Entrepreneurship Coursework 1 – Business Plan

ScholarStack - A Complete Research Workspace

Aryan Chegini, Thamid Miah, Sultan Fawad, Morus D. Bowen

December 2025

Contents

1 Introduction

Writing dissertations and undertaking research projects in academia can be a highly stressful and fragmented process for researchers and students, particularly those who are undertaking dissertations or large research projects for the first time. Many individuals often struggle to locate papers, keep track of their reading and notes, and face the constant need to switch between tools for different tasks within the research workflow. Existing research tools such as ResearchRabbit, ReadWise and NotebookLM require students to constantly switch between interfaces, leading to fragmented workflows and increased cognitive load (Chis et al., 2023).

To address these challenges, we are developing ScholarStack, an integrated research platform designed to support students and early-stage researchers at every step of their research journey. Unlike traditional research tools for discovery and note-taking, our platform brings together the various aspects of research into a single environment. With institutional access, AI-powered explanations, visual research mapping, and personalised recommendations for research papers, the platform enables a more intuitive, structured way to conduct research, reducing the workload and stress for students and researchers.

2 Product

2.1 Description of Problem

Academic research today is fragmented across many tools. PDF readers, discovery platforms, note-taking apps, citation managers, and AI assistants all require researchers to manage separate workflows. Because these systems aren't unified, researchers must constantly switch tools, manually move information, and rebuild their understanding. This increases cognitive load (Liefooghe et al., 2008) and lowers the quality of reading and analysis (Chen et al., 2015).

A constant issue for students and early-stage researchers is "blank-page-anxiety": the struggle to start writing on an unstructured topic (Reigstad, 1985). Information overload, unclear research landscapes, and difficulty connecting literature to arguments all trigger this anxiety. Individuals spend more time initiating writing than critically engaging with the literature or developing insights. This is especially common for first-time dissertation or early research students. Processing large text volumes and converting notes to writing, often without guidance, slows progress. It increases cognitive stress and makes academic research development harder. This need for support motivates the development of a helpful research workspace.

Cognitive Load Theory suggests that human working memory has a limited capacity. When coupled with extraneous load (the way information is presented)—such as poorly structured tasks or context switching—learning and understanding are impaired (Medical College of Wisconsin, 2022). Conducting a literature review is already cognitively demanding. Many researchers, especially early-career academics and undergraduates, struggle with the complexity of literature reviews and the development of conceptual

understanding (Chen et al., 2015). When each stage of the review (e.g., searching, reading, annotating, summarising, citation management) uses different tools, information becomes scattered, connections between papers may be lost, and synthesising across papers becomes difficult.

Frequent task switching and context switching further exacerbate the issue. Empirical research in knowledge-intensive domains finds that switching between tasks and tools can significantly disrupt flows, degrade individual concentration, and slow down information processing (Abad et al., 2018). When it comes to academic researchers, they often need to deeply understand complex topics, draw parallels between different sources, and build out extensive arguments based on their readings. Such processes take intense concentration and are greatly impared when interupted.

Finally, the sheer volume of academic publications is growing very quickly. This means that the burden of tracking, reading relevant papers on the research topic, and organising the literature in a meaningful manner is increasing (Hanson et al., 2023). This increased influx of papers available to researchers further compounds the fragmentation problem. With more papers to read and manage, disjointed workflows, and the use of independent tools for each stage of research, the problem becomes even more unsustainable.

2.2 Description of Proposed Solution

Product Overview: ScholarStack
ScholarStack is a unified, comprehensive "Research Workflow Application" designed to address fragmentation and cognitive overload in academic research. Its innovation is a holistic approach: rather than segmenting tasks into disconnected modules, ScholarStack fully integrates discovery, access, analysis, and synthesis in a single environment. This design advances existing solutions by removing the need to switch between separate applications and reducing interruptions and cognitive demands. By embedding institutional access, a hierarchical notebook system, and context-aware AI into its architecture, ScholarStack enables seamless progression from discovery to drafting, offering an end-to-end workflow not found in current market offerings.

The solution is built upon **seven** core functional components:

1. **Project-Based Architecture**
   To eliminate the chaos of disjointed file storage, ScholarStack organises work into isolated Project Containers that act as folders containing all relevant papers and notebooks. Furthermore, within each project, the system generates a Project Knowledge Graph, which is a dynamic visual map of the the current research being used in a user's workspace. This graph uses "proximity grouping" to automatically position semantically similar papers (based on their vector embeddings) together, creating natural "neighbourhoods" of related research, while visible citation links allow users to trace the flow of evidence.

2. **Integrated Discovery & Seamless Access**
   Retrieval is often the peak of research friction. ScholarStack addresses this via Institutional Access. Users sign in once using Single Sign-On (SSO). The app then

unlocks and downloads PDFs behind a login/pay-wall (for example, from Springer) directly into the workspace. The system displays the "Citation Network" for any paper, including "Ancestors" (papers cited by the text) and "Descendants" (newer papers that cite it). "Smart Recommendations" suggest relevant literature based on the current project papers.

3. **Dual-Layer Notebook System**
   To solve the "blank page" problem, ScholarStack distinguishes between gathering information and building conceptual understanding, and producing work:

   - **Paper-Specific Notebooks:** Allow for granular, source-specific analysis of individual papers.
   - **Project-Wide Notebook:** Serves as the master document for drafting final outputs (e.g., a dissertation chapter).

   Crucially, this system supports a multi-page notebook structure (e.g. separate pages in the project-wide notebook for "Introduction" and "Methodology") and allows users to cross-reference notes between notebooks, aiding the process of writing by providing existing knowledge blocks.

4. **Active Highlighting & Annotation**
   ScholarStack replaces passive reading with an "Active Highlighting" workflow. When text is highlighted in a PDF, users see an "Add to Notebook" option. This instantly copies the quote to the active notebook page and allows for immediate annotation. Users can also apply semantic tags (e.g., #evidence, #critique) to both highlights and notes to organise their arguments effectively.

5. **Scoped AI Companion**
   ScholarStack uses a "Scoped AI Companion" built on a Retrieval-Augmented Generation (RAG) framework (Bolaños et al., 2024) to help users understand project content while avoiding hallucinations. Unlike generic chatbots, this system grounds answers strictly in the user's project documents to ensure verifiability.

   The companion supports a human-centric workflow through various mechanisms, two of which are:

   - **Contextual Querying:** Users can scope their queries to specific notebooks or PDFs (e.g., "Explain this concept based on this chapter"), with the AI providing direct citation links back to the source text for transparency.
   - **Micronote Expansion:** Following recent findings on active reading (Huq et al., 2025), the system encourages users to write short "micro-notes" which the AI then expands into fully developed paragraphs. This ensures the user remains cognitively engaged in the synthesis process rather than passively receiving generated text.

   Additionally, an "Explain This" pop-up allows users to select complex text within a PDF and receive an instant simplification to kickstart their note-taking.

6. **Dynamic Bibliography Generation**
   ScholarStack automates the tedious process of citation management through "Dynamic Bibliography Generation." Instead of dumping every file in a project into a list, the system scans the specific notebook page being worked on and generates a formatted reference list in a specific style, containing *only* the sources cited or linked in that specific section.

7. **Reading Interface**
   The user interface is designed to maintain flow through a split-view UI, placing the PDF viewer side-by-side with the active notebook page. This is enhanced via bidirectional navigation between a PDF and its notebook: clicking a reference in the notebook instantly scrolls the PDF viewer to the original location of that text, and clicking a highlight in the PDF opens the corresponding note, ensuring context is never lost.

Justification for Success:
We believe ScholarStack will succeed because it fundamentally solves the "Fragmentation Gap" identified in our market analysis. While the current EdTech market is saturated with specialised tools, they operate independently. ScholarStack's competitive advantage lies in its features and their integration into a cohesive workflow. Our approach is grounded in recent research on Human-AI interaction. As noted in NoTeeline (Huq et al., 2025), AI is most effective when used to expand user thought rather than replace it. We implement this through "micro-notes", short user fragments that our AI expands, ensuring users remain cognitively engaged. Furthermore, studies on AI-assisted literature reviews (Bolaños et al., 2024) suggest that the best results are achieved through Human-AI collaboration when mapping knowledge at scale. ScholarStack enables this by keeping the human in the loop during the critical synthesis phase and offering insight into where it's pulling its ideas from, distinguishing it from tools that merely offer "chat with PDF" functionality.

