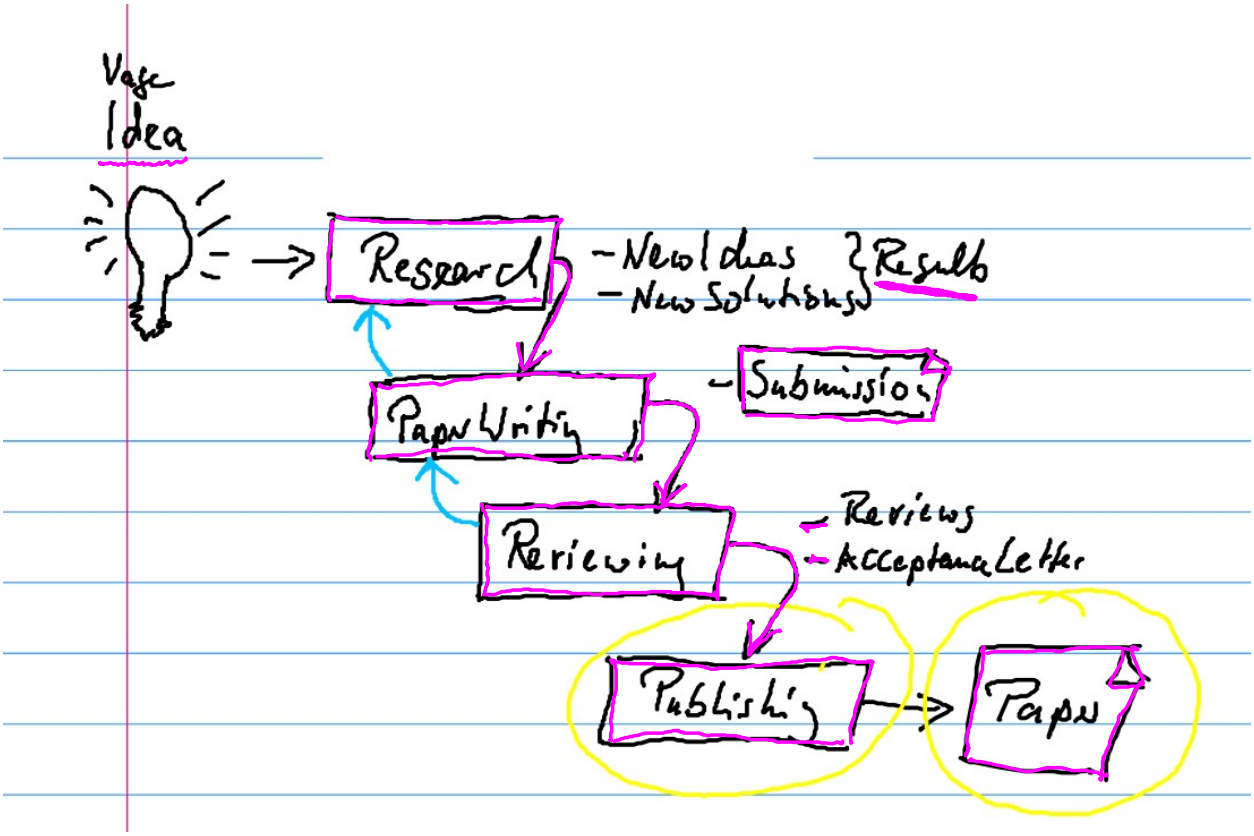
4.6 Publication Process



CHAPTER

FIVE

BIBLIOGRAPHIC DATA AND RANKINGS

5.1 Bibliographic Data

Example of DBLP XML:

Paper xml:

```
<dblp>
  <script/>
  <inproceedings key="conf/tacas/Beyer22" mdate="2022-04-29">
    <author orcid="0000-0003-4832-7662">Dirk Beyer 0001</author>
    <title>Progress on Software Verification: SV-COMP 2022.</title>
    <pages>375-402</pages>
    <year>2022</year>
    <booktitle>TACAS (2)</booktitle>
    <ee type="oa">https://doi.org/10.1007/978-3-030-99527-0_20</ee>
    <crossref>conf/tacas/2022-2</crossref>
    <url>db/conf/tacas/tacas2022-2.html#Beyer22</url>
  </inproceedings>
</dblp>
```

Crossrefs:

- ORCID 0000-0003-4832-7662
  - lookup on DBLP
  - lookup at orcid.org
- Conference conf/tacas/2022-2

5.2 Measurements and Rankings

Be careful about the interpretation of measurements!

