Research Methodology

CS 5001

Instructor: Dr. Ramoza Ahsan

Lecture 3

Agenda

- Research Papers
- Anatomy of Research Paper
- Sample Research Papers
- Project Details
- Sample Survey Papers
- Questions

Difference between research synopsis and research proposal

- A "research synopsis" is a concise summary of a research project, highlighting the key points like objectives, methods, and rationale, while a "research proposal" is a detailed document outlining the entire research plan, including a thorough literature review, methodology, data analysis plan, and justification for the study
- A synopsis is a brief overview of the research idea, whereas a proposal is a comprehensive plan to

Key differences

Length:

- A research synopsis is much shorter and more succinct, often just a few pages long, while a research proposal is typically more extensive, spanning multiple pages depending on the project's complexity.

Level of detail:

- A synopsis provides a high-level overview of the research topic, objectives, and methods, whereas a proposal delves into specific details like sampling strategies, data collection methods, and analysis procedures.

Purpose:

- A synopsis might be used to give a quick overview of a research project to potential supervisors or reviewers, while a research proposal is primarily used to request funding or approval to conduct a study.

Reading and understanding research papers

- Why to read papers
  - To review papers for a conference or a class
  - To keep current in their field
  - Literature survey of a new field
- Sometime, a new researcher wastes much effort in the process and is frequently driven to frustration.

Three-Pass Approach to Read Papers

- Each pass accomplishes specific goals and builds up the previous pass:
  - First gives you a general idea about the paper
  - The second pass lets you grasp the paper's content, but not its details, and
  - The third pass helps you understand the paper in depth

The First Pass

This pass should take about five to ten minutes with the following steps

- Carefully read the title, and the abstract
- Read the section and sub-section headings, but ignore everything else
- Read the conclusion

The Second Pass

Read the paper with greater care, ignore details such as proofs, with the following steps

- Read Introduction – which is itself a short version of an article
- Concentrate on figures, diagrams and other illustrations.
- Remember to mark relevant unread references for further reading (background).
- The second pass should take up to an hour. After this pass, you should be able to grasp the content of the paper. You should be able to summarize the main thrust of the paper, with supporting evidence, to someone else.

The Third Pass

- To fully understand a paper, particularly if you are reviewer, requires a third pass.
  - The key to the third pass is to attempt to virtually re-implement the paper: that is, making the same assumptions as the authors, re-create the work.
  - By comparing this re-creation with the actual paper, you can easily identify not only a paper's innovations, but also its hidden failings and assumptions.
- This pass requires great attention to detail.
  - Read the complete contents
  - Understand the methodology
  - Identify and challenge every assumption in every statement.
  - Think about how you yourself would present this particular idea.
  - Collect the dataset used in the paper and plan about dataset, if existing is not publicly available

Selecting a Base Paper for the Proposed Research – A Tip

- One of the well practiced approach
- Choose one best paper in your research area
  - Understand it completely; read many times
  - Get presentation slides from authors
  - Implement it
    - If original code available then use it to clear some of the concepts, which you don't understand from paper

Research Literature

- A Detailed Analysis

Listing of Premier Journals/Conferences to be used for Reading purpose

- Easy to use formula
- Type
  - “list of top journals in ____” in Search Engine
- Or
  - “list of top conferences in ____” in Search Engine

Research Literature

- Include literature survey that has a significant connection to your work and interest
  - It establishes that your work is indeed novel or innovative
  - Helps you to understand current theory, discoveries, and debates
  - Helps identify newlines of questioning or investigation

Finding Research Papers

- The number of papers published in major computer science venues each year is at least tens of thousands.
  - DBLP Computer Science Bibliography
- **Way-out**: benefit from knowledge (publications) of other credible authors
  - In any active field of research, other researchers have, to a certain extent, already explored and digested the older literature
- However, ***reading about a paper*** that seems relevant is never a substitute for ***reading the paper itself***. If you need to discuss or cite a paper, read it first.

Finding Research Papers

- Sometime, there are several versions of the same paper: a preprint in an online archive, a conference version, and a journal version
- Finding *all* relevant work is hard, But finding all *significant* work is a critical part of doing research
- Searching and reading
  - Separate activities or operate in parallel ?