ScholarStack succeeds by unifying the distinct strengths of current market leaders into a single solution. Where Research Rabbit excels at discovery but fails at consumption, and Readwise Reader masters passive consumption but lacks academic rigour, ScholarStack bridges the gap. We offer the visual discovery of the former and the reading experience of the latter, but integrate them directly with the writing capabilities lacking in pure synthesis tools like NotebookLM. By resolving the friction of switching between these isolated apps, we capture the value lost during context switching. By housing the user's entire research lifecycle from the first citation discovered to the final paragraph written, ScholarStack creates a high "barrier to exit". Once a researcher establishes their Project Knowledge Graph and citations within our ecosystem, the cost of migrating back to a disjointed toolchain of three or four separate subscriptions becomes prohibitively high, ensuring high user retention and distinct value over separate point solutions.

2.3 Feedback & Prototype Solution

2.3.1 Initial Prototype

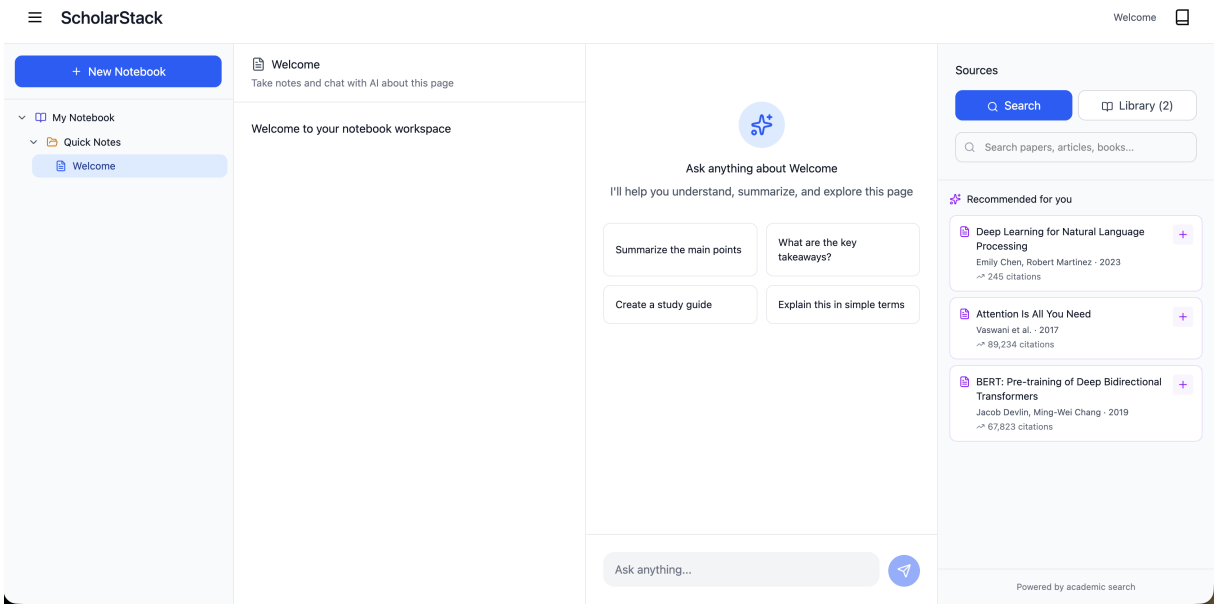

Figure 1: First iteration of prototype. Includes folder structure for general note-taking on the left-hand side, similar to applications like OneNote. Middle section represents AI integration within notes, and right-hand side consists of library of research papers and the search function. Bibliography is automatically created and exportable.

2.3.2 Feedback on Initial Prototype

After development of the first prototype, a mixture of quantitative and qualitative feedback was gathered through the use of a questionnaire ($n = 15$), as well as individual feedback from stakeholders. This was the primary driver of the second prototype, which improved and extended certain features to make AI-assisted research easier. Notable comments included adding cross-platform accessibility across devices such as tablets/iPads, rather than solely laptops and PCs. One interviewee stated that a "self-organising" folder structure would be very useful for creating and modifying sections based on the context of the notes or style of writing (e.g. Literature Review, Dissertation). Another consistent point of feedback reiterated the need for built-in bibliography management and citation tracking, reinforcing the use of that feature within the second prototype. Users also stressed the importance of linking AI-generated summaries back to the original paper, both for verification and to provide more details about the paper's contents.

Concerns about fabricated resources were explicitly stated, with almost half of the participants reporting that AI hallucinated or provided misinformation when used for note-taking or research. One participant reported that current LLM tools frequently invent citations or misattribute academic findings, which reinforced the need for verifiable, source-

grounded summarisation and citation. Another participant stated that it took them longer to check if sources were real than to find relevant ones themselves. Stakeholders also expressed strong interest in integrating research documents, including the ability to upload and query academic PDFs and other reading materials directly within the system. This included allowing the LLM to operate over user-provided research documents rather than relying solely on general knowledge. Some interviewees suggested incorporating domain-specific fine-tuning for educational or scientific contexts to improve factual accuracy and reduce hallucinations during research tasks.

The questionnaire responses provided further evidence that the proposed application addresses genuine pain points in academic research workflows. The vast majority of respondents were students (Figure A.1), and nearly all reported taking digital notes or managing research materials daily or several times per week (Figure A.2), indicating the importance of digital tooling methods as part of academic routine. Most users primarily relied on laptops/PCs (Figure A.3), but a sizeable portion also used tablets, echoing earlier interview feedback about the need for cross-platform accessibility. This reinforces the importance of ensuring app functionality seamlessly across devices, particularly in contexts where users annotate papers, highlight text, or interact with AI-powered summaries on the go.

A central finding from the questionnaire was the fragmented nature of current research workflows, with some respondents even using multiple apps solely for note-taking (Figure A.4), such as Goodnotes, Obsidian, and OneNote. Users frequently move between note-taking, PDF viewers, citation managers, and LLM interfaces, so integrating these functionalities into a unified platform enables greater workflow coherence and less cognitive overhead. We asked the question "How useful do you find AI-assisted note creation/summarisation?" (Figure A.5), where the responses were given on a Likert scale from not useful at all (1) to very useful (5). The average was 3.67, indicating an above-average need for the mentioned feature. Another question was posed on the same scale: "How useful would you find a library for storing papers, articles, and links?" (Figure A.7), with an average of 3.67. This result shows that the paper compilation feature is considered useful and should be integrated into users' workflows. For the question "Which AI note features do you find most valuable?" (Figure A.6), generating summaries was the most prominent, with 53.3% of people selecting it. The next two were expanding notes into detailed explanations and refining messy notes with 46.7% and 40% respectively. Finally, related topic suggestions accounted for 33.3%. Expanding notes reinforces the idea that users can take micro-notes (Huq et al., 2025) and that the AI assistant can elaborate and add context. Suggesting related topics is evidently a high priority, which is why the search-and-suggest feature was implemented in the latter prototype for paper searches. One participant also highlighted the fact that they had to use "multiple different AI tools for different things", claiming that ChatGPT was good for basic questions, Anara/GPT scholar was better for research, and Gemini for sentence formulation. This once again stresses the importance of a unified, end-to-end platform that can perform all these tasks without switching contexts or LLMs.

2.3.3 Final Prototype

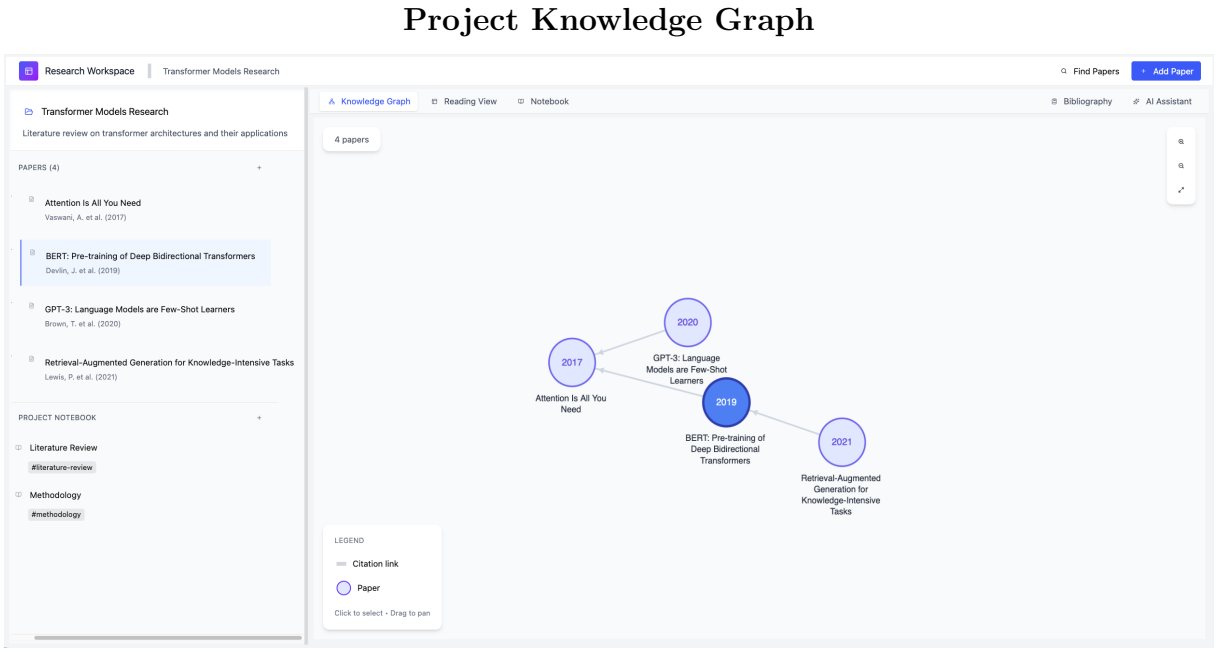

Figure 2: Part of the second iteration of prototype, visualising the Project Knowledge Graph within the workspace. It displays the Citation Network, using Citation Links (visible arrows) to connect papers that directly cite one another. This visual map allows the user to trace the flow of evidence and see the Ancestors and Descendants of the academic arguments currently in their project container.

