

Conducting a Literature Survey

Prof. Pranay Kumar Saha

Contents

1	Introduction to Literature Survey	2
1.1	What is a Literature Survey?	2
1.2	Why is a Literature Survey Important?	2
1.3	Examples in Computer Science	2
2	Steps to Conduct a Literature Survey	2
2.1	Step 1: Clarify the Research Problem	2
2.2	Step 2: Identify Reliable Sources	2
2.3	Step 3: Formulate Keywords and Search Strategies	2
2.4	Step 4: Analyze and Categorize Sources	3
2.5	Step 5: Summarize and Evaluate	3
3	Tools and Techniques	3
3.1	Academic Search Engines	3
3.2	Reference Management Tools	3
3.3	Visualization Tools	3
4	Evaluating Quality and Credibility of Sources	4
4.1	Peer-Reviewed vs. Non-Peer-Reviewed	4
4.2	Indicators of Quality	4
4.3	Avoiding Predatory Journals	4
4.4	Example Activity	4
5	Structuring the Literature Survey	4
5.1	Common Structures	4
5.2	Writing Tips	4
5.3	Visualization Techniques	5
6	Case Study: Literature Survey on a Sample Problem	5
6.1	Problem: Improving Recommendation Systems with Machine Learning	5
6.2	Activity	5

1 Introduction to Literature Survey

1.1 What is a Literature Survey?

A literature survey is a comprehensive review of existing research on a specific topic. It involves systematically searching, analyzing, and synthesizing relevant works to understand the current state of knowledge and identify research gaps.

1.2 Why is a Literature Survey Important?

- **Foundation for Research:** Builds the theoretical base for your study.
- **Avoids Duplication:** Ensures your work adds something new.
- **Identifies Gaps:** Highlights areas that need further investigation.
- **Guides Methodology:** Provides insights into existing methods and tools.

1.3 Examples in Computer Science

- Researching deep learning models for medical image analysis.
- Surveying blockchain technologies for supply chain management.
- Comparing algorithms for anomaly detection in cybersecurity.

2 Steps to Conduct a Literature Survey

2.1 Step 1: Clarify the Research Problem

Define the scope and focus of your research.

- **Example:** For a problem like *Reducing the energy consumption of IoT devices*, focus on existing protocols and energy-efficient designs.

2.2 Step 2: Identify Reliable Sources

- **Databases:** IEEE Xplore, SpringerLink, ACM Digital Library, Scopus, Google Scholar.
- **Other Sources:** Technical blogs, patents, and whitepapers.
- **Gray Literature:** Theses, dissertations, and preprints.

2.3 Step 3: Formulate Keywords and Search Strategies

- Use Boolean operators:
 - **AND:** Combines terms (e.g., *IoT AND energy efficiency*).
 - **OR:** Broadens search (e.g., *IoT OR edge computing*).
 - **NOT:** Excludes terms (e.g., *IoT NOT healthcare*).

- Explore synonyms and related terms:
 - Example: *Machine learning* could include *deep learning* or *neural networks*.

2.4 Step 4: Analyze and Categorize Sources

- Group papers by themes:
 - Algorithms (e.g., A*, Dijkstra's).
 - Datasets (e.g., CIFAR-10, ImageNet).
 - Applications (e.g., healthcare, finance).
- Use tools like Excel, Notion, or Zotero for organizing references.

2.5 Step 5: Summarize and Evaluate

- Compare approaches and methodologies:
 - Strengths and limitations.
 - Novelty of the research.
 - Relevance to your problem.
- Create a matrix or table to visualize comparisons.

3 Tools and Techniques

3.1 Academic Search Engines

- **Google Scholar:** Free, simple interface.
- **IEEE Xplore:** Focused on engineering and computing.
- **ACM Digital Library:** Covers computer science topics extensively.

3.2 Reference Management Tools

- **Zotero:** Free and open source, integrates with browsers.
- **Mendeley:** Includes annotation tools.
- **EndNote:** Advanced but paid.

3.3 Visualization Tools

- **VOSviewer:** For bibliometric analysis (e.g., citation networks).
- **Gephi:** Network visualization.
- **CiteSpace:** Identifies trends in research.

4 Evaluating Quality and Credibility of Sources

4.1 Peer-Reviewed vs. Non-Peer-Reviewed

- **Peer-reviewed:** Higher credibility, validated by experts.
- **Non-peer-reviewed:** May lack rigor; use cautiously.

4.2 Indicators of Quality

- **H-index:** Measures productivity and citation impact of authors.
- **Citation Count:** Indicates influence but check for self-citations.
- **Publisher Reputation:** Recognized publishers like IEEE, Springer, and ACM.

4.3 Avoiding Predatory Journals

- Look for journals indexed in Scopus, Web of Science.
- Avoid journals with high acceptance rates and low impact factors.

4.4 Example Activity

Compare a well-cited IEEE paper with a low-quality blog post on the same topic.

5 Structuring the Literature Survey

5.1 Common Structures

1. **Thematic:** Organized by topics (e.g., cloud computing, security, performance).
2. **Chronological:** Follows the evolution of research (e.g., AI advancements from 2010 to 2023).
3. **Methodological:** Compares research methods (e.g., supervised vs. unsupervised learning).

5.2 Writing Tips

- Use a formal and concise tone.
- Avoid plagiarism by paraphrasing and citing correctly.
- Example structure for a paragraph:
 - **Context:** Introduce the area.
 - **Details:** Summarize findings.
 - **Evaluation:** Highlight limitations or gaps.

5.3 Visualization Techniques

- **Concept Maps:** Show relationships between concepts.
- **Comparison Tables:** Example:

Algorithm	Accuracy	Complexity	Dataset Used
Algorithm A	95%	$O(n^2)$	CIFAR-10
Algorithm B	92%	$O(n \log n)$	ImageNet

6 Case Study: Literature Survey on a Sample Problem

6.1 Problem: Improving Recommendation Systems with Machine Learning

- **Identify Keywords:** *Recommendation systems, collaborative filtering, deep learning.*
- **Search for Papers:** Example: Netflix dataset studies.
- **Summarize Findings:**
 - Existing models (e.g., collaborative filtering, matrix factorization).
 - Limitations: Cold start problem, scalability issues.
- **Identify Gaps:** Lack of personalized recommendations for edge cases.
- **Present Findings:** Use a table or concept map to summarize.

6.2 Activity

1. Assign a topic (e.g., *AI in healthcare*).
2. Task:
 - Identify 3-5 papers.
 - Summarize findings in a table:

Title	Author	Year	Key Findings	Gaps Identified
Example Paper 1	Author A	2022	Improves accuracy by X%	Lacks real-time testing