

IT 314

Group 25

Gen-scholar:-

A Collaborative Research Platform



Mid-evaluation Report

Table of Contents

1. Introduction
2. Stakeholder and User Identification
 - 2.1. Identified Stakeholders
 - 2.2. Stakeholder Rationale
3. Elicitation Techniques Used
 - 3.1. Brainstorming
 - 3.2. Interviews
 - 3.3. Documentation Review
4. Requirement Identification
 - 4.1. Functional Requirements (FRs)
 - 4.2. Non-Functional Requirements (NFRs)
5. User Stories
6. Requirement Division into Epics
 - 6.1. Epic 1: User & Workspace Foundation
 - 6.2. Epic 2: Core Document Collaboration & Deployment
 - 6.3. Epic 3: Contextual Discussion & AI Insights
 - 6.4. Epic 4: Review, Feedback, & Discovery
 - 6.5. Epic 5: System Administration & Reliability
 - 6.6. Epic 6: Publisher & Content Integrity
7. Conflict Identification Between Epics
 - 7.1. Conflict 1: Foundational Dependency on Epic 1
 - 7.2. Conflict 2: User Experience Overlap
8. Proof of Concept for Sprint 1
9. GitHub Project Repository

1. Introduction

The process of collaborative academic research often involves a fragmented workflow. Researchers typically use separate applications for reading papers, taking notes, and discussing findings, leading to disconnected feedback and inefficient knowledge sharing.

Gen-scholar is a proposed web platform designed to solve this problem by transforming research papers into interactive, shared workspaces. The platform will allow research teams to upload documents, generate AI-powered summaries, annotate text in real-time, and hold discussions directly within the context of the paper.

2. Stakeholder and User Identification

Identifying all individuals and groups affected by the system is critical for ensuring the project meets its intended goals.

2.1. Identified Stakeholders

Through initial analysis and brainstorming, we have categorised the stakeholders into two groups:

- **Primary Stakeholders:** These are individuals or groups who will directly interact with the Gen-scholar platform.
 - **Research Groups:** This includes graduate/phd students and undergraduate researchers who collaborate on literature reviews and research projects.
 - **Reviewers:** Faculty members or supervisors responsible for providing feedback on literature reviews and monitoring the progress of the work.
- **Secondary Stakeholders:** These stakeholders have an interest in the project but do not interact with it directly.
 - **Research Websites:** Academic databases and publishers (e.g., IEEE Xplore, ACM Digital Library) who could potentially integrate with our platform in the future.
 - **System Administrators:** Individuals responsible for deploying, maintaining, and securing the platform's infrastructure.

3. Elicitation Techniques Used

We employed a combination of elicitation techniques to gather comprehensive information from our identified stakeholders, ensuring our project goals aligned with real-world needs.

3.1. Brainstorming

Our team held multiple brainstorming meetings at the project's inception. The primary outcomes of these sessions were:

- **Defining the Core Goal:** We established that the system's main purpose is to create an effective collaborative workspace for small research groups, with an ideal size of 4-5 users per workspace.
- **Initial Stakeholder Identification:** We generated the initial list of all stakeholders mentioned in the previous section. We reasoned that while professors are part of the research ecosystem, our interviews later clarified that their collaboration patterns differ from those of student groups, so they are better classified as reviewers or supervisors rather than members of the core research groups the tool targets.
- **Technology Stack Decisions:** We discussed and decided upon the foundational technologies to be used for the application's development.

3.2. Interviews

This was our most critical elicitation technique for understanding user needs. We conducted semi-structured interviews with a total of **30 participants**: 10 professors and 20 students (a mix of B.Tech, MTech and PhD candidates).

We asked the following questions:

1. How do you currently read and manage research papers?
2. How do you usually decide whether a paper is worth reading in full?
3. What tools or platforms do you already use for research collaboration (if any)?
4. What are the biggest challenges you face when reading PDFs for research?
5. How do you currently extract figures, tables, or data from papers?
6. How do you collaborate with your team when discussing a paper and what challenges do you face while collaborating?
7. How do you keep track of questions or debates about a paper in your group?
8. Have you ever lost notes or struggled to find old discussions about a paper?
9. Would you find AI-generated summaries helpful? If yes, in what way?

10. If you could change one thing about how you currently work with research papers, what would it be?
11. What other features would you expect in the future (e.g., citation graphing, linking across papers)?

