ScrumBot
AI-Powered Sprint Planning Assistant

IS 492

OM VYAS • NAKUL VASANI

GITHUB: GITHUB

DEMO: SCRUMBOT

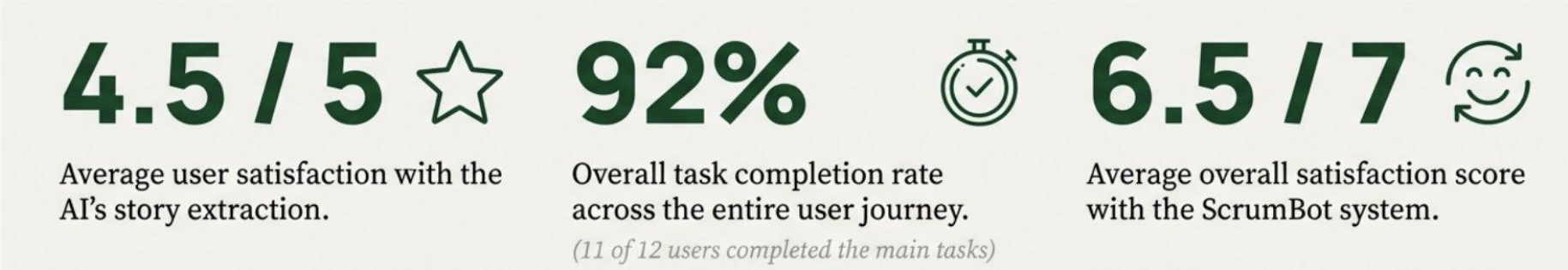
Recap & Final Goals

- PROBLEM STATEMENT

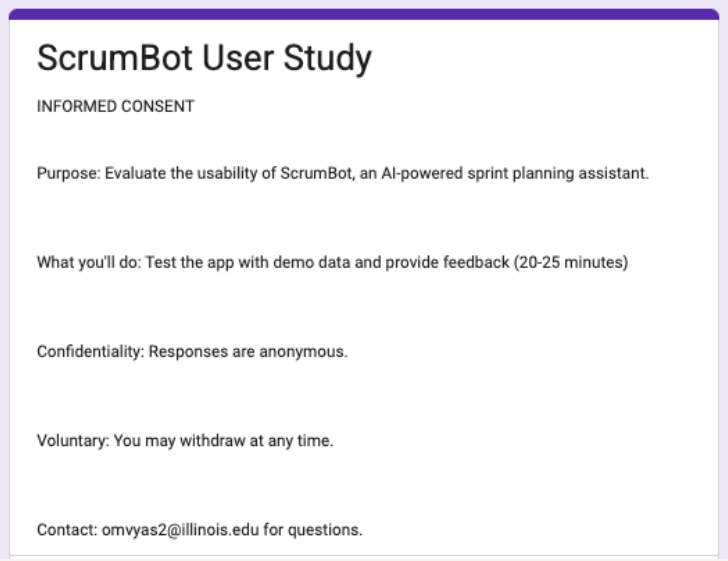
  Sprint planning is time-consuming and subjective. ScrumBot automates story extraction from meeting transcripts and uses AI + team knowledge base to recommend optimal task assignments.

- TARGET USERS

  1. Scrum Masters & Product Managers
  2. Agile development teams (2-10 members)
  3. Organizations using sprint planning



Metrics Collected



ScrumBot User Study

INFORMED CONSENT

Purpose: Evaluate the usability of ScrumBot, an AI-powered sprint planning assistant.

What you'll do: Test the app with demo data and provide feedback (20-25 minutes)

Confidentiality: Responses are anonymous.

Voluntary: You may withdraw at any time.

Contact: omvyas2@illinois.edu for questions.

Quantitative:

- ✓ Task success rate
- ✓ Time-on-task
- ✓ Error rate
- ✓ SUS score (0-100)
- ✓ UMUX-Lite (1-7)
- ✓ Satisfaction & usefulness (1-5)
- ✓ Trust in AI (1-5)

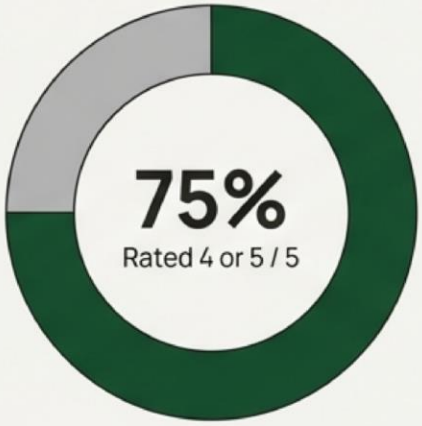
Qualitative:

