University of Pittsburgh

INFSCI 3005: Introduction to Doctoral Program

Lecture 6: Reference and Search Tools

Peter Brusilovsky, Professor
School of Information Sciences and Intelligent Systems Program, University of Pittsburgh
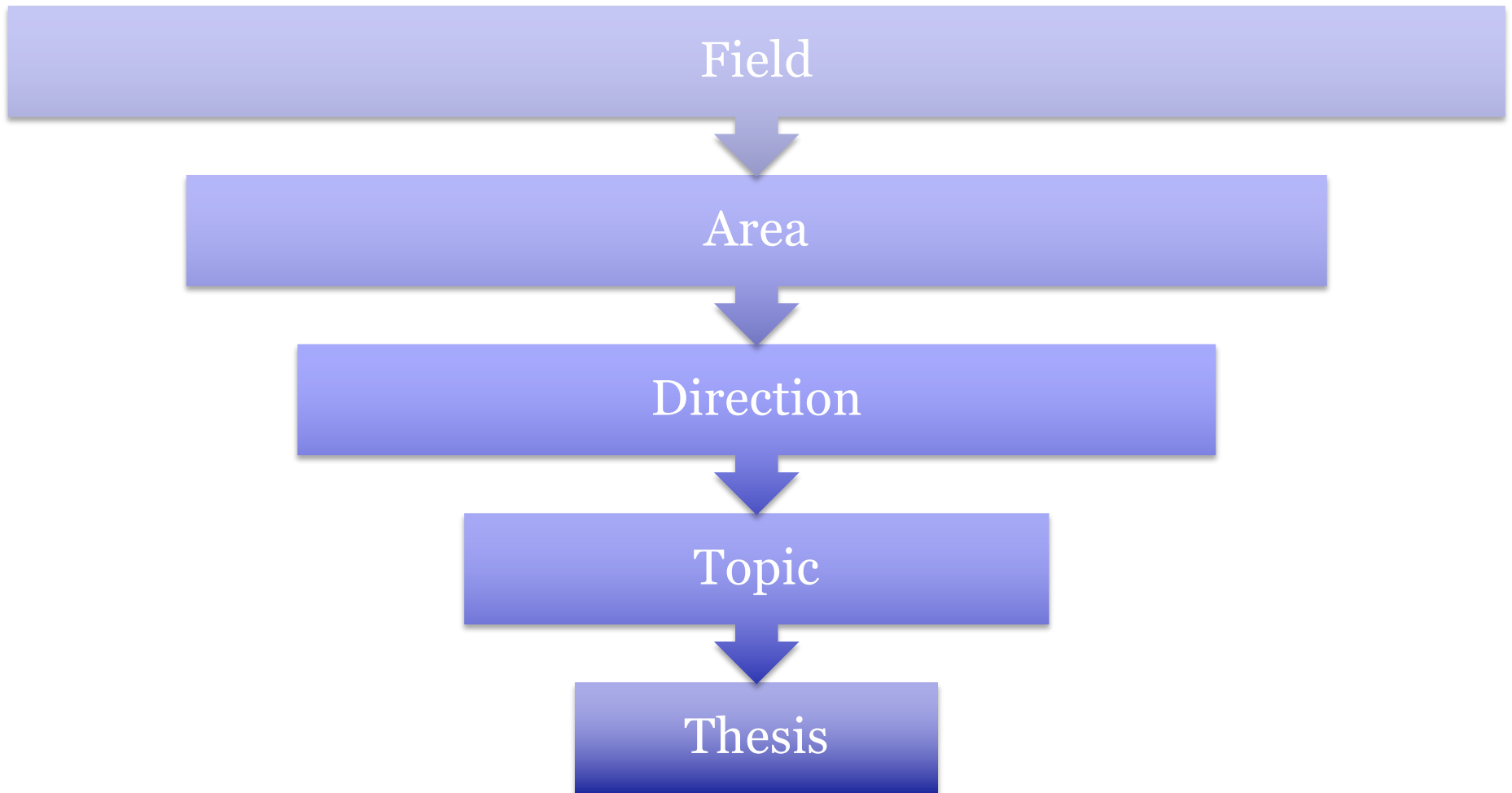
Working with research literature: When?