Summary of Responses:

- **Current Workflow (1, 3, 6):** The responses showed a strong reliance on **offline methods**. Many participants print papers for detailed reading and use in-person meetings with whiteboards for discussions. The primary challenge cited for collaboration was the **logistical difficulty of scheduling common meeting times**. For writing, teams use tools like Overleaf, but noted its lack of integrated discussion features.
- **Information Management and Loss (4, 7, 8):** A significant pain point is the fragmentation of information. Discussions are often captured in personal notebooks or as separate "minutes of the meet" documents. Participants specifically highlighted that these notes are frequently lost. Even when they are not lost, it is **very time-consuming to search** through past notes to find a specific comment or debate, as the information is disconnected from the source paper. This strongly validated the need for a centralised platform that preserves discussions in context.
- **AI Feature Validation (2, 9):** To decide if a paper is worth reading, most participants rely on the abstract, introduction, and conclusion. This confirmed that an AI summary feature would be valuable for accelerating this initial screening process. While question 2 was not directly about collaboration, it was essential for validating this feature. Users expressed more trust in summaries for well-researched papers than for emerging topics, especially concerning mathematical proofs.
- **Feature Prioritisation (5, 10, 11):** Extracting figures and tables was **not a major issue**; simple screenshots were deemed sufficient, leading us to de-prioritise this feature. While there was some resistance to changing established workflows, most agreed that a single, unified platform would be beneficial. Suggestions for future features included integrated video/voice chat and a system for individual team member evaluation, which could be a valuable feature for supervisors/reviewers.

3.3. Documentation Review

- **For Research Websites:** Our platform's workflow involves users downloading research papers from academic databases like IEEE Xplore and the ACM Digital Library and then uploading them to their private Gen-scholar workspace. Therefore, we conducted a review of the terms of service and content usage policies for these websites. The goal was to ensure that downloading a paper for use in a closed, collaborative group for annotation and discussion falls within the scope of fair use or their stated user agreements, and does not constitute unauthorised redistribution.
- **For System Administrators:** We reviewed standard industry documentation for deploying and securing scalable web applications. This included best practices for cloud infrastructure management and security guidelines. This analysis helped inform our non-functional requirements related to reliability, security, and scalability from an administrative and maintenance perspective.

4. Requirement Identification

This section details the functional and non-functional requirements for Gen-scholar.

4.1. Requirement Elicitation Process

Our approach to defining requirements was iterative. We started with the high-level goal of improving research collaboration and then used targeted elicitation techniques to collect specific, actionable requirements. The feedback from interviews was especially important, as it allowed us to validate our initial assumptions and prioritise features that would provide the most value to our primary users.

4.2. Functional Requirements (FRs)

- **User Authentication and Management**
 - **FR1:** The system shall require users to create an account and log in to access any functionalities.
 - **FR2:** Users shall be able to register for a new account using a username and password.
 - **FR3:** The system shall provide a secure password reset mechanism for users who have forgotten their password.
- **Workspace and Document Management**
 - **FR4:** An authenticated user can create a new, private collaborative workspace.
 - **FR5:** The creator of a workspace can invite other registered users to join via email or a unique link.
 - **FR6:** Users can upload PDF documents from their local devices into a shared workspace.
 - **FR7:** Users can view a list of all documents within a workspace and open them for reading.
 - **FR8:** Users with appropriate permissions can delete documents from a workspace.
- **Collaboration and Communication**
 - **FR9:** Multiple users in a workspace can view the same PDF document simultaneously.
 - **FR10:** Users can highlight sections of the PDF text and attach comments to these annotations.
 - **FR11:** All annotations and comments shall be visible to all members of the workspace in real-time.
 - **FR12:** The system shall support threaded discussions, allowing users to reply directly to comments to maintain context.

- **FR13:** Each workspace shall have a general chat area for discussions not specific to any document.
- **AI and Search Functionality**
 - **FR14:** The system shall provide a feature to generate an AI-based summary./Can ask questions to AI about an uploaded research paper.
 - **FR15:** Users can search for papers within their accessible workspaces by title and/or author.

