Literature Research & Literature Management

Principles and Practices of Scientific Work
Sandro Schulze, PhD (based on material of Katrin Krieger)
Summer term 2017
What's scientific literature anyway?

The only difference between messing around and science is writing it down.
<table>
<thead>
<tr>
<th>Scientific Literature</th>
<th>vs.</th>
<th>Popular Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>facts and truths</td>
<td></td>
<td>fiction, fantasy and stories</td>
</tr>
<tr>
<td>physical manifestation of profound truths</td>
<td></td>
<td>artistic form of written word</td>
</tr>
<tr>
<td>peer-reviewed</td>
<td></td>
<td>not peer-reviewed</td>
</tr>
<tr>
<td>follows a chain of argumentation</td>
<td></td>
<td>follows narrative flow</td>
</tr>
<tr>
<td>written to spread knowledge and insights</td>
<td></td>
<td>for entertainment</td>
</tr>
<tr>
<td>embedded in a frame of work of others</td>
<td></td>
<td>stand-alone</td>
</tr>
</tbody>
</table>
Why do I need to read scientific literature?

- to gain understanding of a scientific field
- to find a topic to work on
- to find and understand open questions
- to see how others have tackled open problems
- to understand the scope of the problem you are working on
The scientific manuscript

The structure of a scientific manuscript

• follows a certain purpose
• is mostly the same in all kinds of scientific publications
• not following the standard structure might lead to rejection of the publication
• helps other scientists to understand and judge the publication
Standard structure

• title
• abstract
• introduction – What is the problem?
• (background/related work)
• methods – How was it studied?
• results – What was found?
• discussion – What does that mean?
• Related work (goes usually here)
• Conclusion + Future Work
• references
Title

• appears in databases and archives
• is the first thing most people read
• base for the decision whether to read the article or not
• be careful with the wording
• avoid abbreviations
• as short as possible, as long as necessary
• is rarely a grammatically complete sentence
Abstract

- is a mini version of the publication
- contains:
  - An overview of the publication
  - The motivation that led to the work
  - Essential contents → what is the contribution
  - Most important results
- goal: The reader has to be able to see whether the paper is relevant or not
Abstract (2)

• the abstract MUST NOT contain
  • claims
  • introduction of new terms
  • discussion
  • references and abbreviations
  • background information
• length: one paragraph
• less than 250 words, better: 100
Abstract (3)

Towards Monitoring of Novel Statements in the News

Michael Färber, Achim Rettinger, and Andreas Harth

Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
{michael.faerber, rettinger, harth}@kit.edu

Abstract. In media monitoring users have a clearly defined information need to find so far unknown statements regarding certain entities or relations mentioned in natural-language text. However, commonly used keyword-based search technologies are focused on finding relevant documents and cannot judge the novelty of statements contained in the text. In this work, we propose a new semantic novelty measure that allows to retrieve statements, which are both novel and relevant, from natural-language sentences in news articles. Relevance is defined by a semantic query of the user, while novelty is ensured by checking whether the extracted statements are related, but non-existing in a knowledge base containing the currently known facts. Our evaluation performed on English news texts and on CrunchBase as the knowledge base demonstrates the effectiveness, unique capabilities and future challenges of this novel approach to novelty.

Keywords: Semantic novelty measures · Novelty detection · Statement extraction
Introduction

- The introduction sets the stage for the rest of the paper
- Goal: The reader understands, why the content of the paper represents scientific progress
- The reader knows:
  - the background of the topic
  - why it is interesting and important
  - why the problem has not been solved (appropriately) so far
Where can I find literature?

- Google Scholar
- IEEE Xplore
- CiteSeerX
- ACM Digital Library
- dblp
- Springerlink
- online catalogue of the university library
- Wikipedia
- ...
How do you recognize „good“/relevant literature?