- ✓ Themes from open-ended feedback
- ✓ Likes, frustrations, suggestions

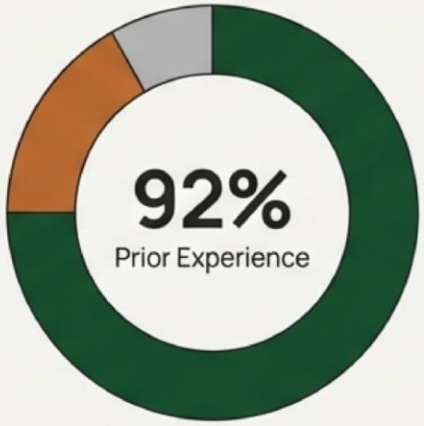
Who We Talked To: An Experienced and Relevant Group of Testers

Our study included 12 participants with a strong mix of technical, Agile, and AI tool experience, ensuring the feedback is grounded in real-world expertise.

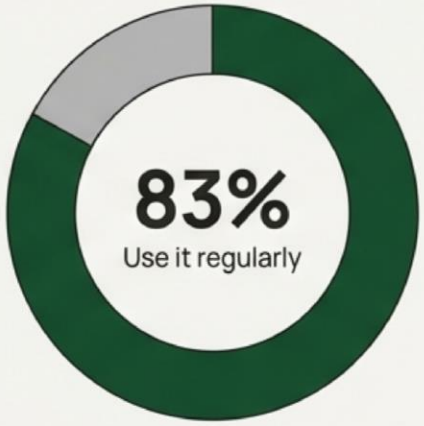
High Technical Proficiency



Deep Agile/Scrum Background

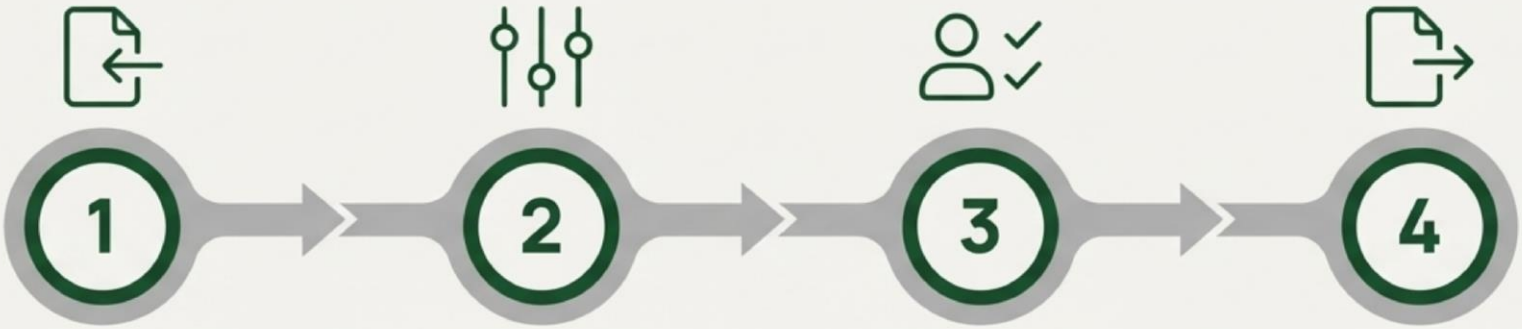


Familiar with AI Tools



The Four-Step Journey We Tested

We designed the user study to follow the product's primary workflow, from initial data input to final export. This allowed us to evaluate each critical step of the user experience.



Extract & Review

User uploads a project document. The AI processes the text and extracts user stories.

Tune Recommendations

User reviews the AI's team member recommendations and adjusts importance-weighting sliders.

