Scientific Literature (@ TUM)

Benedikt Hauptmann, Henning Femmer

With material from
Dr. Florian Deißenböck, and
Prof. Dr. Stefan Wagner
Goals and Content

1. Methodology: Searching for literature in a scientific field
2. Evaluation: Indicators for quality of scientific papers
3. Technical Aspects (@TUM): Getting a paper for a citation

Two parts:
1. What are scientific publications?
2. Literature search
Why literature reviews?

- Part of your job.

Goals of literature reviews:

- Understand the state of science
- Identify currently open questions
- Show relevance
- Define commonalities and differences with other work (and explain why)
- Place your work in the area of research
- Give evidence for your assumptions
- ...
Scientific Publications

(a tiny introduction)
What kind of paper/articles/… exist?

**Research type facets** [Wieringa2005]

- Exploratory
- Solution
- Experience
- Validation / Evaluation
- Philosophical / Opinion

What kind of paper/articles/… exist?

Jackpot: Meta publications / Research Surveys

Conference on Systems Engineering Research (CSER’13)

A Literature Survey on International Standards for Systems Requirements Engineering

Florian Schneider\textsuperscript{a*}, Brian Berenbach\textsuperscript{b}

\textsuperscript{a}Chair for Applied Software Engineering, Technische Universität München, Boltzmannstr. 3, Garching, 85748, Germany
\textsuperscript{b}Siemens Corporation, Corporate Technology, 755 College Road East, Princeton 08540, USA
Which formats?

- **Book**
  - Usually single-author
  - 100 – 1000 pages
- **Book chapter**
  - 20 – 50 pages
- **Journal article/paper**
  - 10 – 30 pages
- **Proceedings article/paper (conference)**
  - 3 – 15 pages
- **Workshop article/paper**
  - 3 – 15 pages
- **Technical reports**
- **Thesis (Dissertation/Master’s Thesis/Bachelor’s Thesis)**
- **Blogs**
- ...
How are things published?

The peer review system

Paper Ready

Submitted

Under Review

Reviewed

Accepted

Rejected

Published

author submits to conference/workshop

Program committee hands out paper to (2-4) reviewers (experts from the field)

reviewers submit evaluation ("reviews")

Program committee decides based on reviews

Major revision

Accepted with minor changes

Rejected

Reviewed

Accepted

Published

starke Überarbeitung, "Rebuttal"

Why?

Published

author includes final changes and presents at workshop/conference
Publication types and peer reviewing

• **Book**
  – *Mostly single author, no peer review*

• **Book chapter**
  – *Peer review (several round trips)*

• **Journal article**
  – *Peer review (several round trips)*

• **Proceedings article/paper (conference)**
  – *Peer review*

• **Workshop article/paper**
  – *Peer review*

• **Technical Reports**
  – *No peer review*

• **Thesis (Dissertation/Master’s Thesis/Bachelor’s Thesis)**
  – *?*

• **Web pages / blogs**
  – *No peer review*

• …
Review examples...

(from a regular phd student…)
Example: Reviews for a workshop

---------------------------- REVIEW 2 ----------------------------
PAPER: 4
TITLE: Who is the Advocate? Stakeholders for Sustainability
AUTHORS: Birgit Penzenstadler, Henning Femmer and Debra Richardson

--------- REVIEW ---------
The paper attempts to devise a systematic process for identifying advocates of sustainability. The process leverages on well-established/classical requirements engineering processes for stakeholders identification.

+ The paper is well-written and presented. The examples are useful.
+ The use of the requirements techniques to identifying sustainability stakeholders looks to be plausible.

- Combining/contrasting the outcomes the four techniques look to be an expensive exercise. Furthermore, sustainability analysis may also require the identification of an equally important slice of stakeholders, who are the devils' advocates for sustainability. This could be important for promoting behavioral change, analyzing obstacles for the sustainability agenda, understanding their incentives, quantifying the risks to adoption and working on solutions for promoting acceptance.