Finding Research Papers

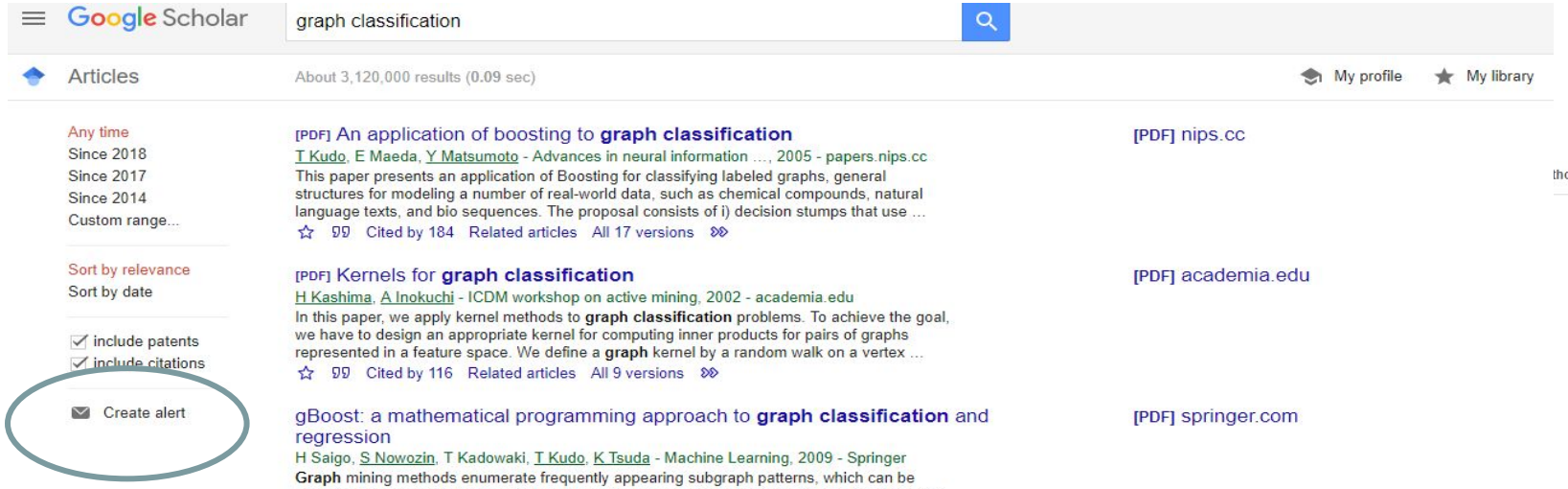
- Having explored the literature, you may discover that your original idea is not so original after all
  - If so, be honest—review your work to see what aspects may be novel, but don't fool yourself into working on a problem that is already solved

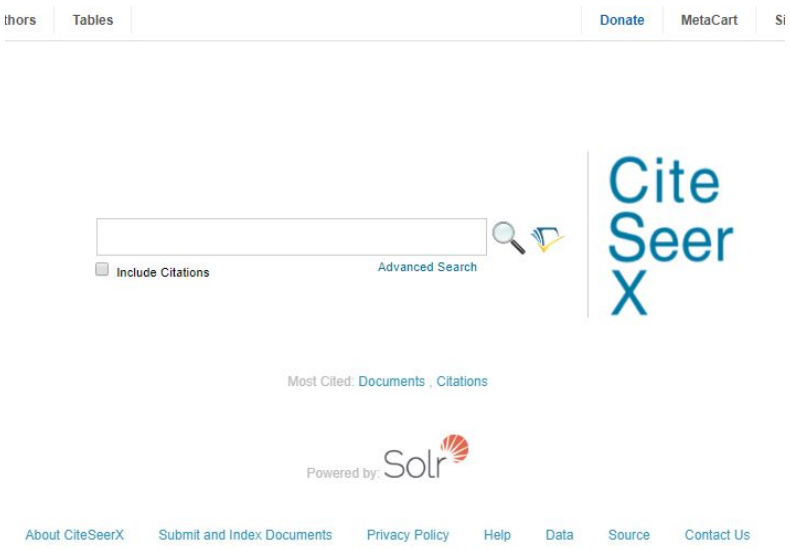
Finding Research Papers – Tools

1. Use obvious search terms/keywords to explore the Web
   - Research Papers
   - Home pages for authors/projects
   - Research groups concerned with the same research area

