Science and Practice in Software Engineering

CHAPTER

THREE

PUBLICATIONS

3.1 Open Access and Open Science

- People **always** have been asking for **access** to knowledge
- But access is limited through society and technology.

Examples:

- 450 BCE: Laws of the Twelve Tables. Written and publicly accessible law
- Science: Lectures
- Bible: Translation 1522 (to understand), Printing 1452 (to spread)
- Technology: Patents and Documentation

Aspects that regulate access:

- Availability (libraries, internet)
- Authorship ("Urheberrecht")
- Copyright ("Verwertungsrecht)
- Personality rights
- Patents

Licenses determine the given rights for published works.

3.1.1 Software:

- Closed Source
  - secret
  - proprietary
  - freeware/shareware
- Restricted Open Source
- Open Source
  - BSD, MIT, Apache 2.0
  - GPL
  - Careful: open != free

more open

3.1.2 Papers:

- Access Restricted ("pay wall"): subscription or payment necessary
- Green Open Access: Authors are allowed to publish their work **on their own homepage** (non-commercial)
- Gold Open Access:
  - Everybody can access the work
  - Different copyright and licensing models:
    - Copyright with author **or** publisher
    - Publisher may hold exclusive rights/license
    - Publisher may hold non-exclusive license; author can share other licenses on their own

Examples for non-exclusive licenses:

- CC0 (often called 'public domain')
- CC BY standard
- CC BY-SA (shared alike)
- CC BY-NC (non-commerical)
- CC BY-ND (no derivatives)

3.2 Publishing Costs

Publishing is expensive. Who pays?

- Depends on license (publisher, author, or community)
- Community: Conference participants, tax payers (in Germany: "Project DEAL")

But why publish with a publisher, and not yourself?

- Peer-reviewing often organized by publisher or conference
- Reputation
- Copy-Editing, rights management, hosting, metadata management, indexing
- Accessibility is guaranteed, easy to find, long-term available
- Added to libraries
- Print versions
- Influence on academic ranking/impact factors
- May be required by funding agency

3.3 Paper Repositories

- Homepage
- ArXiv
- Publishers: Springer Publishing, ACM, IEEE, Elsevier, Wiley, etc. (These organize conference proceedings and journals)
- Popularity and use depends on concrete science.

3.4 Bibliographic Data

- Homepage: Low confidence. May be incomplete or inconsistent
  -> Not suited for citation
- DBLP: Highest confidence. Human-maintained and reliable.
- Google Scholar: Complete, but automatically generated and may be inaccurate (e.g., uses 'et al.').
  But: Nice notifications about new papers that cite own work
- ORCID: Automatically generated, grows in use. Nice notifications about new *own* papers, connected via DOIs.
- CrossRef (papers) and DataCite (data): DOI agencies, manage metadata

3.5 Ideal Requirements for Publishing

Hard requirements:

- long-term available (archived)
- immutable (is always the same version referenced?)
- identifiable (is the correct version referenced?)

Weak requirement: Open Access

3.6 Unique Identification

- DOI: Digital Object Identifier, mostly for papers
  - Resolver: DOI -> URL (e.g., doi.org)
  - points to landing page of publication
  - linked to metadata
  - URLs are ~~ranked~~
- ORCID for scientists