4.2. Non-Functional Requirements (NFRs)

- **NFR1: Usability:** The user interface must be clean, intuitive, and easy to navigate.
- **NFR2: Performance:** The system must process and render PDFs quickly. AI summary generation should complete within a reasonable timeframe.
- **NFR3: Scalability:** The platform must support multiple concurrent users (approx. 4-5 per workspace) without performance degradation.
- **NFR4: Reliability:** User data, especially annotations and comments, must be saved reliably and protected against loss.
- **NFR5: Security:** All user data and documents must be stored securely, with role-based access control to manage permissions.
- **NFR6: Compatibility:** The web platform must be responsive and functional on major web browsers and across different screen sizes.

5. User Stories

No.	User Story (Front of Card)	Acceptance Criteria (Back of Card)
	<i>As a Research Group Member</i>	
1	Story: Account Registration As a new user, I want to register for an account, so that I can join workspaces and start collaborating.	Given I am on the registration page, When I enter my username and a valid password, Then my account is created, and I am automatically logged in,
2	Story: Creating a Workspace As a research group member, I want to create a shared, collaborative workspace, so that my team and I can coordinate our research activities in a single location.	Given a member is logged into the shared workspace, When they add, edit, or comment on a document, Then all team members must see the updates in real time, And a built-in chat feature must allow discussions directly within the workspace.
3	Story: Uploading a PDF. As a research group member, I want to upload research PDFs, so that my team can access and review shared papers.	Given a member is logged into the shared workspace, When they select and upload a research PDF, Then the file must be stored securely and become immediately accessible to all team members, And each member must be able to view the PDF.

4	<p>Story: Real-Time Annotation</p> <p>As a research group member, I want to annotate PDFs in real-time, so that my team can instantly view and respond to my insights.</p>	<p>Given a PDF is open in the workspace, When a member highlights or comments on text, Then all group members should see the annotation immediately, And annotations should auto-save without manual action.</p>
5	<p>Story: Threaded Discussions</p> <p>As a research group member, I want to create threaded discussions on specific document parts, so that our conversations remain contextual and structured.</p>	<p>Given a member highlights text or a figure, When they start a discussion by adding a comment, Then a new thread should be created and linked to that section, And replies should appear in a nested format under the original comment.</p>
6	<p>Story: AI-Generated Summaries</p> <p>As a research group member, I want to get AI-generated summaries of papers, so that I can quickly understand key insights without reading the full text.</p>	<p>Given a research paper is uploaded, When a member requests a summary, Then the system should generate a concise summary, And results must be delivered within a reasonable timeframe. And the summary must include the paper's objectives, methods, and results.</p>
7	<p>Story: Searching for Papers</p> <p>As a research group member, I want to search for uploaded papers,</p>	<p>Given that the workspace contains multiple uploaded papers, When a member searches by title and/or author,</p>

	so that I can quickly locate research documents.	Then the system should display all matching results instantly.
8	Story: Data Reliability As a research group member, I want to ensure annotations and discussions are reliably saved, so that no insights are lost during collaboration.	Given a member creates annotations, When the system crashes or the internet disconnects temporarily, Then annotations must still be preserved and restored upon reconnection, And all saved annotations must remain linked to the correct document section.
9	Story: Main Chat As a research group member, I want to have a main chat area in my workspace, so that I can have general discussions with my team that are not tied to a specific part of a document.	Given that I am in my workspace, When I open the main chat panel, Then I can see the conversation history with my team, And I can send new messages that are visible to everyone in the workspace in real-time.
	<i>As a Reviewer</i>	
10	Story: Reviewing Annotations As a reviewer, I want to view all annotations made by researchers, so that I can assess the quality of their analysis and collaboration.	Given a paper with annotations made by researchers, When the reviewer opens the paper, Then they should see all annotations along with the associated threaded discussions, And the reviewer should not be able

		to delete or modify existing annotations, only read them.
11	Story: Adding Feedback As a reviewer, I want to add comments or feedback on uploaded papers, so that I can guide researchers towards improvement.	Given a reviewer is reading a paper, When they highlight a section, Then they should be able to leave structured comments, And the feedback should be clearly labelled as “Reviewer feedback” to distinguish it from the research group's discussion.
	As a System Administrator	
12	Story: Managing User Accounts As a system admin, I want to manage user accounts and access levels, so that researchers and reviewers only see what they are allowed to.	Given a system admin is logged into the admin dashboard, Then the admin can create, update, deactivate, and delete user accounts. And roles (e.g., Admin, Researcher, Reviewer) must have clearly defined permissions. And role updates must take effect immediately.
13	Story: Backup and Recovery As a system admin, I want to set up automated backup and recovery options, so that valuable research discussions and annotations are never lost.	Given the system is operational, Then it must perform daily backups of documents and annotations. And the restoration process must be tested and functional.

