Unit III: Literature Search, Analysis & Ethics

Dr. Ayush Kumar Agrawal

SoCS
UPES

Introduction to Literature Search in Computer Science Research

Warm-Up Activity

Task:

- Think about your final-year project/research topic.
- Where would you look first for related work? (Google? Wikipedia? IEEE?)
- Share your answer with your neighbor.

Purpose: To highlight how research sources differ from casual searching.

Why Literature Search Matters

- Foundation of every research project.
- Avoids duplication $\rightarrow$ ensures novelty.
- Helps identify benchmarks, tools, datasets.
- Reveals gaps and unresolved questions.
- Builds justification for your contribution.

Sources of Literature in CS

Primary Sources:

- Journals (IEEE Transactions, ACM Surveys).
- Conference papers (NeurIPS, ICSE, SIGCOMM).
- Patents, technical reports.

Secondary Sources:

- Books, review/survey papers.
- Online repositories (arXiv, ResearchGate).
- Databases: IEEE Xplore, ACM DL, SpringerLink.

Mini Case Study

Case: A student wanted to work on "Blockchain for Supply Chain."

- Step 1: He searched Google $\rightarrow$ found blogs, news, random articles.
- Step 2: He searched IEEE Xplore $\rightarrow$ found 1,200+ peer-reviewed articles.
- Step 3: He realized blogs lacked depth; journals revealed real gaps (scalability, energy use).

Lesson: Reliable databases = reliable research foundation.

Objectives of a Literature Review

- Summarize current knowledge in the domain.
- Compare methods and results across studies.
- Identify gaps and potential research directions.
- Highlight contradictions or debates in findings.
- Form the theoretical and methodological base for your study.

Process of Literature Search

1. Define your research question/problem.
2. Choose suitable databases.
3. Identify keywords & synonyms.
4. Use Boolean operators (AND, OR, NOT).
5. Screen abstracts before reading full texts.
6. Record references systematically.
7. Group insights into categories/themes.

Quick Interactive Quiz

Which of these is a reliable academic source?

1. Wikipedia article on “Deep Learning”
2. Research paper in IEEE Transactions
3. Blog on Medium.com
4. YouTube tutorial on CNNs

Quick Interactive Quiz

Which of these is a reliable academic source?

1. Wikipedia article on "Deep Learning"
2. Research paper in IEEE Transactions
3. Blog on Medium.com
4. YouTube tutorial on CNNs

Answer: Option 2 (others may be helpful for learning, but not valid research sources).

Benefits of an Effective Literature Search

- Saves time $\rightarrow$ focus only on relevant studies.
- Prevents "reinventing the wheel."
- Increases credibility of your proposal/paper.
- Helps track key researchers and trends.
- Forms the backbone of your thesis/dissertation.

Common Mistakes by Students

- Relying only on Google or Wikipedia.
- Reading papers without taking notes.
- Collecting too many papers without analysis.
- Ignoring the “Methods” and “Limitations” sections.
- Not organizing references systematically.

Class Activity: Live Search

Task:

- Take a sample topic: "AI for Healthcare Diagnosis."
- Search IEEE Xplore / Google Scholar.
- Find 1 paper, note down: author, year, main result.
- Share with class (2–3 students).

Purpose: Practice identifying quality sources.

Summary

- Literature search = foundation of all research.
- Sources: journals, conferences, books, digital libraries.
- Objectives: summarize, compare, identify gaps.
- Process: define problem $\rightarrow$ search $\rightarrow$ organize.
- Effective search saves time and improves credibility.

Discussion Prompt

Question: If you were starting research in AI Ethics,

- Which 3 databases would you prioritize?
- What keywords would you use?
- How would you check if the papers are reliable?

Databases & Search Strategies in Computer Science Research

Warm-Up

Task:

- Where do you usually search for information when starting a project?
- Google? Wikipedia? YouTube? Research papers?
- Raise your hand if you've ever used IEEE Xplore or ACM DL.

Purpose: To show the gap between casual search and academic search.

Why Databases?

- Provide access to peer-reviewed, credible research.
- Allow advanced search by author, year, keywords.
- Contain citation links to track references.
- Often provide full-text, downloadable PDFs.
- Ensure research quality vs. random web sources.

Major Databases in CS

Core Databases:

- IEEE Xplore – electronics, CS, AI, IoT, robotics.
- ACM Digital Library – computing, algorithms, HCI.
- SpringerLink, Elsevier (ScienceDirect) – broad coverage.
- Scopus, Web of Science – indexing and citation analysis.
- arXiv – preprints in AI, ML, mathematics.
- Google Scholar – free, wide coverage but less filtered.

Case Study: IEEE vs Google

Topic: "Federated Learning in Healthcare"

- Google search $\rightarrow$ Blogs, YouTube videos, Wikipedia pages.
- IEEE Xplore $\rightarrow$ 500+ peer-reviewed papers with experiments, datasets, results.
- **Lesson:** Only academic databases ensure reliable, citable content.