Assign & Finalize

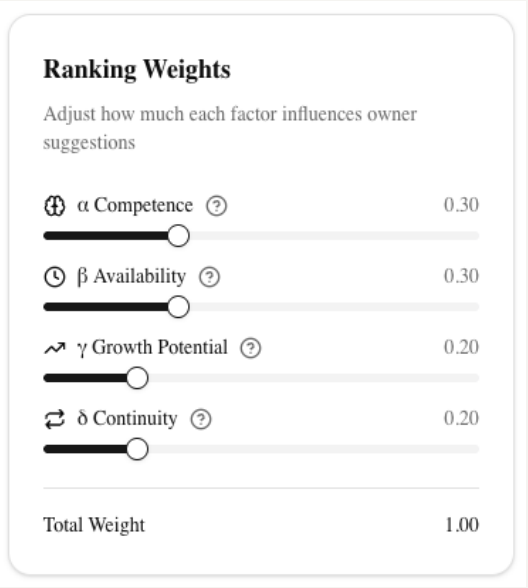
User assigns stories to specific team members based on the tuned recommendations.

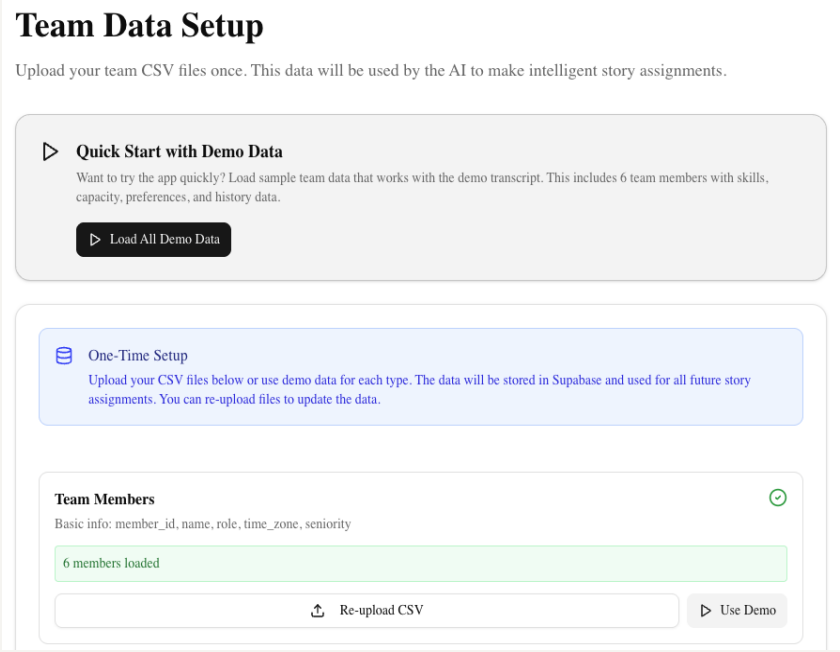
View & Export

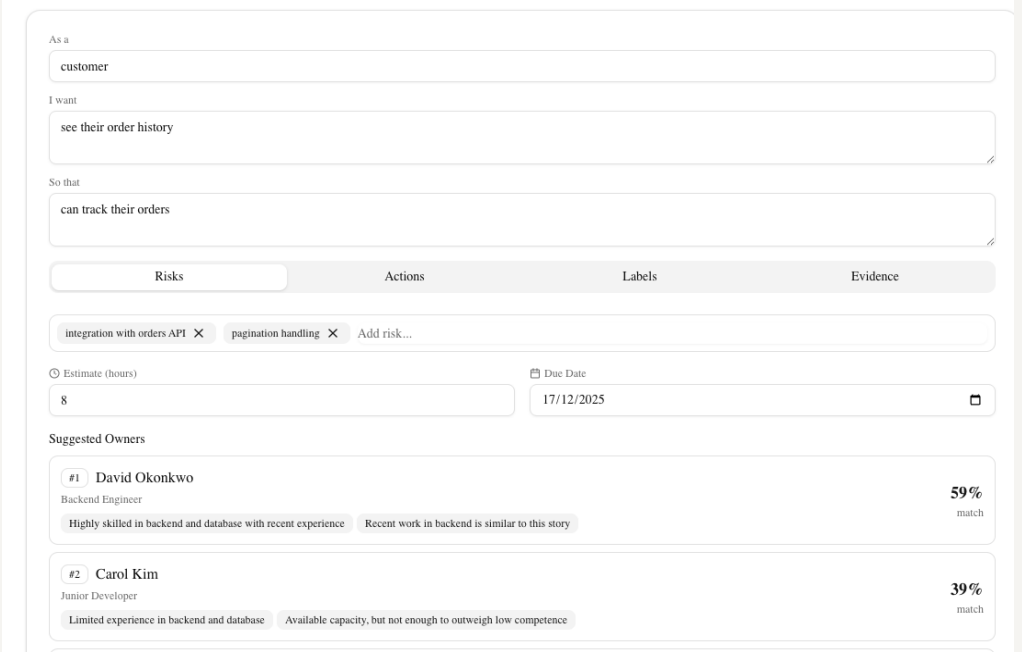
User reviews story dependencies and exports the completed sprint plan to a CSV file.

