TU Graz

SCIENCE
PASSION
TECHNOLOGY

Search Literature

Roman Kern <rkern@tugraz.at>

706.015 - Introduction to Scientific Working

Search for Literature

Where to find relevant literature for a given topic?

Roman Kern <rkern@tugraz.at>, Institute of Interactive Systems and Data Science, Graz University of Technology

Motivation

- Entering a new field
  - e.g., young researcher
  - e.g., new research interest
  - e.g., new topic
- Writing a proposal
- Writing a paper
- ...

Big Picture

- The **body of research** consists of
  - Papers
  - Communities
  - Terminology

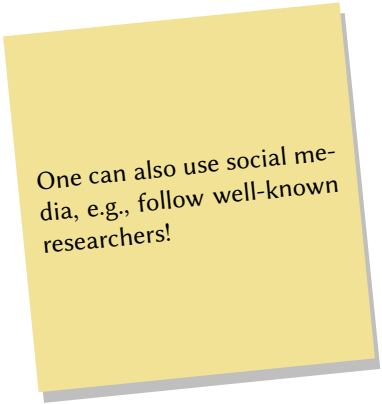
State of the Art

One needs to get a solid understanding of all aspects, not only the literature in isolation.

Roman Kern <rkern@tugraz.at>, Institute of Interactive Systems and Data Science, Graz University of Technology

Strategies

- Strategy “**Search**”
  - Start with a set of key terms
- Strategy “**Traverse**”
  - Start with an initial paper
- Strategy “**Community**”
  - Start with conference, journal, ...
- Strategy “**Projects**”
  - Start with corpus, project, ...



Strategy “Search”

- Identify a number of **keywords**
- Use **search engines** to find papers
- Iterative **process**
  1. Initial keywords
  2. Read papers
  3. Refine keywords

Academic Search Engines

- Google Scholar
  - https://scholar.google.com/
- Semantic Scholar
  - https://www.semanticscholar.org/

Academic Platforms

- ACL Anthology
  - https://www.aclweb.org/anthology/
- ACM Digital Library
  - https://dl.acm.org/
- IEEE Xplore
  - https://ieeexplore.ieee.org/Xplore/home.jsp

Academic Platforms

- ScienceDirect
  - https://www.sciencedirect.com/
- Scopus
  - https://scopus.com/search/
- SpringerLink
  - https://link.springer.com/
- Web of Science
  - https://apps.webofknowledge.com/

Strategy “Traverse”

- Starting with a single (or set) of **papers**
- Go **back** in time
  - Look at papers that are cited by the seed paper
- Go into the **future**
  - Look at papers that cite the seed paper

Strategy “Traverse”

- Need then to select **relevant papers**
- Can be further **restricted**
  - e.g., On year of publication, contain key terms, ...

Strategy “Community”

- Starting with a **conference**
  - Look at the proceedings
- Starting with **journals**
  - Look at the published issues

Roman Kern <rkern@tugraz.at>, Institute of Interactive Systems and Data Science, Graz University of Technology

Strategy “Projects”

- Sometimes **projects** have points to papers
  - e.g., github
- **Datasets** (corpora) often are accompanied by a paper
  - Describes the dataset

Can be combined with other strategies, e.g., use a corpus paper as seed

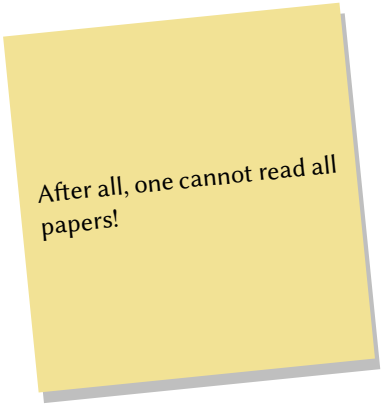
Alternative Sources

Web Pages

- Papers with Code
  - https://paperswithcode.com
- Distill
  - https://distill.pub/
- Mendeley Groups
  - Curated lists of papers (login required)

Paper Selection

- After a number of papers have been collected
- Filter them out
  - Clear criteria
  - e.g., “high-quality” papers, recent papers



Quality Criteria

- Pre-prints
  - https://arxiv.org/
- Conference rank
  - http://portal.core.edu.au/conf-ranks/
  - https://research.com/conference-rankings/computer-science/2021
- Journals
  - Impact factor (e.g., h-index, h5)
- Authors
  - Well-known research groups/universities

Thank You

For your attention!

Roman Kern <rkern@tugraz.at>, Institute of Interactive Systems and Data Science, Graz University of Technology