- I would also expect more justification of the exercise in terms of long term benefits and influences on various software artifacts. The motivation is bit weak to justify the effort.
PAPER: 8
TITLE: Detecting Inconsistencies in Wrappers - A Case Study
AUTHORS: Henning Femmer, Dharmalingam Ganesan, Mikael Stegent

= Overall opinion =
It's a pleasure to read such a self-contained paper that answers pretty much every question the reader raises while reading. It's well-structured and provides both detailed theory and insightful

= Detailed feedback =
- abstract - typo: "instead of to the" -> "instead of the"
- p2 - typo: "an SAL" -> "a SAL"
- p3 - typo: "an analysts" -> "an analyst"
- p3: "The tool detects differences in the function pair and highlights the important ones to the user." -> at this stage we don't understand how the "important ones" are defined. It is also unclear at this stage that the training set starts empty (all functions unclassified) and is incrementally defined/increased. I would suggest to be more explicit when describing step (1) and (3).
- p3 - fig 2: the "very light-red background" is very hard to see; I would strongly recommend to find a better way to highlight.
- p3 - fig 2: why are "uint32" and "free" highlighted although they are identical in both implementations?
- p5: "304 pairwise comparisons": can't thing be optimized here? Eg: if f1 and f2 are equivalent and f2 and f3 too, we don't really need to check for f1 and f3, do we?
- p7 - fig 6: I would suggest to repeat (at least in the label) which OS which side corresponds to (as in "VxWorks on the left and RTEMS on the right")

The paper addresses the equivalence of several given implementations and modification of pa...
4 dimensions vs. views
Btw., why is the technical dimension important? For me it is simply a subset of the economic 'dimension'. Keeping a system sustainable from a technical perspective makes it more profitable from the economic point of view. If it is not 'technically sustainable', it is also not in the economically way.

example of green software code are mobile apps
Why are they green?

#It is essential to explain the values you are using. Why not use some of the examples from the previous chapter?

First some general thoughts.
Sorry to be hard on you, but I do not want to demolish the paper. To me, it looks too much like a list of categories to classify a software product. I would evaluate it as a review paper, might be my fault. Your aim is to review the GSWs. Mine, too, I consider the work of yours. Therefore, the contribution of your paper. When I think back to the paper. Then I remember the contribution of yours.

It should be more clear. If one keeps in mind that the paper is easier to comprehend when evaluating the quality of the work. It might still be interesting. Disturbing/annoying is, that the classification was not done.

Sec 2 Related Work and its Quality

Attention in SW Engineering only recently.
- depends on whether SW E is something different than the production of software
- SW that is now classified as greening (Green by) has been produced since many years in environmental sciences, and there has been environmental informatics for a while. There is the conference series on EnvirolInfo (Environmental Informatics) (27th this year). I had a lectures how to design environmental software in 1994. Early (german) names are e.g. R Grützer, B Page, V. Wohlgemuth. At least since 2003 there is a Journal on Environmental Informatics (methodologies, applications, and policy considerations, the needs for environmental systems analysis, the challenges of environmental systems modeling, and the impacts of environmental informatics are discussed etc., btw., IF 3.619, not that I like Impact Factors).
How *good* is a scientific publication?

*Easier* to judge:

