Lorenz Glißmann, Sascha Safenreider

Working with Scientific Literature

Table of contents

"*Students shouldn't go into life without the ability to communicate. Your success in life will be determined largely by...*

- *your ability to speak,*
- *your ability to write, and*
- *the quality of your ideas,*

in that order."

— Prof. Patrick Winston

Goal: Find and Include Scientific Sources

- Discover relevant literature
- Read and understand state-of-the-art research
- Accurately provide references

Image source: uni-goettingen.de

Need for Accurate Information

- “Standing on the shoulders of giants”
  - All scientific developments extends previous efforts
- Research development commonly does one or a combination of these:
  - Generalize
  - Specialize
  - Correct

Strong Claims require Strong Sources

- Claims that contradict common understanding need
  - Direct proof and/or
  - Strong scientific evidence
- Claims that align with common understanding need either
  - A reference / citation and/or
  - Scientific Evidence
- Common knowledge might be cited but doesn't have to be
  - e.g. "Bubblesort's comp. complexity is in $O(n^2)$"

Good pratice: phase claims carefully, e.g. "In practice, recursion can often perform worse than iteration" instead of "Recursion is slower than iteration"

Non-Scientific Sources

Be careful with non peer-reviewed sources e.g.

- Blogs
- Articles
- Wikipedia
- ChatGPT and AI models
- Books
- Pre-print papers

In the end, it depends on the objectivity and reliability of any source.

Finding Credible, Citable Sources

- **Conferences:** e.g. NeurIPS, ICML, CVPR, SIGMOD.
- **Journals:** e.g. by
  - Professional Associations (IEEE, ACM, GI)
  - Publishing Houses (Elsevier, Springer)
- **Repositories:** arXiv (cross-check peer-reviewed follow-ups).
- Avoid unverified online content.

Conducting Effective Literature Searches

- Use e.g. **Google Scholar, IEEE Xplore, ...**.
- Combine keywords: “federated learning” AND “privacy” AND “benchmark”.
- Apply filters: year, publisher, citations.
- Read abstracts first to gauge relevance.

Classic Approach

Goal: Find papers related to your topic

1. Identify relevant keywords
2. Search e.g. on Google Scholar
3. Focus on recent works (and secondarily higher citation counts)
4. Collect papers in a reference manager (e.g., Zotero)
5. Start reading
6. Snowball: follow references in the papers you read

Practice – Literature Search Task

Group Task 10 Min

Goal: Apply the classic search approach to a real research question.

Task:

- Use Google Scholar or another academic search engine.
- Find 3–5 relevant scientific papers that answer a question:
  - **Does using AI make us dumb / happy / productive?**

Tip: Try keyword combinations like “AI usage”, or “Productivity”.

Classic Approach — Issues

- Identifying relevant keywords is hard
- Search is not semantic
- What about pre-prints without citation counts?
- Reading all papers takes a lot of time
- Exploring references could be automated

Identifying Keywords

- Ask supervisor/expert
  - For important papers
- Ask AI
  - Prompt generative AI for keywords
  - Use semantic search tools
- After you found some paper(s), find more keywords from
  - Paper titles & their (official) keywords
  - Introduction and references

Semantic Search

- Space is quickly evolving
  - AI Research tools read abstracts of papers
- Notable options right now
  - https://www.scienceos.ai/
  - https://consensus.app/
  - https://elicit.com/
  - https://www.openread.academy/

Practice – Try Semantic Search with Consensus

Individual Task 10 Min

Goal: Explore AI-based semantic search helps find relevant scientific results.

Task:

- Open `https://consensus.app/` in a private tab (no login).
- Search for the question:
  - **Does using AI make us dumb / happy / productive?**
- Review the summarized results:
  - Which papers are shown?
  - How is the answer phrased or summarized by the AI?

Tip: Compare Consensus results with Google Scholar — which one gives you more focused and credible findings?

AI Summarization of Papers

- In theory, most modern LLMs can summarize a paper
  - Papers are complex, AI might not fully understand
  - Visual understanding of graphs is relatively new
  - Modern AI may even consider multiple uploaded documents
- https://chat-ai.academiccloud.de/ GWDG ChatAI can summarize and understand a research paper
- https://www.semanticscholar.org/ provides good TLDR of papers

Practice – Summarize with GWDG ChatAI

Individual Task 10 Min

Goal: Use GWDG's ChatAI to summarize and understand a research paper.