- abstract = Mini version of the publication – is the content relevant?
- published in a peer-reviewed outlet
- clear structure and concise language
- metrics such as impact factor and citation rate
Tom Heath
Open Data Institute
Semantic Web - Linked Data - Data Integration - Data Science - Open Data
Bestätigte E-Mail-Adresse bei theodi.org

<table>
<thead>
<tr>
<th>Zitationsindizes</th>
<th>Zitate meiner Artikel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alle</td>
<td>1144</td>
</tr>
<tr>
<td>Seit 2008</td>
<td>0</td>
</tr>
<tr>
<td>Zitate</td>
<td>3934</td>
</tr>
<tr>
<td></td>
<td>3964</td>
</tr>
<tr>
<td>h-index</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>17</td>
</tr>
<tr>
<td>i10-index</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>24</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Titel/Autor</th>
<th>Zitier von</th>
<th>Jahr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linked data-the story so far</td>
<td>1965</td>
<td>2009</td>
</tr>
<tr>
<td>C. Bizer, T. Heath, T. Berners-Lee</td>
<td></td>
<td></td>
</tr>
<tr>
<td>International Journal on Semantic Web and Information Systems (IJISWS) 5 (3 ...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linked data: Evolving the web into a global data space</td>
<td>690</td>
<td>2011</td>
</tr>
<tr>
<td>T. Heath, C. Bizer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synthesis Lectures on the Semantic Web: Theory and Technology 1 (1), 1-136</td>
<td></td>
<td></td>
</tr>
<tr>
<td>How to Publish Linked Data on the Web</td>
<td>336</td>
<td>2007</td>
</tr>
<tr>
<td>C. Bizer, R. Cyganiak, T. Heath</td>
<td></td>
<td></td>
</tr>
<tr>
<td>URL: <a href="http://sites.wws.fu-berlin.de/suhl/bizer/pub/LinkedDataTutorial/">http://sites.wws.fu-berlin.de/suhl/bizer/pub/LinkedDataTutorial/</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linked data on the web (LDCW2008)</td>
<td>244</td>
<td>2008</td>
</tr>
<tr>
<td>C. Bizer, T. Heath, K. Idehen, T. Berners-Lee</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
References


Tools

- Jabref
- Citavi
- Mendeley
- BibSonomy
- LibraryThing

Literature research & management – Sandro Schulze – slide #33
How to read scientific papers

- title, abstract, methods, conclusion, remainder
- set yourself goals for reading – what do you want to know?
- skim the text first, then read it more intensely with markings and notes
- divide the text into parts, read each part separately, write a summary for each part (what have I just read?)
- speed reading

https://www.youtube.com/watch?v=3wO7QsA_EDw
Setting Up the Search Process

- Now you have heard about
  - Why to write
  - How to structure a paper
  - Finding the important information
  - Tools that support you

- But how does this help you to find appropriate literature...
  - ...in a goal-oriented way
  - ...with optimized effort
  - ...with good results

- You need to set up a structured process!
Defining keywords

- Make yourself clear about the domain and specific topics relevant for your literature search
- You may already think of research questions you are going to answer

- Based on the above
  - Determine keywords representing your relevant topic(s)
  - Combine keywords to word phrases
  - Connect word phrases using logical operators (OR, AND)
  - Automate, if possible
Define Places & Tools

- Which libraries you focus on?
- How do they cope with (combining) search terms?
- To what extent can the search be automated?

- Where to store the papers found?
- Which information should I store?
Filter Your Results

- Remove duplicates (found in several sources)
- Define criteria for
  - Inclusion → when is a paper really worth to consider
  - Exclusion → aspects that definitively make the paper irrelevant (e.g., language, missing evaluation, short paper, etc.)
- Skim reading papers (title, abstract, and maybe introduction)
  - Take inclusion/exclusion criteria into account
  - Find a way to classify paper (must read, keep it for later, etc.)
- Create a final list of relevant papers....DONE!
References for these slides

- Billy Uber: How to get published in English. Workshop-Materialien

Links

- Improving your reading skills http://www2.le.ac.uk/offices/ld/resources/study/reading
- Reading: Comprehension, Strategies and Activities  http://de.slideshare.net/jabbusch/reading-comprehension-strategies-and-activities?related=4
- Critical reading techniques http://www2.open.ac.uk/students/skillsforstudy/critical-reading-techniques.php