Finding Research Papers – Tools

2. Some major search engines have search tools that are specifically for academic papers





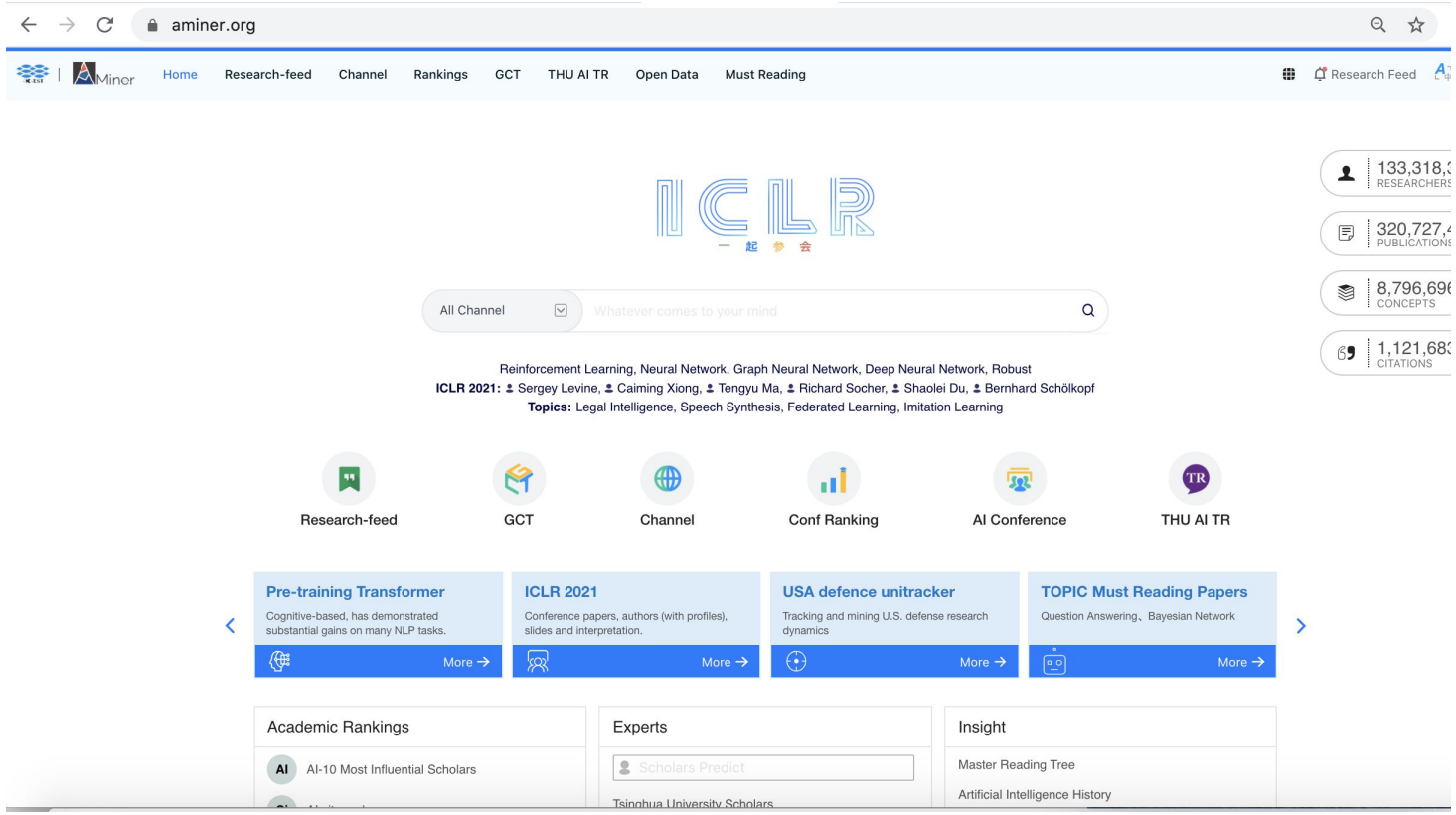
Finding Research Papers – Tools

3. Visit the websites of research groups and researchers working in the area to find out
   - the names of researchers whose work you should investigate
   - the names of their co-authors
   - conferences where relevant work appears
   - papers with lists of references to explore

Finding Research Papers – Tools

4. ArnetMiner

- (also AMiner) is a free online service used to index, search, and mine big scientific data.
- a platform to search for
  - Experts
  - Papers



Finding Research Papers – Tools

5. Follow up the references in promising research papers
   - These indicate relevant individuals, conferences, and journals
   - Forward and backward citations

A parallel graph partitioning algorithm to speed up the large-scale distributed graph mining

ZengFeng Zeng
School of Computer Science
Beijing University of Posts and Telecommunications
Beijing 100876, China
zzfeng1987@gmail.com

Bin Wu
School of Computer Science
Beijing University of Posts and Telecommunications
Beijing 100876, China
wubin@bupt.edu.cn

Haoyu Wang
School of Computer Science
Beijing University of Posts and Telecommunications
Beijing 100876, China
muyiwhy0523@gmail.com

ABSTRACT

For the large-scale distributed graph mining, the graph is distributed over a cluster of nodes, thus performing computations on the distributed graph is expensive when large amount of data have to be moved between different computers. A good partitioning of distributed graph is needed to reduce the communication between computers and scale a system up. Existing graph partitioning algorithms incur high computation and communication cost when applied on large distributed graphs. A efficient and scalable partition-

1. INTRODUCTION

1.1 Large-scale graph mining