	<i>As a Research Publisher</i>	
14	Story: Displaying Metadata As a research publisher, I want to ensure my papers are displayed with proper metadata, so that they are cited correctly within collaborative discussions.	Given that a user is viewing a paper, Then the displayed metadata must include title, authors, journal, year, and DOI. And metadata cannot be edited by standard users.
15	Story: Preserving Branding As a research publisher, I want to have my branding and source links preserved, so that researchers know the original source of the document.	Given a user is viewing a paper, Then the publisher's name must be displayed with the paper information. And a link to the original source must always be visible and accessible.

6. Requirement Division into EPICs (Each EPIC to be done in one sprint(of ~2 weeks))

We have grouped the user stories into larger functional blocks called Epics. This organisation helps in planning our development sprints and provides a high-level overview of the major components of the Gen-scholar platform.

6.1. Epic 1: User & Workspace Foundation

Goal: To allow users to register, log in, and create secure, private workspaces. This Epic covers the basic functionalities required for a user to start using the application.

- **User Stories Included:**
 - Account Registration (1)
 - Creating a Workspace (2)
 - Managing User Accounts (12)

6.2. Epic 2: Core Document Collaboration & Deployment

Goal: To provide the main collaborative features that allow teams to interact with research papers and ensure the system is deployed and accessible.

- **User Stories Included:**
 - Uploading a PDF (3)
 - Real-Time Annotation (4)
 - Main Chat (9)

6.3. Epic 3: Contextual Discussion & AI Insights

Goal: To provide focused, contextual discussions as the primary interaction method, enhanced by AI-powered tools to improve the research process.

- **User Stories Included:**
 - Threaded Discussions (5)
 - AI-Generated Summaries (6)

6.4. Epic 4: Review, Feedback, & Discovery

Goal: To build the specific tools required for the formal review process and to allow users to find relevant documents easily.

- **User Stories Included:**
 - Reviewing Annotations (10)
 - Adding Feedback (11)
 - Searching for Papers (7)

6.5. Epic 5: System Integrity, Administration, & Reliability

Goal: To ensure the platform is stable, secure, and maintainable from the backend, while also ensuring all content is properly attributed and sourced to maintain academic integrity.

- **User Stories Included:**
 - Backup and Recovery (13)
 - Data Reliability (8)
 - Displaying Metadata (14)
 - Preserving Branding (15)

7. Conflict Identification Between Epics

This section outlines potential conflicts and critical dependencies between the defined Epics.

7.1. Conflict 1: Foundational Dependency on Epic 1

- **Epics Involved:** Epic 1 (User & Workspace Foundation) vs. Epics 2, 3, and 4.
- **Reasoning:** This is a critical dependency conflict. The implementation of user roles, permissions, and workspace structure in Epic 1 directly impacts almost every other feature. For example, the permission model must be robust enough to handle the specific needs of the "Reviewer" role (Epic 4), which requires view-only access to annotations but permission to add separate feedback. A simplistic user model in Epic 1 would force significant rework when implementing features in later epics.

Resolution: This conflict will be resolved through development sequencing. Epic 1 will be implemented first. This ensures that the foundation is in place to support the requirements of all subsequent Epics without needing major architectural changes later on.


7.2 Conflict 2: User Experience Overlap

- **Epics Involved:** Epic 2 (Core Document Collaboration) vs. Epic 3 (Contextual Discussion).
- **Reasoning:** Epic 2 includes a "Main Chat," while Epic 3 includes "Threaded Discussions." There is a potential for user confusion if the distinction between these two communication channels is not clear. The user interface must be designed carefully to guide users on when to use the general chat versus a specific, contextual discussion thread. This conflict could lead to a fragmented user experience and redundant technical implementation if not addressed during the design phase.


Resolution: The conflict will be addressed through careful UI/UX design. The "Main Chat" will be implemented as a persistent, workspace-level panel for general communication. In contrast, "Threaded Discussions" will only be initiated by interacting directly with a specific part of a document (e.g., highlighting text and adding a comment). This clear separation in how the features are accessed and presented will make their distinct purposes intuitive to the user.