Keyword Selection

- Use precise and technical terms.
- Include synonyms, acronyms, related words.
- Example Topic: "AI in Cybersecurity"
  - Keywords: Artificial Intelligence, Machine Learning, Intrusion Detection, Malware Detection, Deep Learning.
- Tip: Check keywords used in top papers in your field.

Boolean Operators & Advanced Search

- **AND** $\rightarrow$ narrows search (AI AND Cybersecurity).
- **OR** $\rightarrow$ broadens search (IoT OR "Internet of Things").
- **NOT** $\rightarrow$ excludes terms (Cloud NOT Gaming).
- **" "** $\rightarrow$ exact phrase ("deep reinforcement learning").
- ***** $\rightarrow$ wildcard (comput* $\rightarrow$ computer, computing, computation).

Hands-On Activity

Task:

- Topic: "Drone Applications in Agriculture."
- Step 1: Write 3–4 keywords.
- Step 2: Combine them using Boolean operators.
- Step 3: Try searching in Google Scholar / IEEE Xplore.

Share 1 interesting paper title you find.

Citation Chaining

- **Backward chaining:** Look at references in a paper.
- **Forward chaining:** Look at newer papers that cited it.
- Helps discover related and influential research quickly.
- Example: A 2015 paper on "Generative Adversarial Networks" $\rightarrow$ thousands of follow-up papers.

Search Alerts & Profiles

- Many databases allow email alerts for new papers matching keywords.
- Researchers can create profiles (Google Scholar, ResearchGate).
- Alerts save time by automatically delivering new relevant papers.
- Example: Set an alert for "IoT security in smart grids."

Quick Quiz

Which search is better?

1. "AI in Healthcare"
2. ("Artificial Intelligence" OR "Machine Learning") AND (Healthcare OR "Medical Diagnosis")

Quick Quiz

Which search is better?

1. "AI in Healthcare"
2. ("Artificial Intelligence" OR "Machine Learning") AND (Healthcare OR "Medical Diagnosis")

Answer: Option 2 $\rightarrow$ more precise and comprehensive.

Common Mistakes in Search

- Using only 1–2 keywords $\rightarrow$ missing relevant work.
- Relying only on Google search.
- Ignoring Boolean operators.
- Not keeping a systematic record of searches.
- Skipping abstracts and downloading too many irrelevant papers.

Summary

- Databases ensure access to credible, peer-reviewed research.
- IEEE, ACM, Scopus, Springer, arXiv are essential in CS.
- Smart keywording and Boolean logic improve search quality.
- Citation chaining and alerts save time and reveal trends.
- An effective search strategy = foundation of strong research.

Literature Analysis & Synthesis in Computer Science Research

Warm-Up

Task:

- Pick one paper you recently read.
- How did you decide whether it was useful for your topic?
- What sections did you focus on most? (Abstract, Method, Results, Conclusion)

Purpose: To realize critical reading differs from casual reading.

Why Literature Analysis?

- Not enough to just collect papers $\rightarrow$ must evaluate them.
- Helps separate high-quality research from weak studies.
- Reveals methodologies, strengths, and limitations.
- Guides researcher in selecting suitable approaches.
- Prevents "data dumping" in the literature review.

How to Critically Read a Paper

- Start with Abstract $\rightarrow$ relevance check.
- Introduction $\rightarrow$ motivation, research questions.
- Methodology $\rightarrow$ algorithms, datasets, experimental design.
- Results $\rightarrow$ performance, validity of conclusions.
- Discussion/Limitations $\rightarrow$ what is missing, scope for improvement.
- References $\rightarrow$ backward chaining.

Extracting Key Elements

While analyzing, record:

- Problem addressed (research question).
- Method/approach used.
- Dataset/experimental setup.
- Key findings (results, accuracy, performance metrics).
- Limitations acknowledged by authors.
- Opportunities for further research.

Mini Case Study

Topic: Intrusion Detection using ML

- Paper A: SVM model with 92% accuracy, dataset = KDD Cup 99.
- Paper B: Deep Learning with 96% accuracy, dataset = NSL-KDD.
- Paper C: Hybrid model, high accuracy but computationally expensive.

Analysis: - SVM: Simple, good baseline. - DL: More accurate, modern dataset. - Hybrid: Best accuracy but not scalable.

From Analysis to Synthesis

- **Analysis** = Breaking down each study individually.
- **Synthesis** = Integrating findings across studies.
- Move from "what each paper says" $\rightarrow$ "what the field says."
- Build themes, clusters, and trends.

Example: Many studies use ML for intrusion detection, but fewer address real-time deployment $\rightarrow$ research gap.

Techniques for Literature Synthesis

- **Thematic Synthesis:** Group studies by common themes.
- **Methodological Comparison:** Compare algorithms, datasets.
- **Chronological Review:** Show progress over time.
- **Conceptual Mapping:** Create diagrams linking key ideas.
- **Systematic Review:** Use PRISMA or similar frameworks.

Class Activity

Task:

- Instructor provides 3 abstracts on a topic (e.g., IoT security).
- Students identify: - Research question. - Method/approach. - Key finding.
- Then compare across abstracts $\rightarrow$ find a gap or difference.