Graph datasets we face today are becoming much larg
than before. The modern large search engines crawls mo
than one trillion links in the internet and the social ne
working web site contains more than 800 million active use
[9]. Besides the large graph on Internet and social networ...,
the biological networks which represent protein interactions

Natural Science Foundation of China (Grant No.61074128, 60905025, 90924029).

8. REFERENCES

[1] A. Abou-Rjeili and G. Karypis. Multilevel algorithms for partitioning power-law graphs. *IPDPS'2006*, 2006.

[2] M. Bailly-Bechet, C. Borgs, A. Braunstein, J. T. Chayes, A. Dagkessamanskaia, J.-M. Franfois, and R. Zecchina. Finding undetected protein associations in cell signaling by belief propagation. *CoRR*, 2011.

[3] A. Barabasi and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509, 1999.

[4] S. Barnard. Pmrsb: Parallel multilevel recursive spectral bisection. In *Supercomputing, 1995. Proceedings of the IEEE/ACM SC95 Conference*, page 27, 1995.

[5] J. C.Chu, J.A.George and E.G.Ng. User'sguide for sparspak-a: Waterloo sparse linear equations package. Technical report, Computer Science, University of Waterlo o,Ontario., 1984.

[6] A. Chan, F. K. H. A. Dehne, and R. Taylor. Cgmgraph/cgmlib: Implementing and testing cgm graph algorithms on pc clusters and shared memory machines. *IJHPCA*, 19(1):81–97, 2005.

[7] W. D. and W. F. Between mincut and graph bisection. In *In Proceedings of the 18th International Symposium on Mathematical Foundations of Computer science (MFCS)*, pages 744 – 750, 1993.

[8] J. Dean and S. Ghemawat. Mapreduce: simplified

Technical Journal, 49(2):291–307, 1970.

[18] J. Leskovec. Snap, 2011. http://snap.stanford.edu/snap.

[19] G. Malewicz, M. H. Austern, A. J. C. Bik, J. C. Dehnert, I. Horn, N. Leiser, and G. Czajkowski. Pregel: a system for large-scale graph processing. In *SIGMOD Conference'10*, pages 135–146, 2010.