- $\text{citationCount} : \text{Authors} \to \mathbb{N}$
  $\text{citationCount}(a) = \left|\{p_2 \in \text{Papers} \mid \exists p_1.(a, p_1) \in \text{published} \wedge (p_2, p_1) \in \text{cites}\}\right|$

- $\text{number papers} : \text{Authors} \to \mathbb{N}$
  $\text{number papers}(a) = \left|\{p_1 \in \text{Papers} \mid (a, p_1) \in \text{published}\}\right|$

- $\text{number of coauthors} : \text{Authors} \to \mathbb{N}$
  $\text{number of coauthors}(a) = \left|\{a_2 \in \text{Author} \mid \exists p \in \text{Papers}: (a, p) \in \text{published} \wedge (a_2, p) \in \text{published} \wedge a \neq a_2\}\right|$

- $\text{citation count} : \text{Papers} \to \mathbb{N}$
  $\text{citation count}(p) = \left|\{p_2 \in \text{Paper} \mid (p_2, p_1) \in \text{cites}\}\right|$

- $\text{papers with n citations} : \text{Author} \times \mathbb{N} \to \mathbb{N}$
  $\text{papers with n citations}(a, n) = \left|\{p \in \text{Paper} \mid \text{citation count}(p) \geq n \wedge (a, p) \in \text{published}\}\right|$

- $\text{h-index} : \text{Author} \to \mathbb{N}$
  $\text{h-index}(a) = \max\{n \in \mathbb{N} \mid \text{papers with n citations}(a, n) \geq n\}$

Operational computation of h-index:

1. Sort all papers of authors in decreasing order of citation count.
2. Count downwards as long as citation count >= counter.

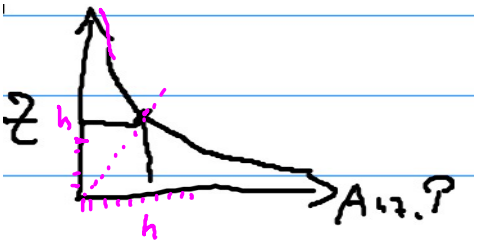
Example:

- Publication A: 5 citations
- Publication B: 10 citations
- Publication C: 3 citations
- Publication D: 1 citation
- Publication E: 7 citations

1. Sorted:
   1. B 10
   2. E 7

3. A 5
4. C 3
5. D 1

2. Count until 3 => h=3



Common understanding of h-index for a person:

- > 10: W2 professor
- > 20: W3 professor
- > 30: excellent reputation
- > 40: outstanding

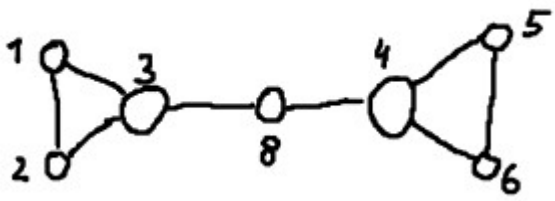
Heavily biased:

- age of scientiest/journal/institution
- area of research (some cite more broadly, others more narrowly)
- collaboration culture
- size and volume of institutions/conferences/journals

Other measures:

- h5 for publication venues (h-index for publications of last 5 years)
- i10 for people (#papers >= 10 citations)

5.3 Why Rankings?



- Milestones / career steps
- Quantify objects
- Select universities (by rank, research fields, . . .)
- Funding decisions
- Hiring decisions (select University/Department/Professor for your PhD)
- Legitimization of decisions

- Select publication venues

5.4 Important Rankings and Platforms

- Google Scholar: Scientists (h, i10, #citations, #papers) and Venues (h5)
- Core Ranking (Australia): Conference rankings by tier (A*, A, B, C, ...)
- csrankings.org
- csmetrics.org

CHAPTER

SIX

CAREER

6.1 Selecting a research topic

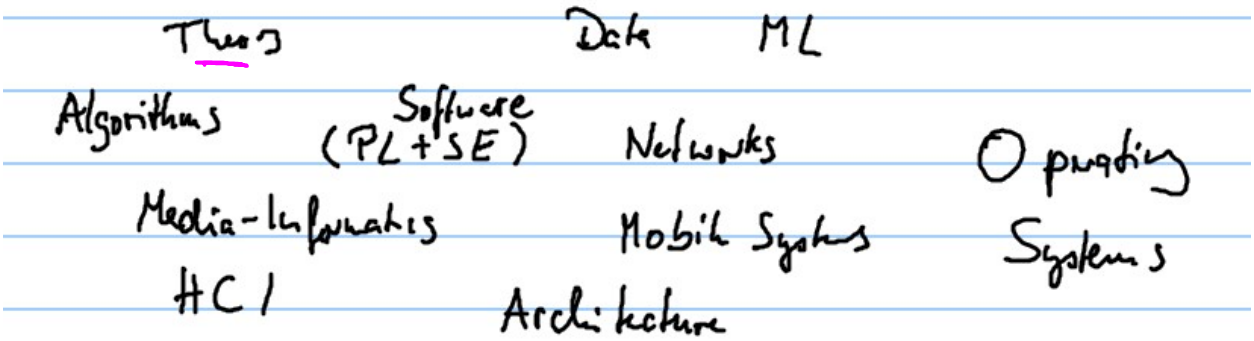
- What interests me?
- What motivates me?
- What am I good at, what do I have a vocation for? (“sich berufen fuehlen”)
- Strengths vs. weaknesses, skills
- With who do I want to/can I work?

Questions to the Professor:

- Why this Job?
- How does a research group work?
  - “family”
  - “liking each other”
  - “Doktormutter” Academic parent
  - mentor

6.2 Identifying Research Areas

- Structure, Divide and Conquer
- Theory vs. Practice



But what is *Computer Science*?

1. look at publications
2. create research clusters
    - list of computer scientists
    - list of their publication venues
3. Invent cluster names (e.g., collect and sort words that appear in conference names)
4. Rank topics (– which rankings?)