8. Proof of Concept for Sprint 1


Sign in page



Welcome to GenScholar
Your collaborative research platform for academic excellence



Research Hub
Organize and manage your research projects




Collaborate
Work together with researchers worldwide

Welcome Back

Sign in to your GenScholar account to continue your research journey


Username

Password
 


[Forgot Password?](#)

Log In


or continue with

 Continue with Google


Sign up page



Welcome to GenScholar
Your collaborative research platform for academic excellence




Research Hub
Organize and manage your research projects



Collaborate
Work together with researchers worldwide

Create Account


Join GenScholar and start your collaborative research journey


 Sign up with Google

or sign up with email

Full Name

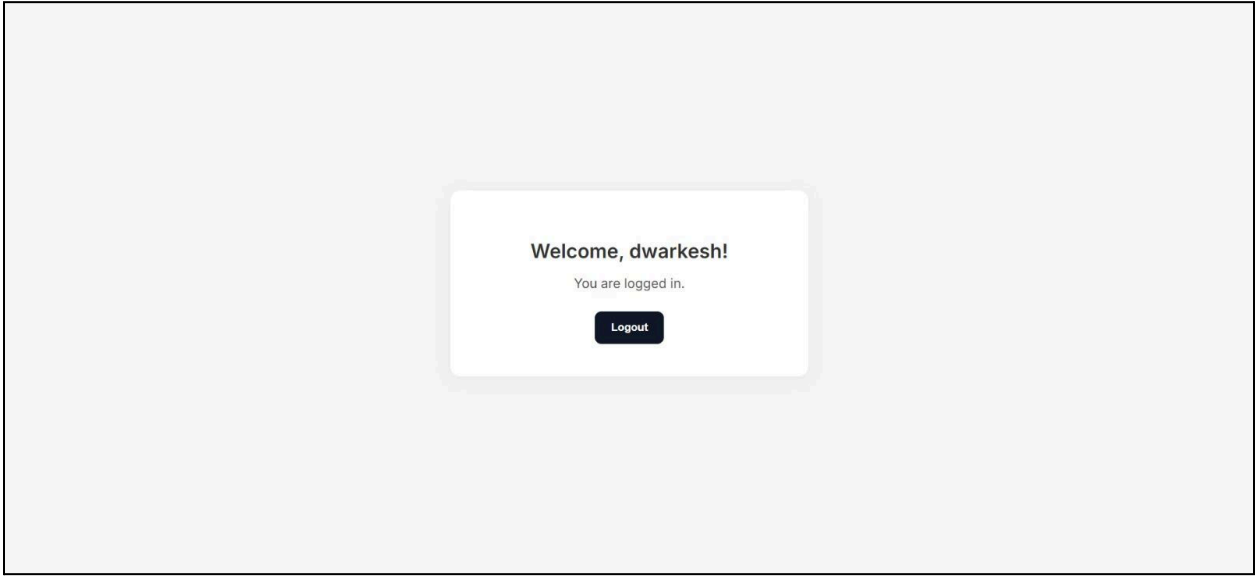
Email Address

Password
 

Confirm Password
 

Create Account

Logout page



9. GitHub Project Repository

https://github.com/yugsavalia/GenScholar-Collaborative_Research_Paper_Explorer

10. Contribution-

Yug Savalia(202301263) - Managed all team meetings, took interviews, wrote a summary for the findings of the interviews, made EPICs and Conflict resolution of EPICs. Made final report.

Archan Maru(202301217) - Took interviews, Identified Functional and Non-Functional Requirements, and then discussed them amongst the group. Worked on the Backend for the website.

Dhruvil Patel(202301201) - Made Questions for interviews, made user stories for Research Groups. Worked on the Backend for the website.

Vedant Patel(202301227) - Took interviews, made user stories for Research Groups. Worked on the Frontend part of the login page.

Manan Ghonia(202301240) - Took interviews, made user stories for Reviewer. Worked on the Backend for the website.

Dwarkesh Vaghasiya(202301225) - Took interviews, made user stories for research groups. Worked on the Backend for the website.

Vatsal Somaliya(202301210) - Took interviews, made user stories for Reviewer. Helped with report writing.

Dhyey Raval (202301253) - Made user stories for research websites. Made UI for the frontend using Figma.

Arav Vaitha(202301267) - Took interviews. Helped in Identifying Functional and Non-Functional Requirements.

Kanu Bhadraka (202301257) - Made user stories for system admin. Made UI for the frontend using Figma.