Task:

- Go to `https://chat-ai.academiccloud.de/` and log in.
- Select a better model e.g. GPT OSS 120B or Qwen 3 235B Thinking
- Upload a paper e.g. from the previous task.
- Ask ChatAI to e.g.:
  - Summarize the main research question and findings.
  - Explain the method in simple terms.
  - Suggest one possible application of the work.

Automatic Snowballing

- **Snowballing** is a method to expand your literature review by exploring references linked to an initial paper (“seed” paper).
  - *Backward snowballing:* reviewing papers **cited by** the seed paper.
  - *Forward snowballing:* finding papers that **cite** the seed paper.
- This approach helps identify:
  - Seminal or highly influential works,
  - Research trends, and
  - Gaps or emerging areas around your topic.
- **Automatic tools** make snowballing faster and visual:
  - `https://www.researchrabbit.ai/`
  - `https://inciteful.xyz/`
  - `https://openknowledgemaps.org/`
- They show citation networks and topic clusters, allowing you to discover relevant papers you might have missed.

Practice – Explore Snowballing with ResearchRabbit

Individual Task 10 Min

Goal: Learn how to expand your literature review by exploring citation networks of a paper.

Task:

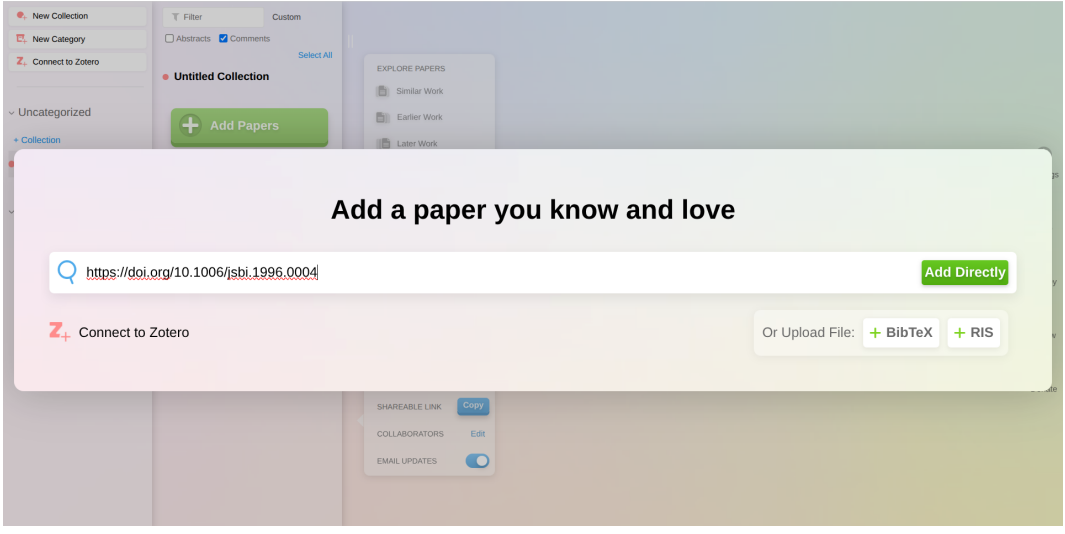
- Go to `https://www.researchrabbit.ai/` and **sign up** with your university or personal email address.
- After logging in, add a paper by entering its **DOI**, title, or direct link.
- Explore the visualization and identify:
  - Papers **cited by** your paper (backward snowballing)
  - Papers **citing** your paper (forward snowballing)
- Observe how papers are connected and how new relevant works can be discovered.

What is a DOI?

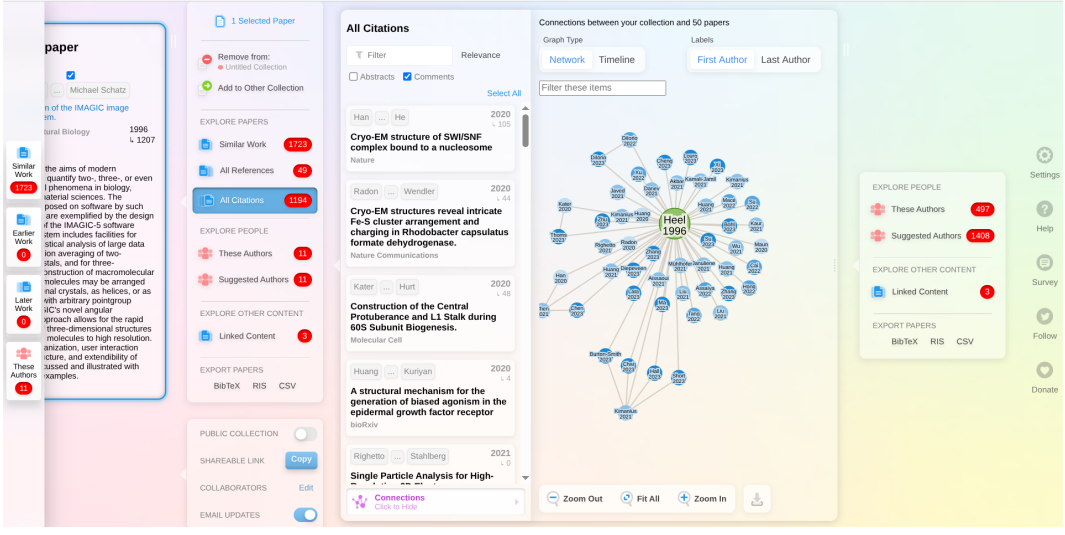
Digital Object Identifier (DOI)