[20] G. L. Miller, S.-H. Teng, W. Thurston, and S. A. Vavasis. Automatic mesh partitioning. In A. George, J. Gilbert, and J. Liu, editors, *Graphs Theory and Sparse Matrix Computation*, The IMA Volumes in Mathematics and its Application, pages 57–84. Springer-Verlag, 1993. Vol 56.

[21] U. N. Raghavan, R. Albert, and S. Kumara. Near linear time algorithm to detect community structures in large-scale networks, Sept. 2007.

[22] A. H. D. Schult and P. Swart. Networkx. http://networkx.lanl.gov/.

[23] I. Stanton and G. Kliot. Streaming graph partitioning for large distributed graphs, 2011. http://research.microsoft.com/apps/pubs/default.aspx?id=

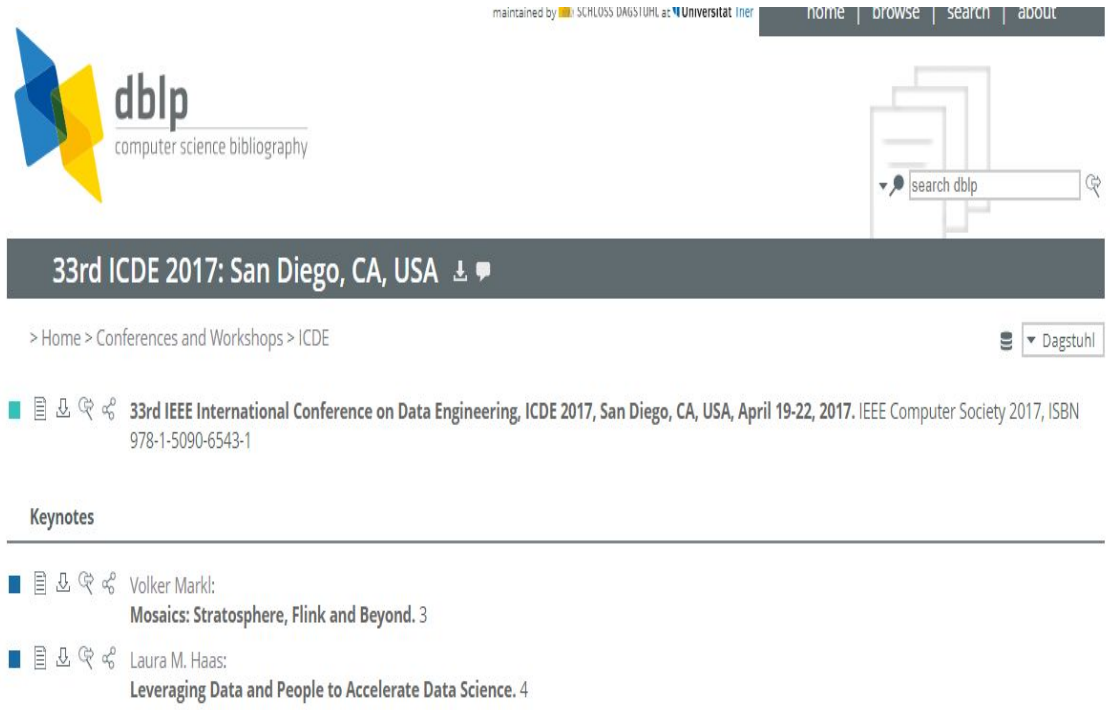
[24] X. Sui, D. Nguyen, M. Burtscher, and K. Pingali. Parallel graph partitioning on multicore architectures. In *LCPC'10*, pages 246–260, 2010.

[25] C. Walshaw. Graph partitioning archive. http://staffweb.cms.gre.ac.uk/ wc06/partition.

[26] D. J. Watts and S. H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393(6684):440–442, June 1998.