1. Where was it published (venue)?
   - Peer-reviewed venue
   - Impact factor of venue
   - Acceptance rate of venue

2. How was it received in literature?
   - Number of citations

*Use with care...*
### Top General SE Conferences

<table>
<thead>
<tr>
<th>Year</th>
<th>ICSE</th>
<th>FSE/ESEC</th>
<th>ASE</th>
<th>SPLASH/OOPSLA</th>
<th>ECOOP</th>
<th>ISSTA</th>
<th>FASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>85/461 (18%)</td>
<td>51/251 (20%)</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>32/124 (26%)</td>
<td>26/112 (23%)</td>
</tr>
<tr>
<td>2012</td>
<td>87/408 (21%)</td>
<td>34/201 (17%)</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>31/108 (29%)</td>
<td>33/134 (25%)</td>
</tr>
<tr>
<td>2011</td>
<td>62/441 (14%)</td>
<td>34/203 (17%)</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>25/117 (21%)</td>
<td>29/99 (29%)</td>
</tr>
<tr>
<td>2010</td>
<td>50/380 (14%)</td>
<td>34/169 (20%)</td>
<td>37/252 (15%)</td>
<td>61/166 (37%)</td>
<td>?</td>
<td>24/105 (23%)</td>
<td>24/96 (25%)</td>
</tr>
<tr>
<td>2009</td>
<td>50/405 (12%)</td>
<td>32/217 (15%)</td>
<td>38/33/222 (17%)</td>
<td>25/144 (17%)</td>
<td>?</td>
<td>25/93 (27%)</td>
<td>30/124 (24%)</td>
</tr>
</tbody>
</table>

**Conferences (Proceedings)**
- International Conference on Software Engineering (ICSE)
- Foundations of Software Engineering (FSE)
- International Conference on Automated Software Engineering (ASE)
- OO Programming, Systems, Languages and Applications (OOPSLA)
- International Symposium on Software Testing and Analysis (ISSTA)
- International Conference on Software Maintenance (ICSM)

http://web.engr.illinois.edu/~taoxie/seconferences.htm

*in no particular order*
Hands-on: Literature reviews
Where can we find papers?

Publishers:
- ACM Digital Library
- IEEE Xplore
- Springer Link
- Elsevier
- TUM library

Direct sources
- Authors homepages

Meta sources
- scholar.google.com
- Research Gate

We use mostly Google Scholar!
- Pro
  - All results in one place
  - Direct meta-information (citations)
  - Sometimes direct link to PDF
  - Author graphs
- Con
  - No quality filter
  - Few search filter options
Search strategies

1. Manually searching for keywords
2. Searching through author’s pages
3. (Literature-) Snowballing
4. Systematic strategies
   - Systematic Mapping Studies
   - Systematic Literature Reviews

REFERENCES

I know a paper. How do I get the pdf?

- Papers are licensed by the publishers (ACM, IEEE, Springer, Elsevier, …)
- TUM has bought most of the licenses

How do you get access?
- Library computers
- [https://EACCESS.ub.tum.de](https://EACCESS.ub.tum.de) and search for venue/journal
- VPN + Proxy server

Details can be found here:
[https://www.lrz.de/services/netzdienste/proxy/journals-access/](https://www.lrz.de/services/netzdienste/proxy/journals-access/)
Reference management

- BibTeX (+LaTeX)
  - Classical, useful format for LaTeX
  - Most sources offer information in bibtex format
  - Use proper types: @article, ...
Reference management

- BibTeX (+LaTeX)
  - Classical, useful format for LaTeX
  - Most sources offer information in bibtex format
  - Use proper types: @article, ...

- JabRef
  - Plattform independant
  - Uses BibTeX as data format
  - Link to pdf
  - Group

- Mendeley
  - Modern, “in the cloud”
  - Cooperative references
  - Notes, highlighting, etc
  - Bibtex export

- Papers (Mac only)
  - ...

Goals and Content

1. Methodology: Searching for literature in a scientific field
2. Evaluation: Indicators for quality of scientific papers
3. Technical Aspects (@TUM): Getting a paper for a citation

Two parts:
1. What are scientific publications?
2. Literature search
More details

- Kent Beck: How to get a paper accepted at OOPSLA:
  http://plg.uwaterloo.ca/~migod/research/beckOOPSLA.html
  and the corresponding presentation:
  https://www.youtube.com/watch?v=NM3ClIbuVoM
- Zugang zu Wissenschaftlichen Publikationen für Mitarbeiter und Studierende der TUM
  https://www.lrz.de/services/netzdienste/proxy/zeitschriftenzugang/
- Paper-Verwaltung:
  http://www.mendeley.com
  http://jabref.sourceforge.net
  …