- A **DOI** is a permanent and unique identifier for digital publications such as journal papers.
- It provides a stable web link to access the paper, e.g.: `https://doi.org/10.1145/1234567.8901234`
- You can find it on the paper's title page or publisher's website.

ResearchRabbit Visualization Example



ResearchRabbit Visualization Example



Identifying Relevant Papers

- Consider
  - Title
  - Publication year
  - Journal
  - Related organizations
  - Citations
- Read and consider if relevant
  - Abstract
  - Introduction
  - Conclusion
- If all this passes, read/skim the entire paper

How to Read a Paper Strategically

1. **Skim Pass:** Abstract, intro, figures, conclusion.
2. **Focused Pass:** Methods, data, evaluation.
3. **Critical Pass:** Identify gaps, assess rigor.

Ask: What's new? Is it reproducible? How does it fit my question?

Evaluating Research Quality

Criterion	Description	Example
Rigor	Logical, supported methods	Controlled experiments
Reproducibility	Accessible code/data	GitHub repo
Novelty	Extends knowledge	New algorithm
Impact	Practical significance	Improves fairness
Clarity	Clear, structured writing	Visuals, flow

How to Read a Paper

- Start with survey paper (if possible)
  - Provides wide introduction to field of research
- Read the paper from front to back
  - Take notes on the significant points
  - Look up unknown words
  - Note down interesting (key) references
  - Write down your ideas
- Answer these questions:
  - Why does this matter?
  - How is this useful to me?
- In your notes, ensure you can map information to references

Reference Manager

- What it does:
  - Collects and organizes scientific sources
  - Exports references into your documents
- Recommended tool: **Zotero**
  - Open-source desktop application + browser extension
  - Rich plugin ecosystem and integrations
  - We will introduce and use it in this course (in a few slides)

Introducing Zotero

Reading is not just about consuming papers —
it is about **organizing**, **annotating**, and **building knowledge**.

In this session, we will **install and introduce Zotero**,
a tool that helps you follow good scientific practice
and manage your reading efficiently.

Why use a system for reading?

Common problems:

- PDFs scattered across folders and devices
  - hard to find later
- Duplicate files and broken metadata waste time
- Forgetting key insights weeks after reading
- Manual bibliographies prone to errors and inconsistencies

What you gain with a system like Zotero:

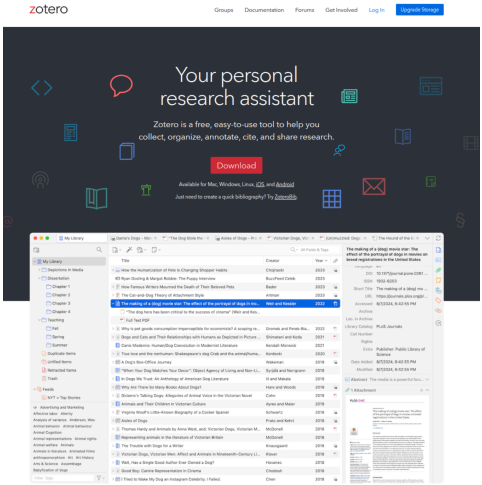
- Traceable and reliable sources
- Organized notes and searchable annotations
- Consistent and automated citation management
- Secure storage on **GWDG Cloud** (50 GB default, free for students)
- AI assistance (e.g., ChatAI API for translations and more)

Introducing Zotero

Zotero is a free and open-source reference manager that helps you:

- Collect and organize scientific sources
- Save PDFs, web pages, and metadata with one click
- Add notes and tags for better reading management
- Automatically generate citations and bibliographies

Website: https://www.zotero.org/



Practice – Download Zotero

Individual Task 5 Min

- Open `https://www.zotero.org/download/`
- Select the installer for your operating system:
  - Windows
  - macOS
  - Linux
- Download and run the installer

Tip: Make sure you choose the correct version for your system.