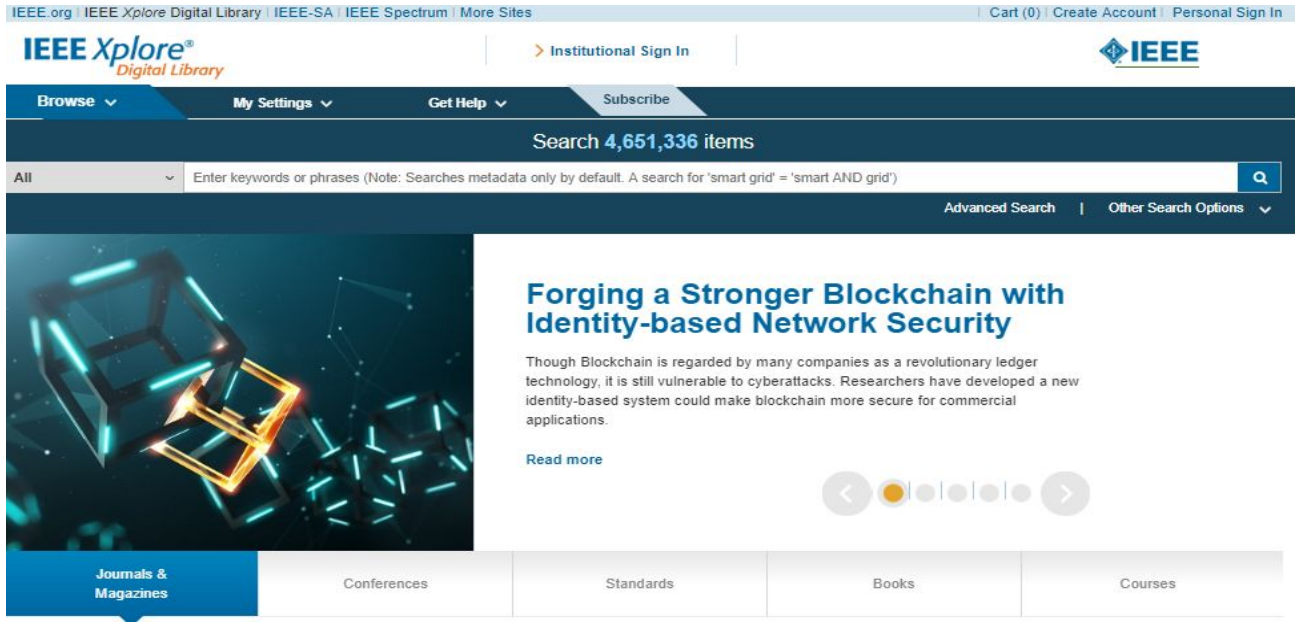
Finding Research Papers – Tools

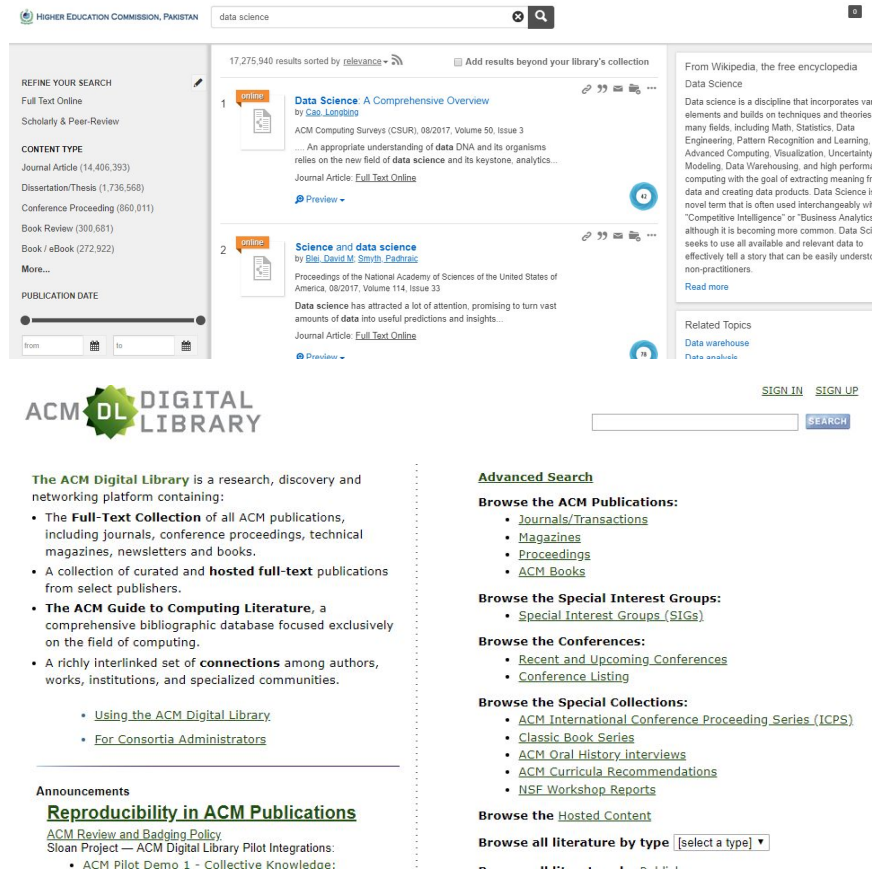
6. Browse the recent issues of the journals and conferences in the area