The Reading Interface & AI Companion

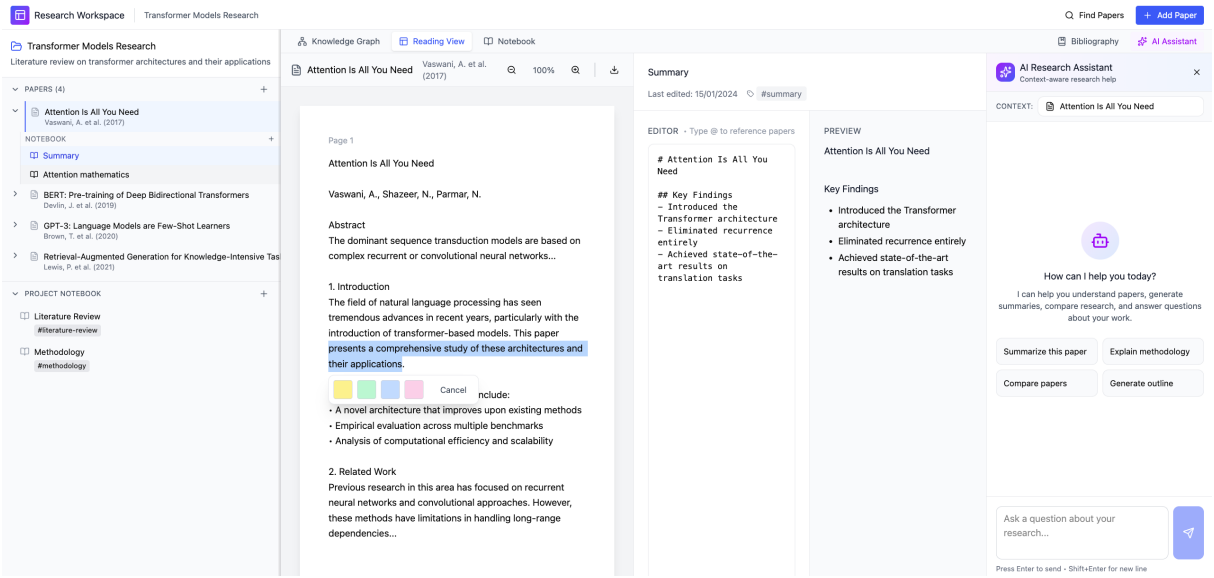

Figure 3: Displays the Reading Interface with a Split-View UI (PDF left, Scoped AI Companion right) with highlighting. The sidebar illustrates the Project-Based Architecture, separating the work into Paper-Specific Notebooks and the Project-Wide Notebook.

Project Notebook & Dynamic Bibliography

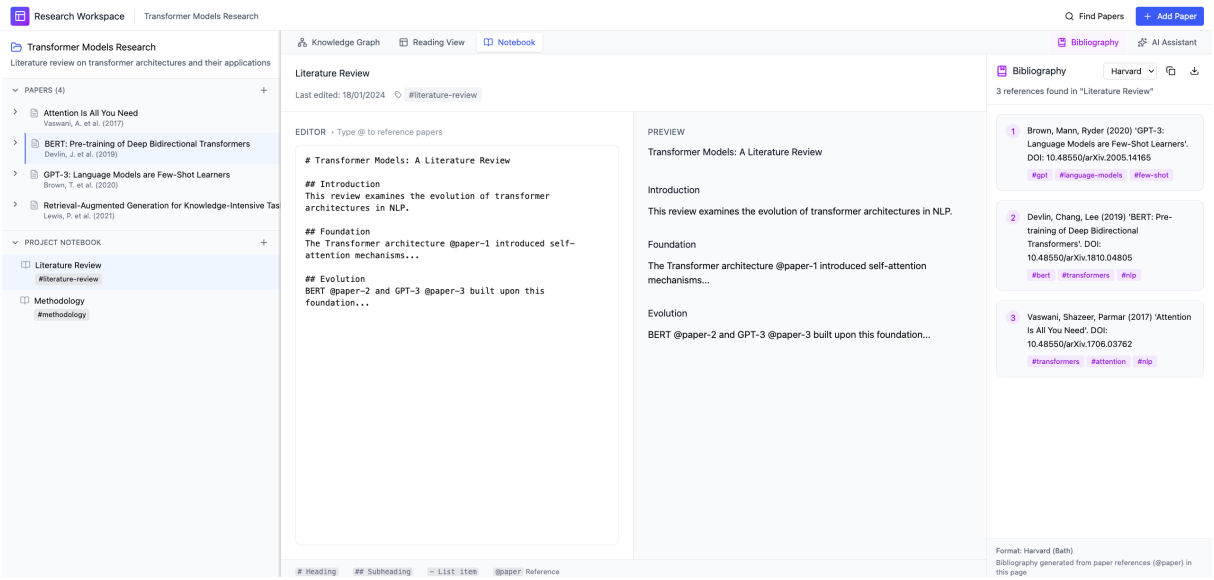

Figure 4: Shows the Project-Wide Notebook where users draft outputs using Markdown Support. The editor features Cross-Referencing (currently linking papers via @), which triggers the Dynamic Bibliography Generation on the right. This automatically creates a Usage-Based Compilation of references specific to the current section.

Integrated Discovery & Smart Recommendations

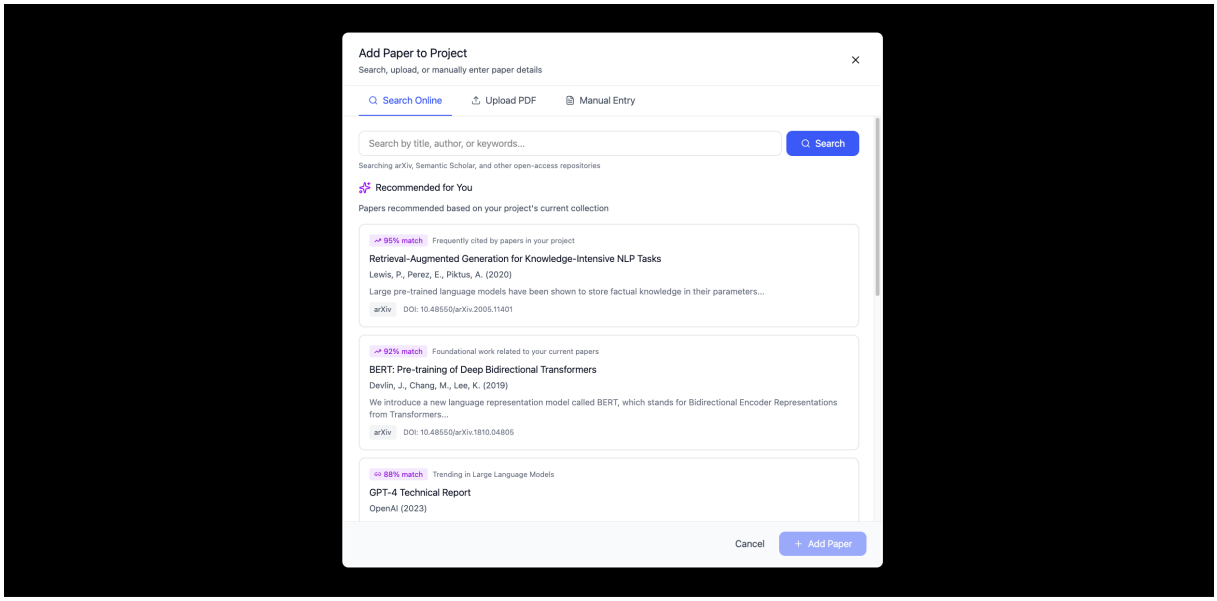

Figure 5: Showcases Integrated Discovery via In-App Search for open-access repositories. The "Recommended for You" section demonstrates Smart Recommendations, passively suggesting relevant literature based on the project's current content rather than just keywords.