- To start exploring the field

- To look for research ideas and things to do

- To find similar works for a project or a paper

- To prepare review part of your thesis

- To choose next research goals

Stages of the Doctoral Study

Field

Area

Direction

Topic

Thesis

Working with research literature: Needs

Find

Assess

Get

Organize

Share

Use

Working with research literature: Needs

- Find introductory books, textbooks

- Find research papers on a specific subject

- Find papers of a specific author or team

- Evaluate an impact of authors / research papers

- Organize your [pdf] paper collection

- Organize your collection of references

- Manage citations when writing a paper (thesis)

- Share references with co-authors and groups

Working with research literature: Tools

- Libraries

- Digital libraries and publisher's portals

- Search tools

- Social research portals

- Citation bookmarking and sharing systems

- Citation and impact services

- Citation management tools

- Paper organization tools

The Library

- One of the most valuable local resources
  - Main Pitt Library
  - Carnegie Library
- Interlibrary loan
- Reserves
- Local delivery
- Search tools:
  - http://pittcat.pitt.edu/
  - https://librarycatalog.einetwork.net/

Fee-Based Digital Libraries

- Digital Portals of professional societies
  - ACM: http://portal.acm.org/
  - IEEE CS: http://www.computer.org/portal/
  - IEEE: http://ieeexplore.ieee.org/
- Digital content from leading publishers
  - Springer: http://www.springerlink.com/
  - Taylor and Francis: http://www.tandfonline.com/
  - Elsevier: http://www.sciencedirect.com/
- Free access from Pitt and remotely with VPN
  - http://www.library.pitt.edu/offcampus
  - http://technology.pitt.edu/service/secure-remote-access-connect-with-the-pulse-secure-client

Open-Access Digital Libraries

- National Initiatives
  - https://www.ncbi.nlm.nih.gov/pubmed/
- Author's institutional repository
  - https://d-scholarship.pitt.edu/
- Centralized Research Repositories
  - Contribute (ResearchGate)
    - https://www.researchgate.net/home
  - Automatic (CiteSeerX)
    - http://citeseerx.ist.psu.edu/
- Author's own home pages

Global Search Tools

- Google Search and Google Scholar
  - https://scholar.google.com/
- New Microsoft Academic Search
  - https://academic.microsoft.com/
- DBLP: http://www.informatik.uni-trier.de/~ley/db/
- Semantic Scholar
  - https://www.semanticscholar.org/
- ArnetMiner
  - https://aminer.org/
- ACM DL
  - ACM: https://dl.acm.org/
- Most search tools allow post-search navigation
  - Authors, sources, series, co-authors

Social Research Portals

- Research Gate
  - https://www.researchgate.net/home
- Academia.edu
  - https://www.academia.edu/
- More than access to sources
  - Social connections (follow!)
  - Topics
  - Projects
  - Authority

Citation Bookmarking and Sharing

- Allows you to collect and share *citations* online
  - Citation scrapping when adding to database
  - Tagging for citation organization
  - Import/Export to your collection (EN & BibTeX)
  - Groups for collaborative work
  - Finding papers from people with similar interests (who bookmarked what)
- Main services

  - BibSonomy: http://www.bibsonomy.org/
  - Mendeley: https://www.mendeley.com/library/

Online Social Citation Services

- Social Tagging systems for Citations
  - Bibsonomy: http://www.bibsonomy.org/
  - Manage and share
  - Explore:
    - Open!
    - Power of tags!
  - Import in EndNote, Bibtex and other formats
- Challenges of social services

  - CiteULike: https://en.wikipedia.org/wiki/CiteULike
- Mendeley: http://www.mendeley.com/

  - Bringing it all together
  - Managing paper sources, bookmarking, collaboration, paper writing, integration with Web site
  - Less attention to tags now, lots of closed information

Citation Indexing and Impact

- History of Abstracting/Citation indexing
- Citation tracing: Who cites whom?
- How many citations for a paper, person, journal?
- ISI Web of Science (now Thomson Reuters)
  - http://webofknowledge.com/
  - Maintains Science/Social Sciences/Humanities Citation Index
  - Allow to search for both sources and citations
  - Calculates journal *Impact Factor*
  - Requires subscription
- Elsevier Scopus: http://info.scopus.com/