Finding Research Papers – Tools

7. Search the publisher-specific digital libraries
   - IEEE, ACM, Springer, Elsevier, HEC

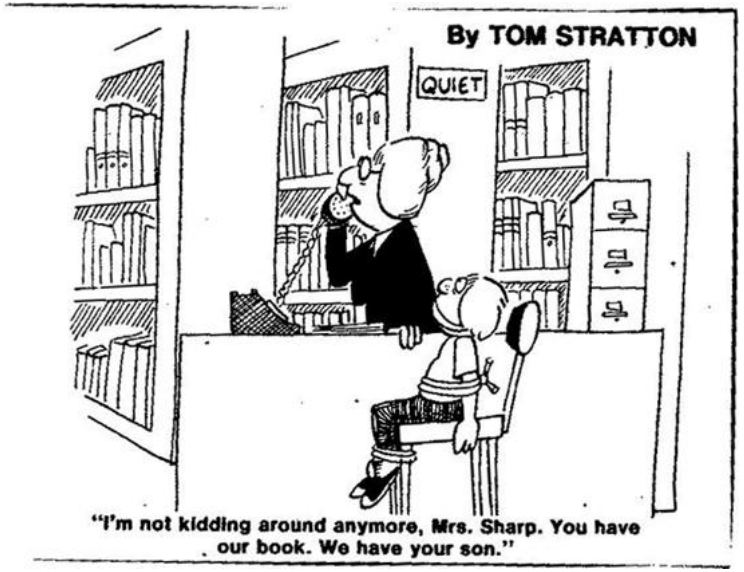




Finding Research Papers – Tools

8. Visit Library
   - having similar material shelved together often leads to unexpected discoveries
   - without the distractions that arise when browsing the Web

"I love to visit the library … perfect place to text."



"I'm not kidding around anymore, Mrs. Sharp. You have our book. We have your son."

Finding Research Papers – Tools

9. Discuss your work with as many people as possible
   - Establish research collaborations
     - Cross-domain research has high impact
   - “working in silos”—mean that people investigating similar problems can be unaware of one another.

How to Write a Research Paper

Getting Ready With Data

- Gather all important data, analyses, plots and tables
- Organize results so that they follow a logical sequence
- Consolidate data plots and create figures for the manuscript

First Draft

- Identify two or three important findings emerging from the experiments. Make them the central theme of the article.

- Note the readership of the journal that you are considering to publish your work.

- Prepare figures, schemes and tables in a professional manner

Structure of a Scientific Paper

- Title
- Abstract
- Introduction
- Related Work
- Methodology
- Experimental Section
- Results and Discussion
- Conclusions
- Acknowledgments
- References
- Supporting Information

Title

- As you craft a name for your paper, you should consider these potential objectives for the title you choose. A title should:

- Describe the content of the paper
- Distinguish the paper from others on a similar topic
- Catch the reader's attention and interest
- Match search queries so people will find your paper (and cite it)

Title Examples

- Example: Suppose you're a robotics researcher, and you've discovered that probabilistic path finding is far superior to earlier methods that depended on deterministic methods. You might choose a title like

- "Probabilistic Pathfinding: Beyond Deterministic Methods for Navigation in Rough Terrain"

Title Examples

- It’s also useful to create a title that sticks with people. You might consider a few devices to help you there. How about a snappy name or acronym for your approach?