Metadata Entry & Institutional Access

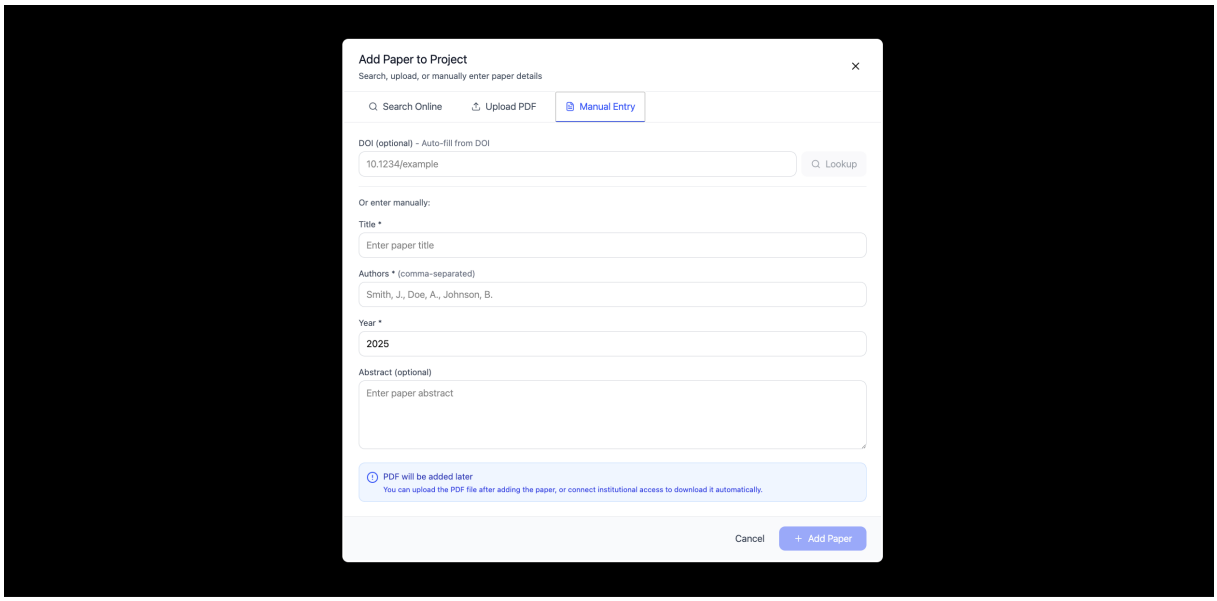

Figure 6: Demonstrates Smart Metadata entry, where a DOI lookup auto-fills paper details to eliminate manual typing. The notification highlights Seamless Institutional Access, confirming the system will automatically bridge university credentials to unlock and download paywalled PDFs.

2.4 Validation

2.4.1 Rigour in Research Gathering

To ensure the product addressed a scientifically valid problem, we looked at research from a variety of different angles. We grounded our initial hypothesis in secondary research, strictly selecting peer-reviewed literature from Google Scholar. By prioritising academic sources over general web articles, we ensured our "Problem Description" was based on established psychological principles rather than anecdotal evidence. This foundation was then cross-validated against the empirical data gathered from our primary user research.

2.4.2 Ethical Considerations on User Research

Ethical standards were strictly maintained throughout the data collection process. Participants were explicitly informed of the questionnaire's purpose and intended data usage via the Google Form preamble. To prioritise privacy, all responses were anonymised, and no personally identifiable information was requested. Additionally, users were notified of the data retention period (until 1st June 2026) and were assured of their right to opt out or stop the questionnaire at any stage prior to submission without their data being saved.

2.4.3 Software/AI Tools

Various digital tools were utilised to support the project's development. Figma AI assisted in creating both the initial and final prototypes, where prompts were carefully engineered to ensure the generated interface elements aligned with our functional requirements. For report structuring, AI tools helped generate complex LaTeX table formatting. In all instances, AI-generated outputs were evaluated and used solely for drafting ideas. Finally, Google Forms was used to distribute questionnaires and collect anonymous user data.

3 Market Section

3.1 Value Proposition Canvas

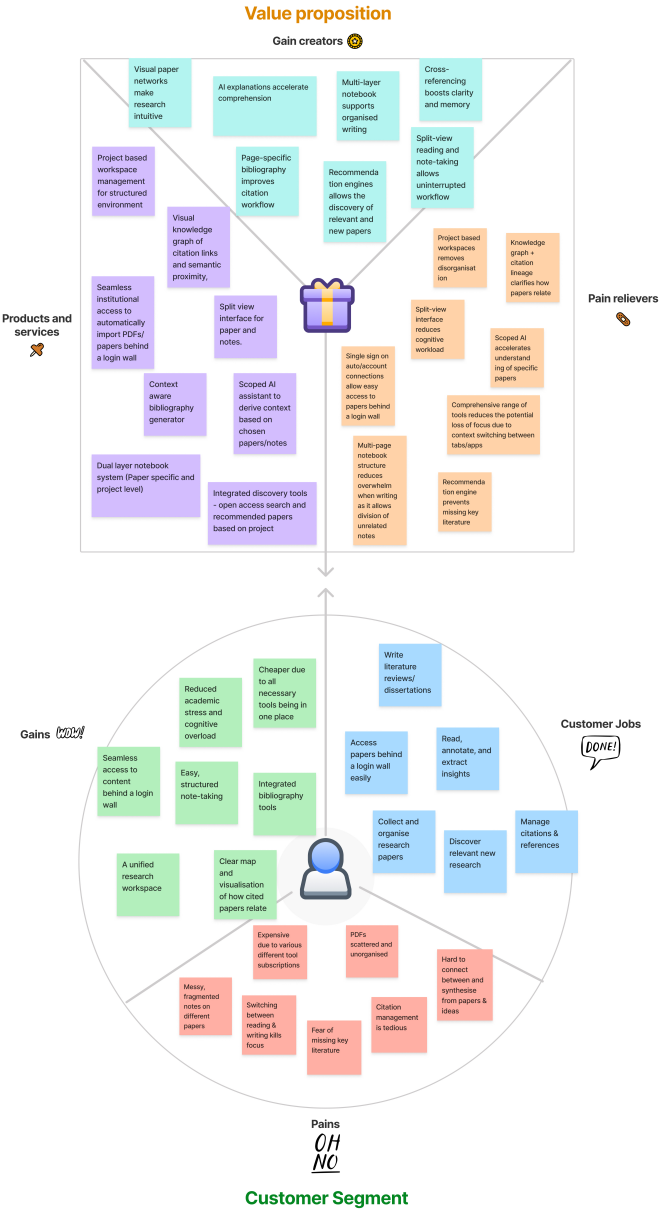

Figure 7: Value Proposition Canvas

The value proposition canvas shows how the proposed research platforms aligns with the needs and challenges that the intended users such as, university students, researchers and PhD students face. The canvas is split up into two segments, the Value Proposition, which address how the platforms features create value and solve the needs of the customers, and the customer segment, which describes the user's tasks, pains and desired outcomes.

Together, they demonstrate a positive fit with what the user requires and the proposed solution.

3.1.1 Customer Segment

Customer Jobs: Academic users must continuously manage complex research workflows. Their main tasks include finding relevant research papers, accessing paywalled papers, reading and comprehending their literature whilst annotating, organising large sets of research material in a coherent manner, writing structured assignments, and managing citations and references. This means they need a system that supports this kind of workflow and helps them unify the different aspects of research, rather than having to use separate tools for each task.

Pain: The current research workflow, especially for those early in their PhD or just about to begin their research projects, is fragmented, with many tools scattered across different services. Students and researchers need to maintain multiple subscriptions to access all the tools they need. PDFs can often be scattered across many web tabs, making it difficult to connect different pieces of literature. Note-taking for these research papers is fragmented across different platforms, leading to an inconsistent understanding of the papers. Switching between reading tools, writing tools, and citation managers can also be quite disruptive to users and increase their cognitive load. The above pains the user faces mean they struggle to understand how papers relate to one another and fear overlooking important literature. Citation management is also perceived as tedious and time-consuming, with no way to automatically add literature to a citation manager by simply importing it into the desired platform.

