
Finding and reading research papers

INFO 4940: Advanced NLP for Humanities Research

Papers are how you know a field

Papers (and, in the humanities, books) are how researchers in a field talk to one another

Every project eventually wants to be a paper

Every paper wants to make a contribution to knowledge

A contribution to knowledge is a finding that is **novel**, **important**, and **evidence-based**

novel
important
evidence

Finding papers

Humanities

These are harder, because more diffuse. Mostly **journals** and **books**. Some standard sources:

- MLA bibliography (literature)
 - Philosopher's Index
 - America History and Life

 - Cornell Library
 - Google Scholar
 - SSRN (social sciences, not peer reviewed)
-

NLP, CL, CS

Mostly conference proceedings.

- **NLP:** ACL, NAACL, EMNLP, EACL (collectively, *ACL)
 - **CL:** COLING
 - **CS:** SIGIR, AAAI, NeurIPS, ICML, ICLR

 - arXiv (not peer reviewed)
 - Twitter, Bluesky, whatever else the kids are using
-

Finding good papers

Anyone *can* write a good paper ...

... but not everyone does. Look for signs of quality and significance:

- In a big journal, press, conference?
 - PI at a good school/company/lab?
 - Who's the PI?
 - What else has this person/group published?
 - Do people cite this person/group?
 - Is the paper itself well written? Good figures? Free of obvious errors of content or formatting?
-

How to *skim* a paper

You will find a lot of papers. It is not profitable to read them all in depth. But how do you find the ones that are worth reading carefully?

- Scan for quality markers
 - Read the abstract
 - If good, read the intro and figures/tables
 - If good, read the whole thing
-

How to *read* a paper

Remember: **novel, important, evidence**

- What is the problem area?
 - What is the state of knowledge?
 - What is the gap in knowledge?
 - What did they expect to find?
 - What did they do?
 - How did they do it? (methods, data)
 - What did they find? (results)
 - What do they conclude? (conclusions)
 - What are the limitations of the study?
 - How do you evaluate the quality of their findings? (strength of evidence)
 - What might you borrow from this work?
-

An annotated paper

Answer all the questions from the last slide.

You should, in effect, write a very short version of the paper itself (about 250 words).

Pay special attention to what in the paper is of value to you.
You will not remember the whole thing later. You want a quick way to remind yourself why the paper is important to you.

An example

Questions to answer

Remember: **novel, important, evidence**

- What is the problem area?
 - What is the state of knowledge?
 - What is the gap in knowledge?
 - What did they expect to find?
 - What did they do?
 - How did they do it? (methods, data)
 - What did they find? (results)
 - What do they conclude? (conclusions)
 - What are the limitations of the study?
 - How do you evaluate the quality of their findings? (strength of evidence)
 - What might you borrow from this work?
-