Identifying Research Gaps

- Look for contradictions between studies.
- Check if datasets are outdated.
- See if real-world validation is missing.
- Notice if scalability, ethics, or cost issues are ignored.
- Gaps = opportunities for new contributions.

Quick Quiz

Which of these represents **synthesis** rather than analysis?

1. "Paper A uses CNN, Paper B uses SVM, Paper C uses LSTM."
2. "Across studies, deep learning methods outperform traditional ML in image recognition tasks."

Quick Quiz

Which of these represents **synthesis** rather than analysis?

1. "Paper A uses CNN, Paper B uses SVM, Paper C uses LSTM."
2. "Across studies, deep learning methods outperform traditional ML in image recognition tasks."

Answer: Option 2 $\rightarrow$ integrates multiple findings.

Common Mistakes in Literature Analysis

- Collecting too many papers without evaluating.
- Summarizing but not comparing.
- Ignoring limitations mentioned by authors.
- Treating every paper as equally credible.
- Failing to connect insights into a bigger picture.

Summary

- Analysis = evaluate each study; Synthesis = integrate across studies.
- Focus on methodology, datasets, findings, and limitations.
- Use thematic, chronological, or conceptual synthesis.
- Research gaps emerge from contradictions, outdated methods, or missing aspects.
- Strong synthesis builds the foundation for your research problem.

Mapping the Literature in Computer Science Research

Warm-Up

Question:

- Suppose you collected 30 research papers on "AI in Healthcare."
- How will you organize them so that you can make sense of trends?
- By topic? By year? By method?

Purpose: To show why mapping literature is essential.

What is Literature Mapping?

- Process of visually or systematically organizing research papers.
- Helps identify clusters, themes, and connections between studies.
- Makes it easier to see gaps and research directions.
- Moves beyond summaries $\rightarrow$ provides a big-picture view.

Why Map Literature?

- Helps avoid information overload when dealing with many papers.
- Highlights influential authors and frequently cited works.
- Shows evolution of topics over time.
- Provides evidence for choosing your own research direction.
- Useful for writing the literature review section of a thesis.

Methods of Mapping

- **Thematic Mapping:** Group papers by recurring themes.
- **Chronological Mapping:** Track progress across years.
- **Methodological Mapping:** Compare approaches used (e.g., ML vs DL).
- **Citation Mapping:** Track influential papers and citation networks.
- **Conceptual Mapping:** Use diagrams/mind maps to connect ideas.

Case Example

Topic: "IoT Security"

- Early works (2010–2014): Basic encryption methods.
- Mid works (2015–2018): Intrusion detection with ML.
- Recent works (2019–2023): Blockchain, federated learning.

Mapping Result: Trend shift from classical security $\rightarrow$ AI-based $\rightarrow$ decentralized approaches.

Tools for Literature Mapping

- **VOSviewer:** Builds co-citation and keyword maps.
- **CiteSpace:** Visualizes citation networks and trends.
- **Connected Papers:** Free tool to explore related works.
- **Mind Mapping Tools:** XMind, MindMeister for thematic clusters.
- **Spreadsheets:** Manual mapping for smaller studies.

Hands-On Demo (Instructor-Led)

Task:

- Open VOSviewer with a sample Scopus/IEEE dataset.
- Generate a keyword co-occurrence map.
- Show clusters of terms (e.g., AI, Deep Learning, Image Recognition).

Students observe how topics cluster and evolve.

Class Activity

Task:

- Divide into groups of 3–4.
- Topic: "Cloud Computing."
- Each group lists 3 themes from 5 sample papers (e.g., cost efficiency, scalability, energy usage).
- Share themes on the board $\rightarrow$ class builds a thematic map together.

Quick Quiz

Which of these is an example of **citation mapping**?

1. Grouping papers on AI ethics by sub-themes.
2. Showing which later papers cited a famous 2014 GAN paper.
3. Listing research methods (SVM, CNN, LSTM).

Quick Quiz

Which of these is an example of **citation mapping**?

1. Grouping papers on AI ethics by sub-themes.
2. Showing which later papers cited a famous 2014 GAN paper.
3. Listing research methods (SVM, CNN, LSTM).

Answer: Option 2.

Common Mistakes in Literature Mapping

- Collecting papers without categorization.
- Over-focusing on quantity instead of quality.
- Ignoring citation impact of papers.
- Mapping only keywords, not ideas.
- Failing to update the map with recent work.

Best Practices

- Start mapping early in your research.
- Use both manual notes + digital tools.
- Update maps every few months with new publications.
- Focus on influential works, not just the latest ones.
- Use mapping visuals in your thesis/paper review sections.

Summary

- Literature mapping organizes papers into meaningful structures.
- Approaches: thematic, chronological, methodological, citation, conceptual.
- Tools: VOSviewer, CiteSpace, Connected Papers, mind maps.
- Mapping helps identify trends, clusters, and gaps in research.
- A well-built map strengthens your literature review chapter.