Gains: Users want a streamlined experience that brings the entire research workflow together, reducing their stress and improving productivity. They want seamless access to paywalled content, structured and intuitive note-taking for their literature, and a single, unified workspace that brings all their tasks together. Ideal benefits would include integrated bibliography tools, a clear visual map of how the papers they are exploring link to one another and automated assistance to reduce manual workload. A single platform which replaces multiple paid subscriptions would also lower any research-related costs.

3.1.2 Value Segment

Product and Service
Our research platform provides a comprehensive research environment with the following features:

- **Visual Knowledge Graphs** - to visualise the connections between different papers.
- **Dual-layer-notebook** - split into paper specific note-books and project level note-books
- **Split-view interface** - allows reading and note-taking to occur simultaneously.
- **Automated sign-on** - Institutional access to unlock paywalled PDFs.

- **Bibliography tools** - Note-book specific bibliography tools.
- **AI assistant** - Co-pilot like assistant that has context on notes and papers within the workspace

Pain Relievers
The product we came up with directly addresses user frustrations by introducing a structured, project-based workspace that eliminates disorganisation. The split-view interface reduces cognitive workload by allowing the user to simultaneously take notes and read each paper, instead of manually splitting their screen each time for different papers and sets of notes. The visual knowledge graph clarifies relationships between papers, enhancing users' understanding and enabling them to explore additional topics that haven't been covered. AI-driven contextual summaries and recommendations minimise the risk of missing key literature and help users understand complex texts.

Gain Creators
Beyond eliminating the different pain points, the platform actively enhances the entire research experience for students and academics. Visual paper networks make it easier to interpret academic landscapes and explore topics depending on the network. The platform enhances the academic writing experience through a dual-layer notebook structure that allows users to create organised notebooks tailored to their preferences. Integrated bibliography features simplify referencing, and an AI explanation and recommendation engine provides clearer insights into research content and new materials to explore. When these features all come together, they create a more intuitive, engaging academic workflow.

3.1.3 Value Proposition Fit

Overall, our proposed platform demonstrates a strong fit between user needs and the value it provides. Each of the major pains identified in our research is directly addressed by specific product features, as discussed above. At the same time, the solution enables the desired gains of structured writing, integrated bibliography management, intuitive knowledge graphs and access to literature in a unified environment. Instead of functioning as a collection of isolated tools, the components of the workspace come together to give the user a full research workflow experience.

3.2 Product Comparison

3.2.1 Local Competitor - University of Bath Research Support Tools

The University of Bath offers their students and researchers a wide range of library support services including subscription based and open-access databases to access a wide variety of academic journals, research papers, books (University of Bath, 2024). There are extensive guides on referencing, library loans, document supplies, research data services and dedicated specialist support for their researchers. These services provide researchers with

the required tools they need to have access to a wide range sources that are reliable and peer reviewed such as: journals, e-books, datasets, and other scholarly resources. They are often accessed through institutional authentication making them accessible to only enrolled staff and students.

However, while Bath's system excels at providing content and discovery, it generally lacks when it comes to integrated support workspaces to assist researchers annotating, taking structured notes, literature mapping. After downloading or accessing a paper user's often have to use external tools for notes, references and organisation. This lack of continuity forces researchers to juggle multiple tools increasing their cognitive load and efficiency when carrying out research.

In contrast out product, aims to complement rather than replace institutional services like those available at Bath. By creating a unified environment, our platform fills a crucial gap, transforming raw access to content into a manageable and reusable personal research system.

Therefore, whilst Bath's library services represents a very strong local competitor in terms of access and the breadth of content available, they do not address the full workflow needed for modern researchers. Our solution targets this exact need by providing value through a unified workspace and a wide range of tool integrations.

3.2.2 National Competitor - CORE (UK Open Access Aggregator)

CORE is a UK-based open access aggregation service developed The Open University's Knowledge Media (CORE, 2025). CORE harvests research outputs from institutional repositories as well as open-access and hybrid journals, and provides open-access to hundreds of millions open-access papers around the world. It's primary purpose is to improve the discoverability and accessibility of research outputs to support the scientific infrastructure for the UK and international institutions.

In comparison to ScholarStack, CORE operates mainly as a large-scale search service rather than a personal research workspace. It is highly effective at locating relevant papers and literature and provides APIs and analytic services for institutions. CORE helps users find papers, our product is designed to help users work with those papers overtime by combining reading, structured note-taking, AI-assisted comprehension of unfamiliar texts in a single environment. With this in mind CORE does represent a powerful national level infrastructure for open access discovery, whilst our prototype focuses on the individual researcher's workflow and the creation of a personal and organised knowledge base.

3.2.3 International Competitors

Table 1: Comparison of Our Product With Similar International Software

Criterion	Our Product	NotebookLM	Readwise Reader	Research Rabbit
Input File Types	Web published papers and PDFs	PDFs, Docs, ePUBs and web articles.	Amazon and Apple Books, Web Articles, Ebooks, PDfs, docs, Twitter, Podcasts, CSV, txt.	Web published papers
Core Functionalities	Notebook Storage, Browsing research publications, annotation, highlighting and note taking, Graph displaying citations and references, LLM co-pilot	Upload documents, AI co-pilot, visuals generation, find web sources, note-taking	Upload documents, annotation, highlighting and processing, AI co-pilo (Ghostreader), cross-device access.	AI-powered citation based search engine, graph of citations and references, grouping and saving found papers, share researcher papers, receiving updates.
Platform Coverage	Website	Website	Website and App	Website
Primary Usage	Researching papers and aggregating understanding of research.	Note-taking and chat bot help particularly on topics with a complex and wide landscape of information.	Digesting, annotating and retaining long form content.	Browsing research papers, particularly when creating literature reviews or a dissertation.
Privacy	User data encrypted via AWS SSO integration ensures no instituational data is saved Full GDPR compliance	User files not used for LLM training Continuos encryption for user data User data saved on Google servers, sensitive data should not be included	External parties can't access user data, Minimal LLM training on user data, Cookies collected only for security, maintenance and product improvement.	Uses standard web analytics (including marketing and recommendations) Users can choose to keep collections private or public. Supports third-party integrations for reference synchronisation.
Pricing	Free tier, Pro tier: 13.00 monthly	Free tier, Paid tier: 18.99 monthly	Annual Plan: 9.99 monthly, Monthly Plan: 12.99 monthly	Free tier, Annual Plan: 7.50 monthly, Monthly Plan: 7.50 monthly
Links		https://notebooklm.google.com	https://read.readwise.io/	https://www.researchrabbit.ai

NotebookLM

NotebookLM, a comparable product to ours, is a service by Google that offers an AI co-pilot to aid a user's process of understanding information sources. Its purpose is to provide AI-powered research and learning assistance whilst digesting large and complex

information, particularly when multiple information sources are required. The key similarity to our proposed product is the focus on creating a knowledge base of sources and using AI tools to aid understanding and note-taking. However, NotebookLM's note-taking capacity is limited, as it doesn't allow users to read the sources on the site; thus, users cannot add annotations or highlighting to the source text. The site has a very limited capacity to find new sources but a broader range of input data types, as it allows the user to upload media content and websites.

Readwise Reader

A site that allows a user to add books, articles, PDFs, and online content. The site's highlighting, note-taking, and annotation tools facilitate the digestion of information and the creation of a knowledge base across multiple devices. Its core purpose is to offer tools that deepen their comprehension of long-form content. Our product also enables annotating text sources and utilising AI chatbot features to facilitate comprehension of long-form text. Both are designed to enable users to accumulate a knowledge base from their findings. However, Readwise Reader is focused on creating notes from a single source at a time, whereas our product allows note-taking to occur at multiple levels, from single to multiple sources, or with no sources, and accumulates notes from multiple sources in one location.

Research Rabbit

The site offers an AI-powered search capability with adaptive and personalised recommendations. Its graphing feature allows users to see the relationships between papers, aiding in finding relevant papers in their previous domain of interest and visualising those relationships. Its core purpose is to improve the ease and enjoyment of browsing large volumes of research publications through its tools. Both products aim to increase the ease of browsing for papers and using graphical visuals to aid in finding relevant papers. However, Research Rabbit doesn't allow users to read papers on its site, annotate sources, or create notes, making it a limited tool for building a knowledge base of research findings.

