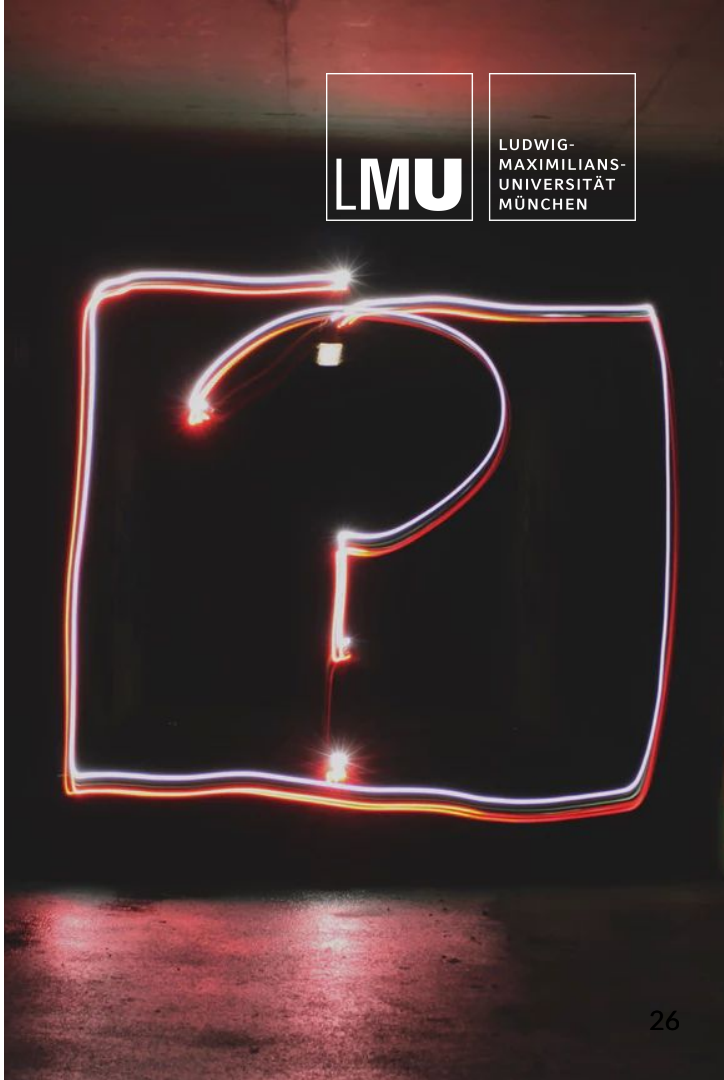
→ your supervisor will write a meta review

→ use your reviews to improve your paper

→ **submit** your final perfect A* paper until 12.07.19 via Uniworx



# Schedule

| | | |
|---|---|---|
| 30.04.19 | Kick-Off | session - compulsory attendance |
| 28.05.19 | 1st draft paper submission | get feedback meet your supervisor before! |
| 02.06.19 | 60s pitch slides submission | |
| 04.06.19 | 60s pitches | session - compulsory attendance |
| 18.06.19 | Review-ready paper submission | |
| 28.06.19 | Review submission | |
| 12.07.19 | Final paper submission | get feedback meet your supervisor before! |
| 14.07.19 | 1st draft slides submission | |
| 21.07.19 | Final slides submission | practice talk with your supervisor! |
| 23.07.19 | Presentation | session - compulsory attendance, about **~5 hours!** |

# Q & A