Practice – Account & Browser Connector

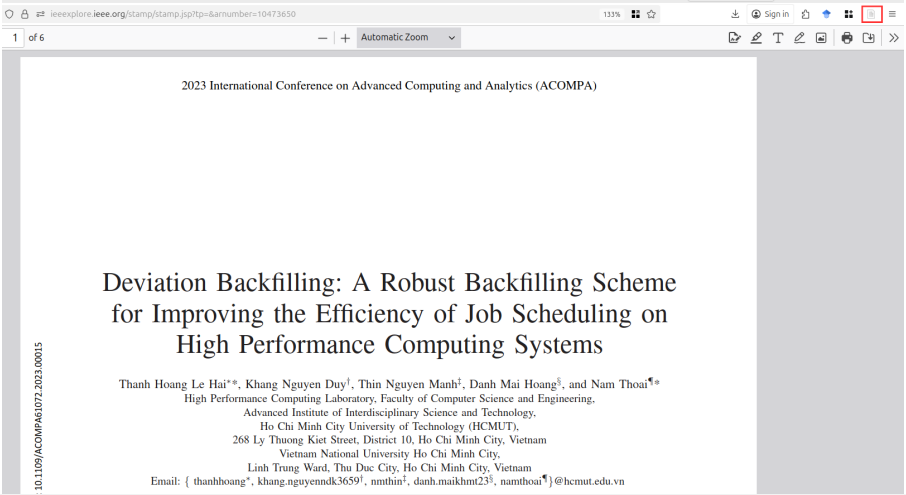
Individual Task 5 Min

- Create a free Zotero account: `https://www.zotero.org/user/register`
- Install the Zotero Connector for your browser:
  - Chrome / Edge (Web Store)
  - Firefox (Add-ons)
  - Safari (built-in option)
- Log in with your Zotero account to sync across devices

Tip: Connector lets you save papers and metadata directly from your browser with one click.

Zotero Browser Extension (Zotero Connector)

- **Collect** references via browser connector



What is WebDAV?

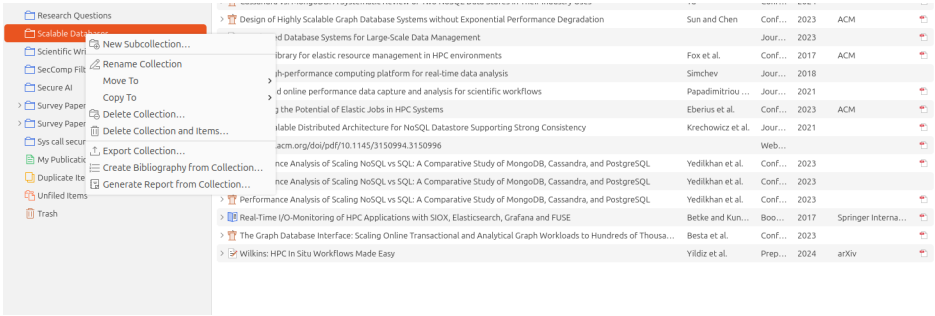
WebDAV – Web-based Distributed Authoring and Versioning

- A standard protocol that lets you **access and sync files over the web** like a remote drive.
- Supported by many cloud systems — including the **GWDG Academic Cloud (ownCloud)**.
- With your GWDG account, you have access to **50 GB of storage with full WebDAV support**.
- In Zotero, WebDAV allows you to **store and sync PDF attachments** across all your devices.

Tip: Think of WebDAV as a “bridge” between Zotero and your GWDG Academic Cloud storage.

Citing with Zotero and LaTeX

- **Export Collection**:
  - Right-click a collection → *Export Item ...* → Format: BibLaTeX → save as `ref.bib`



Citing with Zotero and LaTeX

- **Quick copy single reference:**
  - Open Pdf in Zotero
  - On the Pdf → press `Ctrl + Shift + C`
  - Paste in your LaTeX .bib file with `Ctrl + V`
- BUT: This is probably a format that we do not want

Citing BETTER with Zotero and LaTeX(1/2)

1. **Open Zotero.**
   Make sure you have the **desktop version** of Zotero installed.
2. **Download Better BibTeX:**
   `https://retorque.re/zotero-better-bibtex/`
   Scroll down and download the latest `better-bibtex.xpi` file.
3. **Install in Zotero:**
   - Go to `Tools` → `Add-ons`
   - Click the gear icon → `Install Add-on From File...`
   - Select the downloaded `.xpi` file
   - Confirm installation
4. **Restart Zotero.**

Citing BETTER with Zotero and LaTeX(2/2)

Set Export Format

1. Open `Edit → Preferences → Export`
2. Under “Default Format for Quick Copy,” choose:

Better BibTeX: BibLaTeX

3. From now on:
   - `Ctrl + Shift + C` (Windows/Linux)
   - `Cmd + Shift + C` (Mac)

   will copy selected references as `BibLaTeX` entries to your clipboard.