3.3 Business Model Canvas

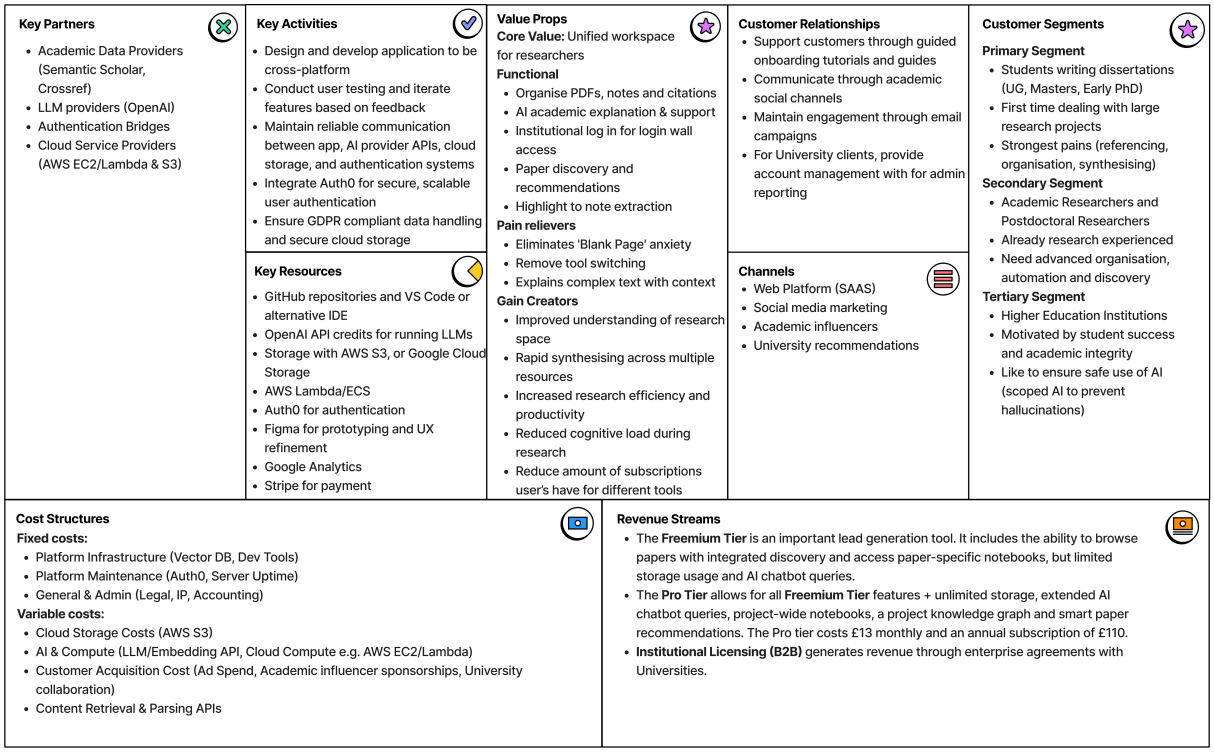

Figure 8: Business Model Canvas

3.3.1 Customer Segments

The primary customer segment consists of undergraduate, master's, and early-stage PhD students who are writing dissertations or undertaking a substantial amount of research for the first time. This particular group faces the biggest challenges, as many are conducting their first-ever academic research, which can be very overwhelming. They are typically unfamiliar with research workflows, face a steep learning curve, and have difficulty synthesising large amounts of academic information; therefore, they benefit most from the structured support our platform provides, along with its integrated tools. The secondary customer segment includes academic and postgraduate researchers who already have research experience but require more advanced support for organisation, automation, and discovery. This segment values the ability to connect research materials across multiple sources and improve their productivity through automation and AI assistance. The tertiary segment includes higher education institutions that seek to support their students. This segment is motivated by institutional goals, including academic compliance and quality of research outcomes. Universities can be valuable enterprise customers through licensing agreements.

3.3.2 Value Propositions

ScholarStack offers a unified academic workspace that brings together all aspects of research from discovery, reading, annotations, and writing within one single environment. The core value lies in integrating multiple research activities that are often fragmented across many tools. This ensures that the research workflow is simplified and the cognitive load, as discussed during our research, is reduced. The functional elements discussed in section 3.1.2 directly address the customer's pain points. Pain relievers address major academic challenges by eliminating blank-page anxiety, reducing the need to switch tools, automating access to content, and fostering a deeper understanding of complex materials. The gain creators discussed in 3.1.1 and Figure 8 extend beyond the problem environment. By enabling improved understanding of the research landscape, rapid synthesis of ideas across multiple papers, increased productivity and lower costs through the consolidation of multiple subscriptions, we allow our customers to focus on what is necessary without having to compromise the quality and depth of their research, but rather allow them to enhance and excel their experience.

3.3.3 Channels

Our channel strategy mixes the use of digital outreach and institutional partnerships to minimise Customer Acquisition Cost (CAC). The core product is delivered as a web-based Software as a Service (SaaS) application, allowing for instant deployment and seamless updates without requiring users to install heavy software. To reach the student market effectively, we aim to partner with niche academic influencers on YouTube and TikTok; individuals who have a fanbase who trust their judgment. These creators demonstrate the workflow benefits of ScholarStack to their engaged audiences, serving as a high-conversion marketing channel that feels organic rather than corporate. For the enterprise segment, our channel strategy focuses on exposure within University Library dashboards. We aim to embed ScholarStack directly into the resources page of partner universities to gain immediate legitimacy and direct access to the entire student body.

3.3.4 Customer Relationships

The platform establishes and maintains customer relationships in many different forms, including automated onboarding systems, targeted communication tools and personalised support. New users will be introduced to the platform through a guided onboarding process that explains, step by step, how to use the product and its features. Interactive walk-throughs and videos to watch in case they forget how to use a specific feature will be available at all times. To keep costs low, the onboarding flows and videos will be produced in-house, and later we can switch to external services to provide further support. Ongoing engagement will be maintained through academic-focused communication channels, including emails, newsletters, and educational content campaigns. Tools such as Mailchimp can be used to convert users to the paid subscription tier, share research tips via newsletters, and provide personalised insights via email. We can also highlight new features or product updates through our email channels to keep our users informed. Social academic

channels and university-affiliated networks can help build product credibility and support peer learning about academic workflows. For our institutional clients, the relationship will be a little different and become account-based, with dedicated support, integration assistance to onboard all their users, and other administrative analytics. Account management may include regular check-ins with specific institutions, university onboarding, and customer support whenever required. This mixed model ensures that individual users receive a self-service experience throughout online support and walkthroughs, whilst university-affiliated customers receive guided deployment, deeper support structures and ongoing communication from their institutions. Through both channels, the relationship evolves from initial onboarding and support to long-term engagement with ScholarStack, product adoption and institutional integration.

3.3.5 Key Resources

The business relies on a combination of technical, cloud, and software resources essential for building and maintaining the platform. Source code and collaborative development are managed through GitHub repositories and IDEs such as VS Code. This improves efficiency when building features in parallel with other developers. AI-powered features depend on OpenAI API credits, enabling access to large-scale language models for summarisation, organisation, and research assistance. The application's backend would consist of a cloud storage provider such as AWS S3, while compute and service orchestration would be handled by a service such as AWS ECS. For user authentication and authorisation, especially in organisational contexts, Auth0 would be used as a secure framework. Design and user experience are supported in Figma through AI-assisted prototyping for rapid development and user feedback. Additional resources include Google Analytics for monitoring user behaviour and performance, while Stripe would be used to securely manage subscription payments and billing processes. These resources collectively enable robust development, secure operations, and long-term scalability.

3.3.6 Key Activities

We start by designing our application after having conducted the preliminary step of conducting structured user testing and obtaining feedback, with which we iteratively improve upon the design and features. The core development will focus on building a cross-platform research and note-taking application that meets users' needs. Operationally, the team must maintain seamless communication between the app, AI provider APIs, cloud storage systems, compute resources, and authentication services to guarantee fast, stable performance. A major step is integrating Auth0 to deliver secure, scalable authentication and account management, enabling required features such as Single Sign-On. Data protection is also a key operational responsibility, requiring strict adherence to GDPR principles, secure cloud infrastructure setup, and regular security reviews to ensure user trust and regulatory compliance.

3.3.7 Key Partners

ScholarStack utilises its key partners to outsource strategically to maintain a lean operation and keep its development simple. One of the key partners is Academic Data Providers, such as Semantic Scholar and Crossref, which supply us with raw metadata (titles, abstracts, DOIs) to aid with our discovery engine. Another main partner is the companies that provide the technology infrastructure we require. These include an LLM provider (likely OpenAI), a Cloud service provider for compute and storage (AWS), and authentication services (likely Auth0) to seamlessly bridge university students to providers in-app.

3.3.8 Financial Plan

The following analysis outlines our projected revenue streams and cost structures across three growth stages.