Citation Tracing in Global Portals

- Google Scholar
  - https://scholar.google.com/
- Old Microsoft Academic Search
  - http://academic.research.microsoft.com/default.aspx
- New Microsoft Academic Search
  - https://academic.microsoft.com/
- Semantic Scholar
  - https://www.semanticscholar.org/
- ArnetMiner
  - https://aminer.org/
- ACM Portal
  - ACM: http://portal.acm.org/
- CiteSeerX: http://citeseerx.ist.psu.edu/

Multiple Ways to find Good Papers

- Classic search
- Identifying and following key authors
  - Impact tools and social following
- Examining top papers by topics
  - Online research portals
- Citation tracing – who cites relevant papers, what these papers cite
  - Portals with citation tracing
- Bookmark tracing – who bookmarks similar papers, what they bookmark
- Citation recommenders!
  - Google Scholar, Mendeley, ResearchGate

Multiple Ways to Get Papers

- No subscription at Pitt, problems with VPN?
- Google Scholar
  - https://scholar.google.com/
- Social Research portals
- Author's home page
  - http://www.pitt.edu/~peterb/papers.html
- Author's institutional repository
  - https://d-scholarship.pitt.edu/
- CiteSeerX
  - http://citeseerx.ist.psu.edu/

Managing Your Stuff: Needs

- You accumulate things over the years, you need to organize and use it

- Managing source papers
  - Why? Not available online, ease of sharing, ease of searching, your notes
  - Add notes to the paper, comment, highlight
  - Find easily the right paper on your drive, search...
  - Lower end – collection of PDFs organized in folders
  - Upper end – dedicated app (Mendeley, Papers for Mac)

- Managing citations
  - Why? Collect, organize, classify, share, use in writing
  - Lower end – bag of references (BibTex file)
  - Upper end – dedicated system with advanced search, tagging, and integrated use in writing (personal or social)

Managing Your Stuff: Tools

- Personal collections of references
  - BibTex files
    - Universal format for citation management using LaTeX
  - EndNote: http://www.endnote.com/
    - Citation management, paper processing, some collection management tools, Pitt has license
  - Also ProCite and other tools
- Online reference management
  - Bibsonomy (social!) http://www.bibsonomy.org/
  - RefWorks: http://www.refworks.com/
  - EndNote Web: http://www.endnoteweb.com/
- Integrated paper and reference management
  - Mendeley, Papers (https://www.papersapp.com/)
- Zotero: http://www.zotero.org/
  - Free extension that can be downloaded and installed in a browser

How to Assess Importance/Ranking

- Field Ranking
  - Scholar, Old Microsoft Academic Search
- H-Index
  - https://en.wikipedia.org/wiki/H-index
  - Field dependent
  - http://web.cs.ucla.edu/~palsberg/h-number.html
- How to find H-Index
  - Google Scholar
  - Various tools – better search!

Unique Identity: Papers

- One of the key problem – accessing, counting impact, citing

- DOI Infrastructure
  - DOI = Unique ID of a paper
  - Managed by CrossRef https://en.wikipedia.org/wiki/Crossref
  - Accessed through http://doi.org/

- Benefits
  - Your paper will be ingested by most search engines, portals, and digital libraries (including Pitt!)
  - Will be counted for impact everywhere
  - Readers could access your paper with one-click
  - Publish in sources, which provides DOI!

Unique Identity: Authors

- A harder problem – resolving authors identity to credit papers and citations
    - Many authors with the same name or initials (citations!)
    - Several ways to cite the same author
    - Less interests from publishers, more by global portals
- Social management of identities

    - Create your profile in Google Scholar, Aminer, RG, ...
    - Claim your papers!
- New Orcid framework - http://orcid.org/

    - Create your ID!
    - http://www.library.pitt.edu/orcid
    - http://orcid.org/0000-0002-1902-1464

Quick Summary: What to Use?

- **Finding and Getting Books**
  - Library, Google, Amazon (used books!)
- **Finding papers**
  - Scholar, ACM DL, Semantic Scholar
- **Getting papers**
  - ACM DL, IEEE, Publishers, CiteSeerX, Scholar, Social portals, authors' home pages, e-mail
- **Sharing papers and references**
  - Mendeley, BibSonomy, ResearchGate, D-scholarship, home page
- **Investigating impact**
  - Google Scholar, ISI, Scopus, other portals
- **Managing your citations and paper processing**
  - Focus on tools that support several needs!
  - BibSonomy with EndNote or BibTex, Mendeley, Zotero