Citing with Zotero and LaTeX

- **Cite in LaTeX**:
- **`\cite{key}`**
  - Provided by standard LaTeX (with natbib, biblatex, or even without extra packages).
    - With default LaTeX: `\cite{key}` → [1] (numerical reference).
    - With natbib: `\cite{key}` → Author (Year) style, depending on options.
- **`\parencite{key}`**
  - Provided specifically by the biblatex package (not standard LaTeX).
  - Always puts the citation in parentheses (round brackets)
    - (Author, Year) (author–year style)
    - [1] (numeric style, but still in brackets)
- **Bibliography (show only, do NOT run here):**

  `\addbibresource{ref.bib}` (preamble) `\printbibliography` (at end)

Questions?

Cloud Sync with GWDG (WebDAV) — Step 1

🔁 **Goal:** Keep your PDFs available on all devices

- Open **Zotero** → **Edit** (macOS: **Zotero**) → **Preferences/Settings** → **Sync**
- Sign in to your **Zotero account**
- Enable:
  - **Sync automatically**
  - **Sync full-text content** (for PDF search)

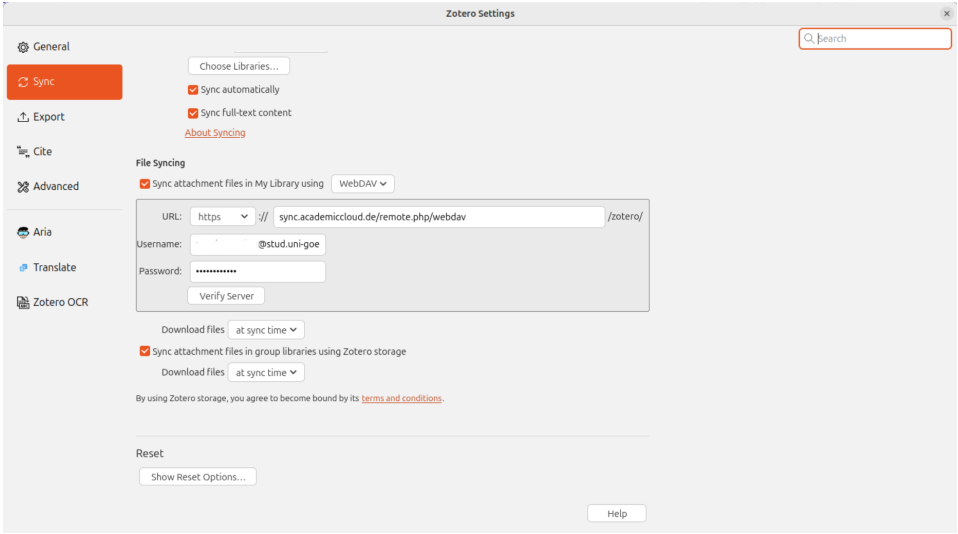
💡 Tip: Being signed in is required before file syncing can work.

Cloud Sync with GWDG (WebDAV) — Step 2

Store attachments in your Academic Cloud

- In **Sync** → **File Syncing**, tick **Sync attachment files in My Library using WebDAV**
- Set **URL**: `https://sync.academiccloud.de/remote.php/webdav/zotero/`
- Confirm the Creation of the Directory
- Enter your **Academic Cloud username** and **password**
- Click **Verify Server**

Cloud Sync with GWDG (WebDAV) — Step 2 (Screenshot)



Practice — Connect Zotero to GWDG Cloud

Individual Task 5 Min

- Open: **Settings** → **Sync**
- Sign in to your **Zotero account**
- Enable **WebDAV** for **My Library**:
  - URL: `https://sync.academiccloud.de/remote.php/webdav/zotero/`
  - Username/Password: your Academic Cloud credentials
  - Click **Verify Server**
- Choose *at sync time* or *as needed* for downloads

Tip: If verification fails, check the URL, credentials, or create the `zotero/` folder in the cloud web UI.