Revenue Model ScholarStack employs a hybrid revenue strategy, leveraging a low-barrier Freemium B2C model to build a user base that can be monetised through premium upgrades and high-value B2B institutional contracts. A detailed comparison of the Free vs. Pro tiers is available in Appendix A.1.

Stage 1: Development

- **Strategy:** Zero Revenue. Focus is strictly on product validation and code stability via a closed beta.
- **Funding:** Costs are covered by self-funding and university innovation grants (e.g., Bath Entrepreneurs Society).

Stage 2: Establishment

- **Total Addressable Market:** The University of Bath has approximately ∼21,000 students and academics (University of Bath, 2025*a*) and ∼1600 academic staff (HESA, 2025), totalling and addressable market of ∼23000 users.
- **Acquisition Target (Freemium):** We aim to capture 5% of the campus population (∼1,150 users) as active free users. Acquisition will be driven by low-cost channels including word-of-mouth, the Entrepreneurship Society, and lecture shoutouts.
- **Conversion Target (Pro Users):** We project a conservative conversion rate of 2% of our Freemium user base (approx. 23 users) upgrading to Pro. This aligns with the lower bound of industry benchmarks for Freemium B2C products (typically 2–5%) (Pathmonk, 2024).
- **Projected Revenue:** For the Pro tier, there is a £13 per month price point. We believe this is a fair price as it consolidates multiple existing subscriptions (Readwise,

ChatGPT+, Reference Managers), offering a net saving to the student. 23 users $\times$ £13 = **£299/month**. While nominal, this stage validates the customer base' willingness to pay before scaling.

Stage 3: Growth We expand beyond Bath to target a "Cluster" of 10 similar UK universities (e.g., Russell Group institutions), creating a new addressable market of ∼300,000 students (HESA, 2024).

- **B2C Scaling:** Maintaining a 5% market penetration on this larger TAM yields ∼15,000 freemium users. We aim to raise the conversion rate to **4%** (600 paid users) by utilising high-intent email marketing campaigns (Mailchimp).
- **Institutional Licensing (B2B):** We will aim to leverage our established user base to secure a bulk contract with a single large faculty (e.g., Faculty of Science) in the University of Bath which contains ∼4500 students and academics (University of Bath, 2025*b*).
- **B2B Revenue:** A bulk license for 1,000 students at a discounted rate of **£8/user** generates **£8,000/month**.
- **Total Projected Revenue:** £7,800 (B2C) + £8,000 (B2B) = **£15,800 Monthly Recurring Revenue (MRR)**.

Cost Structure & Profitability Costs are categorised into Fixed (Infrastructure) and Variable (Usage-based), scaling appropriately across the three development stages.

Stage 1: Development Costs Capital allocation during the bootstrapping phase is strictly focused on technical validation.

- **Infrastructure:** We leverage free tiers for Auth0 (<25k users) and Pinecone Serverless.
- **LLM Prototyping:** Minimal costs for OpenAI API tokens during alpha testing.
- **Total Burn:** Negligible (100), covered by self-funding or micro-grants.

Table 2: Stage 1 Development Cost Estimates

Cost Item	Lower Estimate (£)	Upper Estimate (£)
LLM Usage (Alpha Testing)	0.08	8.45
Domain Registration	0.01	77.99
Total Development Cost	**0.10**	**87.19**

Stage 2: Establishment Costs (1,150 Freemium & 23 Pro Tier Users) Upon public launch, we transition to paid infrastructure (Amazon RDS, Auth0 Standard) to ensure reliability for our initial cohort.

Table 3: Stage 2 Monthly Cost Breakdown

Cost Item	**Lower Estimate (£)**	**Upper Estimate (£)**
Fixed Costs		
Pinecone Vector DB (Standard)	37.50	37.50
Auth0 Free	0	0
Variable Costs		
LLM Inference (GPT-5)	88.00	103.00
Amazon RDS (PostgreSQL)	54.04	91.67
Amazon S3	3.07	8.60
Total Monthly Costs	**182.61**	**240.77**

Table 4: Stage 2 Financial Viability (Monthly)

Metric	**Lower Estimate (£)**	**Upper Estimate (£)**
Revenue (23 Users @ £13)	299.00	299.0
Total Costs	182.61	240.77
Net Profit / (Loss)	**58.23**	**116.39**

Stage 3: Growth Costs (15,000 Freemium & 600 Pro Users and 1000 Institutional Pro Accounts) Costs scale with enterprise requirements, including a significant increase in marketing spend and dedicated cloud infrastructure (S3/RDS scaling).

Table 5: Stage 3 Monthly Cost Breakdown

Cost Item	**Lower Estimate (£)**	**Upper Estimate (£)**
Fixed Costs		
Auth0 Standard (20k users)	1050.00	1050.00
YouTube Ads Budget	1,000.00	1,000.00
MailChimp Ads Monthly	147.71	173.78
Variable Costs		
LLM Usage	1,506.94	3,013.88
Amazon RDS (Scaling)	390.94	955.08
Pinecone Vector DB	130.50	246.75
Amazon S3 Storage	26.40	276.00
Total Monthly Costs	**4,252.49**	**6,715.49**

Table 6: Stage 3 Financial Viability (Monthly)

Metric	**Lower Estimate (£)**	**Upper Estimate (£)**
Revenue (B2B + B2C)	15,800.00	15,800.00
Total Costs	4,252.49	6,715.49
Net Profit / (Loss)	**9,084.51**	**11,547.51**

Implementation Timeline We project Stage 1 (Development) to require approximately 6 months for technical validation and beta testing. Stage 2 (Establishment) is scheduled to align with a single full academic year, allowing us to capture a complete student lifecycle from enrolment to assessment. Finally, Stage 3 (Growth) is forecasted to scale over the subsequent two academic years as we expand to other national universities.

3.4 Validation

3.4.1 Verification of Competitor Data

To ensure our market comparison was accurate, we tested the international competitor products directly rather than just relying on their marketing. The team created accounts on all key platforms (Readwise, NotebookLM, Research Rabbit) to verify their features first-hand. By using these tools ourselves, we found specific limitations that were not immediately obvious. We also checked user reviews on academic forums to confirm that these were genuine user frustrations.

We were already familiar with the Bath library services and ensured our sources were directly from their website. When it came to CORAL we relied on their official website to conduct relevant research.

3.4.2 Financial Reliability

All infrastructure costs for Stages 2 and 3 were calculated using the official pricing pages of our vendors (Pinecone, Auth0, AWS) as of December 2025. This ensures our cost predictions are realistic. For revenue, we used official data (HESA, 2024, 2025; University of Bath, 2025*a*,*b*) to confirm the exact number of students and staff at the University of Bath, ensuring our earnings predictions were not exaggerated.

3.4.3 Software Utilisation

Figma was used to create the Business Model and Value Proposition Canvases, as it offered better design tools than standard word processors. Generative AI was used only to help format the comparison and financial tables in LaTeX to ensure they look professional; all financial calculations and numbers were performed manually by the team to ensure accuracy.

References

Abad, Z., Karras, O., Schneider, K., Barker, K. and Bauer, M. (2018), 'Task interruption in software development projects: What makes some interruptions more disruptive than others?', *arXiv preprint arXiv:1805.05508* .
URL: *https://arxiv.org/abs/1805.05508*

Bolaños, F., Salatino, A., Osborne, F. and Motta, E. (2024), 'Artificial intelligence for literature reviews: opportunities and challenges', *Artificial Intelligence Review* **57**, 259.
URL: *https://doi.org/10.1007/s10462-024-10902-3*

Chen, D.-T. V., Wang, Y.-M. and Lee, W. (2015), 'Challenges confronting beginning researchers in conducting literature reviews', *Studies in Continuing Education* **38**(1), 47–60.
URL: *https://www.researchgate.net/publication/281498502*$_C$*hallenges*$_c$*onfronting*$_b$*eginning*$_r$*esearc*

Chis, M., Li, H., Zheng, K., Lewis, M., Hughes, D. and Sycara, K. (2023), 'The cognitive load – productivity tradeoff in task switching', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* **67**(1), 666–671.
URL: *https://doi.org/10.1177/21695067231193677*

CORE (2025), 'About core'. Accessed 10 Dec. 2025.
URL: *https://core.ac.uk/about*

Hanson, M. A., Gómez Barreiro, P., Crosetto, P. and Brockington, D. (2023), 'The strain on scientific publishing', *arXiv preprint arXiv:2309.15884* .
URL: *https://arxiv.org/pdf/2309.15884*

HESA (2024), 'Higher education student statistics: Uk, 2022/23 - student numbers and characteristics'. Accessed: 2025-12-10.
URL: *https://www.hesa.ac.uk/news/11-01-2024/sb269-higher-education-student-statistics*

HESA (2025), 'Who's working in HE?'. Accessed: 2025-12-09.
URL: *https://www.hesa.ac.uk/data-and-analysis/staff/working-in-he*

Huq, F., Samee, A., Lin, D. C.-E., Tang, A. X. and Bigham, J. P. (2025), Noteeline: Supporting real-time, personalized notetaking with llm-enhanced micronotes, *in* 'Proceedings of the 30th International Conference on Intelligent User Interfaces', IUI '25, Association for Computing Machinery, New York, NY, USA, p. 1064–1081.
URL: *https://doi.org/10.1145/3708359.3712086*

Liefooghe, B., Barrouillet, P., Vandierendonck, A. and Camos, V. (2008), 'Working memory costs of task switching', *Journal of Experimental Psychology: Learning, Memory, and Cognition* **34**, 478–494.