- “GRAMMPS: A generalized mission planner for multiple mobile robots in unstructured environments”

Abstract

- An abstract, or summary, is published together with a research article, giving the reader a "preview" of what's to come. Such abstracts may also be published separately in bibliographical sources, such as Biologic al Abstracts.
- They allow other scientists to quickly scan the large scientific literature, and decide which articles they want to read in depth. The abstract should be a little less technical than the article itself; you don't want to dissuade your potential audience from reading your paper.
- Your abstract should be one paragraph, of 100-250 words, which summarizes the purpose, methods, results and conclusions of the paper.

Abstract

- It is not easy to include all this information in just a few words. Start by writing a summary that includes whatever you think is important, and then gradually prune it down to size by removing unnecessary words, while still retaining the necessary concepts.

- Don't use abbreviations or citations in the abstract. It should be able to stand alone without any footnotes.

Introduction

- What question did you ask in your experiment? Why is it interesting? The introduction summarizes the relevant literature so that the reader will understand why you were interested in the question you asked. End with a sentence explaining the specific question you asked in this experiment.

MATERIALS AND METHODS

- How did you answer this question? There should be enough information here to allow another scientist to repeat your experiment. Look at other papers that have been published in your field to get some idea of what is included in this section.

- If you had a complicated protocol, it may be helpful to include a diagram, table or flowchart to explain the methods you used.

MATERIALS AND METHODS

- Do not put results in this section. You may, however, include preliminary results that were used to design the main experiment that you are reporting on. ("In a preliminary study, I observed the owls for one week, and found that 73 % of their locomotor activity occurred during the night, and so I conducted all subsequent experiments between 11 pm and 6 am.")

- Mention relevant ethical considerations. If you used human subjects, did they consent to participate. If you used animals, what measures did you take to minimize pain?

RESULTS

- This is where you present the results you've gotten. Use graphs and tables if appropriate but also summarize your main findings in the text. Do NOT discuss the results or speculate as to why something happened;
- You don't necessarily have to include all the data you've gotten during the semester. This isn't a diary.
- Use appropriate methods of showing data. Don't try to manipulate the data to make it look like you did more than you actually did.

DISCUSSION

- Highlight the most significant results, but don't just repeat what you've written in the Results section. How do these results relate to the original question? Do the data support your hypothesis? Are your results consistent with what other investigators have reported? If your results were unexpected, try to explain why. Is there another way to interpret your results? What further research would be necessary to answer the questions raised by your results? How do your results fit into the big picture?

Conclusion

- Your conclusion is your chance to have the last word on the subject. The conclusion allows you to have the final say on the issues you have raised in your paper, to summarize your thoughts, to demonstrate the importance of your ideas, and to propel your reader to a new view of the subject. It is also your opportunity to make a good final impression and to end on a positive note.

- Your conclusion can go beyond the confines of the assignment. The conclusion pushes beyond the boundaries of the prompt and allows you to consider broader issues, make new connections, and elaborate on the significance of your findings.

Conclusion

- Your conclusion should make your readers glad they read your paper. Your conclusion gives your reader something to take away that will help them see things differently or appreciate your topic in personally relevant ways. It can suggest broader implications that will not only interest your reader but also enrich your reader's life in some way. It is your gift to the reader.

References

- It is important to properly and appropriately cite references in scientific research papers in order to acknowledge your sources and give credit where credit is due. Science moves forward only by building upon the work of others. There are, however, other reasons for citing references in scientific research papers. Citations to appropriate sources show that you've done your homework and are aware of the background and context into which your work fits, and they help lend validity to your arguments.
- Reference citations also provide avenues for interested readers to follow up on aspects of your work -- they help weave the web of science. You may wish to include citations for sources that add relevant information to your own work, or that present alternate views.

Thank you

Questions?