

Project Overview

ScholarFlow is a SaaS platform designed for researchers, students, professors, and academic teams to upload, manage, annotate, and collaborate on academic papers. It leverages Al-powered summarization, semantic search, and citation tooling to enhance research workflows. Users can create workspaces, organize papers into collections, track activities, and collaborate in real-time while maintaining version history of annotations and Al summaries.

Core Features







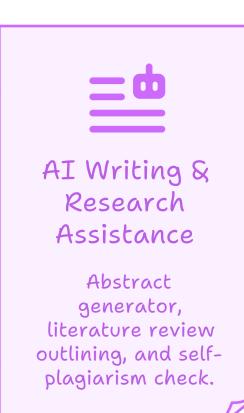
Format citations, visualize citation graph, suggest missing citations, and export citations.



Full-text search, embeddings for semantic search, filters, similar papers, and trends.



version history, collections per workspace, activity logs, and real-time updates.





feed.



Institutional plans, webhook handling.





1. Paper Management

Detailed Features

• Upload PDF, DOCX, LaTeX documents.

- Import papers via DOI, arXiv, OpenAlex, Semantic Scholar.
- OCR for scanned documents using Tesseract (Advanced-Optional if time-limited).
- Text parsing to extract title, authors, and abstract.
- Al-based smart tagging (topics, methods) and short summaries. • Bulk import with progress tracking.

Citation formatting: APA, MLA, IEEE.

2. Citation & References

- Citation graph visualization (Advanced-Optional).
- Missing citation suggestions using LLM similarity (Advanced-Optional).
- Export citations per collection as bibliography.

• Full-text search with keyword or semantic toggle.

3. Semantic Search & Discovery

- Embeddings for semantic search using pgvector (Advanced-Optional: vector index for large-scale).
- Filters: workspace, year, tags. • Similar papers recommendations.
- Trends via keyword extraction (Advanced-Optional).
- 4. Collaboration & Annotation
 - Annotations: highlights, threaded comments, notes. • Version history of annotations with revert capability.
 - Real-time updates via WebSocket/SSE (Advanced-Optional for minimal MVP).

• Collections per workspace with access control: Owner, Editor, Viewer.

• Self-plagiarism check using similarity across user's papers (Advanced-Optional).

5. Al Writing & Research Assistance

• Activity logs per collection/workspace.

• Abstract generator (Advanced-Optional).

• Literature review outlining (Advanced-Optional).

- Multiple workspaces per user. Invite members via email.
- Role-based access: Owner, Editor, Viewer. Workspace activity feed.
- 7. Payments & Plans • Freemium: Up to 100 papers, basic AI features, no shared collections.

6. Workspaces & Teams

• Payment integrations: Stripe (global), SSLCommerz (Bangladesh) (Advanced-Optional for MVP).

• Institutional: Workspace-wide access, SSO (Advanced-Optional).

Webhook handling for payment updates (Advanced-Optional).

• **Pro**: Unlimited uploads, AI tools, collaboration (Advanced-Optional for course MVP).

8. Backend & Data Management (Prisma Schema Highlights) • **User**: profile, role, workspace memberships.

• Paper: metadata, file URL, extracted text, tags, summaries.

• Annotation: highlights, comments, versions, timestamps.

• Collection: workspace-linked, paper links, members. • Workspace: owner, members, settings. • **Subscription & Payment** (Advanced-Optional).

• Al embeddings / vector storage for semantic search (Advanced-Optional).

• Search endpoints: semantic search, similar papers, AI summaries (Advanced-Optional).

Phase 1 MVP Development Sequence

comment, and save

annotations on

Users perform

keyword searches,

- 9. API / Integration Highlights • Paper endpoints: upload, import, list/filter, detail, delete. • Annotation endpoints: create, list, version history, revert.
 - Collection endpoints: CRUD, add/remove papers, invite members. • Workspace endpoints: create, invite, manage roles, activity log.
- Phase 1 Prioritized MVP
 - 1. Paper upload (PDF/DOCX) + metadata extraction. 2. Simple reference management (add, list, delete papers).
 - 6. Keyword search (semantic search can be marked Advanced-Optional). 7. Basic authentication & role-based access (JWT + workspace roles).

3. Collections per workspace (CRUD + membership).

4. Basic annotations (highlight, comment, save). 5. Dashboard with recent uploads and collection overview.

Phase 1 core MVP should focus on:

- **Paper Upload** Management **Basic** Users upload Users manage **Annotations** documents in PDF references by **Keyword Search** or DOCX format. adding, listing, or Users highlight,

deleting papers.

Reference

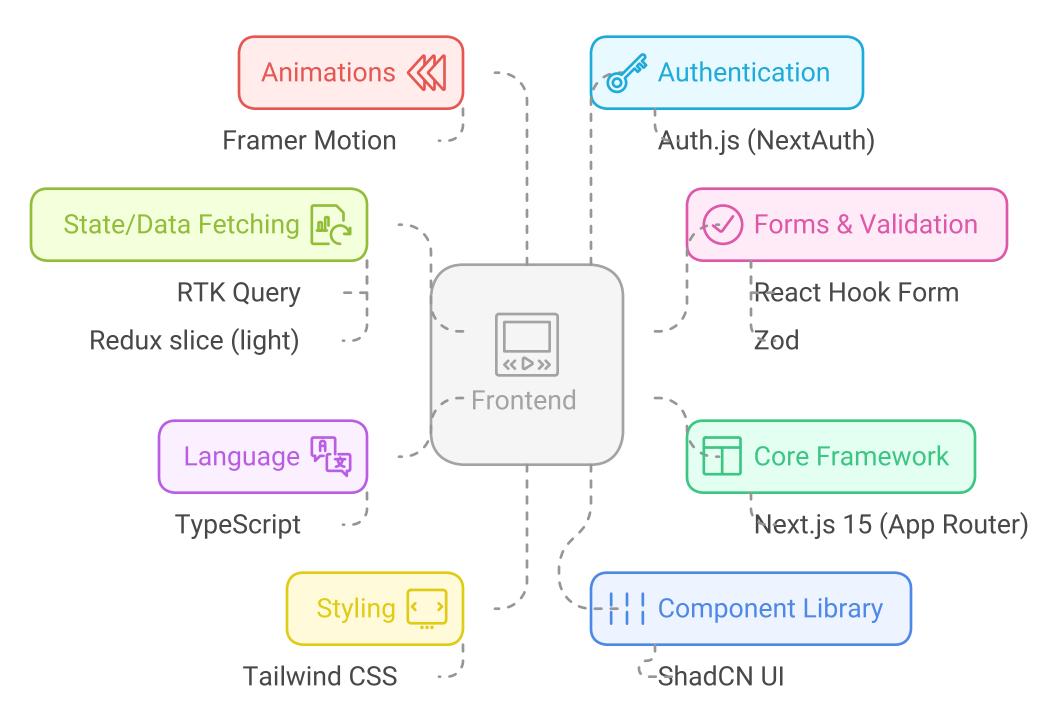


- Advanced/Optional features for post-course or Phase 2: • Al summarization & abstract generator.
 - Citation graph & similarity suggestions.
 - Full semantic search with pgyector. Real-time updates via WebSocket/SSE. Payment, subscription tiers, and institutional SSO.

• Literature review assistance & self-plagiarism detection.



Frontend



Made with 🍃 Napkin

Backend