Medical College of Wisconsin (2022), 'Cognitive load theory: A guide to applying cognitive load theory to your teaching'. Accessed: 2025-12-05.
URL: *https://www.mcw.edu/-/media/MCW/Education/Academic-Affairs/OEI/Faculty-Quick-Guides/Cognitive-Load-Theory.pdf*

Pathmonk (2024), 'Average free to paid conversion rate for saas'. Accessed: 2025-12-10.
URL: *https://pathmonk.com/average-free-to-paid-conversion-rate-saas/*

Reigstad, T. J. (1985), 'Perspectives on anxiety and the basic writer: Research, evaluation, and instruction', *Journal of Basic Writing* **4**(1), 68–77.
URL: *https://doi.org/10.37514/jbw-j.1985.4.1.06*

University of Bath (2024), 'Library guides'. Accessed 10 Dec. 2025.
URL: *https://www.bath.ac.uk/guides/library-guides/*

University of Bath (2025*a*), 'Facts and figures: Student and staff numbers'. Accessed: 2025-12-10.
URL: *https://www.bath.ac.uk/corporate-information/facts-and-figures/*

University of Bath (2025*b*), 'Faculty of science'. Accessed: 2025-12-10.
URL: *https://www.bath.ac.uk/faculties/faculty-of-science/*

4 Group Contribution Form - Group H

Group Member	Contribution	Signature
Aryan Chegini	8	
Sultan Fawad	8	
Morus Bowen	8	
Mashfiqur Thamid Miah	10	

Figure 9: Group Contribution Form

A Appendix

A.1 Subscription Tiers

Table 7: Feature Breakdown: Freemium vs. Pro Tier

Feature	**Freemium Tier (Free)**	**ScholarStack Pro (£13/mo)**
Core Access	Limited Project Containers (Allowed 3) SSO Authentication	Unlimited Project Containers SSO Authentication
Notebook System	Dual-Layer Notebooks Active Highlighting	Dual-Layer Notebooks Active Highlighting
Research Tools	Knowledge Graph Citation Network (Limited Search) Bibliography Generation	Full Knowledge Graph Expanded Citation Network Bibliography Generation
AI Capabilities	Scoped AI Companion *(Limited to 2k tokens/day)*	Scoped AI Companion *(Extended to 40k tokens/day)* Smart Paper Recommendations
Constraints	*Allowed a maximum of 30 papers per project*	*Unlimited papers per project*

A.2 User Research Data ($n = 15$)

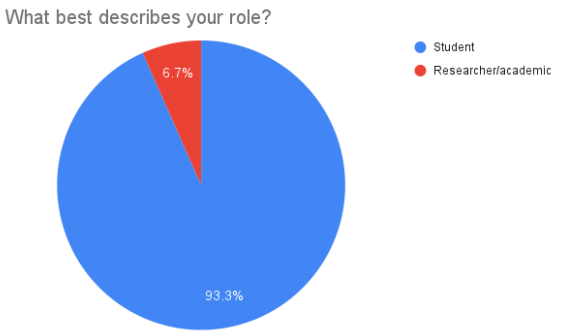

Figure A.1: What best describes your role?

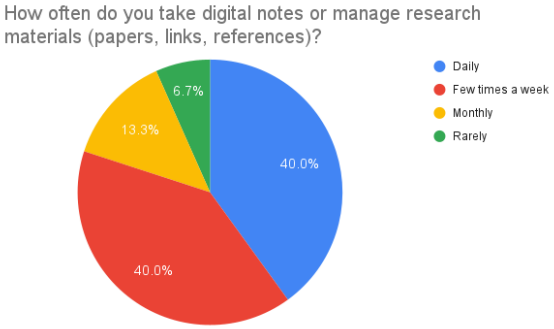

Figure A.2: Frequency of managing research materials

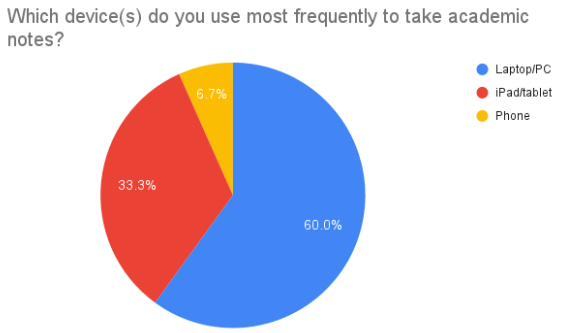

Figure A.3: Devices used for academic notes

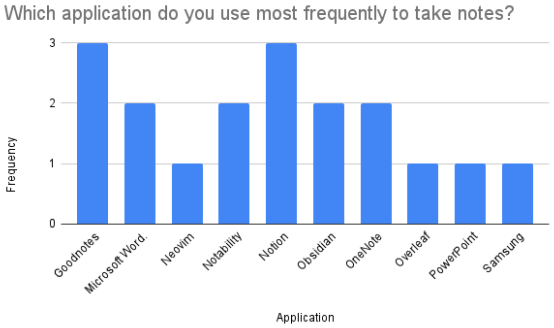

Figure A.4: Applications used for note-taking

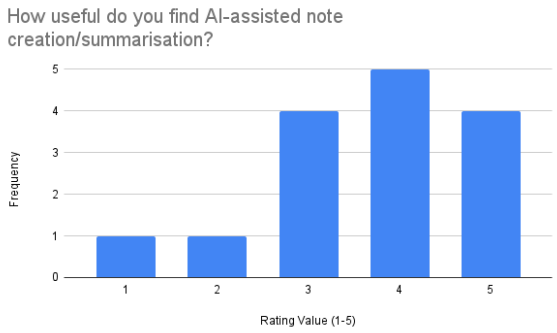

Figure A.5: Utility of AI note creation (1-5)

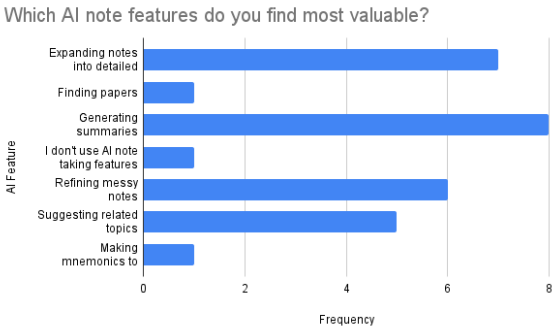

Figure A.6: Most valuable AI features

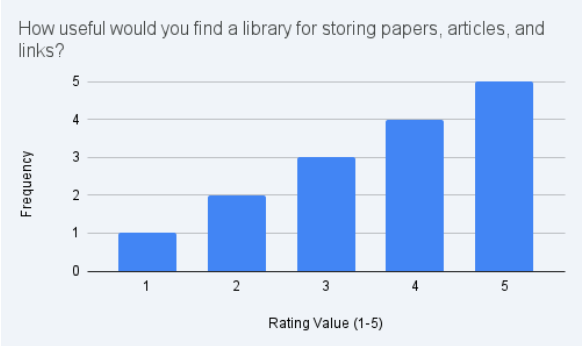

Figure A.7: Utility of a paper library (1-5)

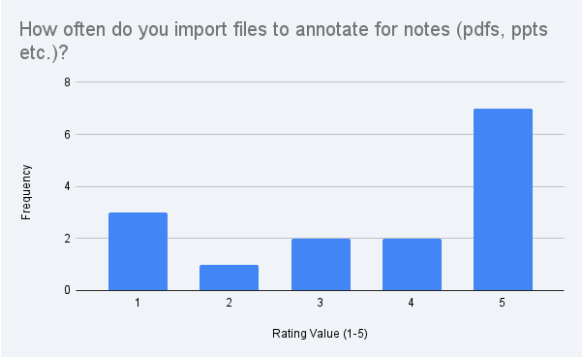

Figure A.8: Frequency of file imports