Study Materials & Protocol

- TASK PROMPTS EXAMPLE
  1. Task 1: "Process the demo transcript and review extracted user stories"
  2. Task 3: "Adjust weight sliders to prioritize 'Availability', then assign 2 stories"
- CONSENT & PROCEDURE
  1. Participants provided informed consent (anonymous, voluntary)
  2. Remote, unmoderated study via Google Form
  3. Study flow: Intro (2 min) → Tasks (20 min) → Survey (15 min)







- [1] Extract & Review
- [2] Tune Recommendations
- [3] Assign & Finalize
- [4] View & Export

Journey Step 1: Story Extraction is Fast, Seamless, and Highly-Rated

Users had an overwhelmingly positive first impression. The automated extraction process was successful and met or exceeded their expectations for quality and speed.

- **11 out of 12 users** successfully completed the task.
- **4.5 out of 5** average rating for 'How well did the AI extract user stories?'
- **4.7 out of 5** average rating for satisfaction with the results.
- **Clear Progress Indicators:** 11 out of 12 users confirmed they could see clear progress indicators during processing, which builds trust and reduces uncertainty.

[1] Extract & Review

[2] Tune Recommendations

[3] Assign & Finalize

[4] View & Export

Journey Step 2: The AI's Logic is Clear and Trusted, But UI Feedback is Too Subtle

This is the core of the user interaction. While users understood and trusted the AI's reasoning, the interface did not provide immediate, clear feedback when they adjusted the recommendation weights.

Validated Strengths

High Marks for Clarity & Trust

4.6 / 5 average rating for clarity of the 4 score components (Competence, Availability, etc.).

4.7 / 5 average rating for how helpful the AI's justifications were.

4.6 / 5 average rating for trust in the AI's recommendations based on those justifications.

Actionable Opportunities

42%

of users (5 of 12) explicitly described the recommendation updates as "subtle changes."

25%

of users (3 of 12) found the weight sliders "confusing" to use, likely linked to the lack of clear feedback.

Deep Dive: Why ‘Subtle Changes’ Undermine User Confidence

The lack of strong visual feedback after adjusting a slider is the single biggest point of friction in the current user experience. When users perform an action, they expect a clear and immediate reaction from the system. Subtlety here is perceived as unresponsiveness or a bug.

> *“Yes, subtle changes”*

Impact Analysis

- **Erodes Confidence:** Users are unsure if their adjustments had any effect.
- **Slows Workflow:** Users may over-adjust or re-adjust sliders, trying to provoke a more obvious system response.
- **Masks AI Power:** The impressive underlying calculations of the AI are hidden if the user can’t perceive the result of their input.

- [1] Extract & Review
- [2] Tune Recommendations
- **[3] Assign & Finalize**
- **[4] View & Export**

Journey Steps 3 & 4: The Final Steps Are Sound, But Reveal Minor Hurdles

Users were able to complete the final tasks of assigning stories and exporting their plan. However, minor usability issues in both steps indicate opportunities for simplification and bug fixing.

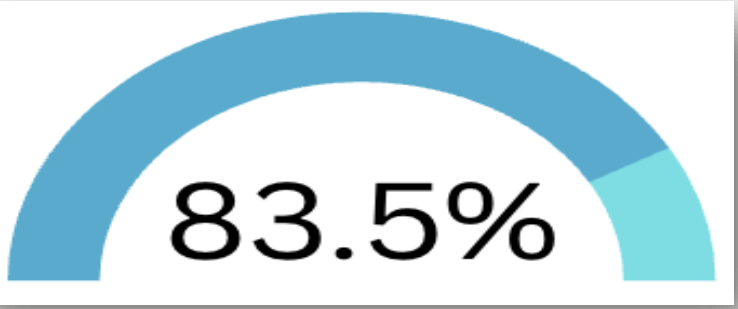
Task 3: Multi-Assignment Is Not Intuitive for Everyone

Insight: The workflow to assign more than one story may be a point of friction or lack discoverability.

Task 4: Export Function Works, But Needs Reliability Check

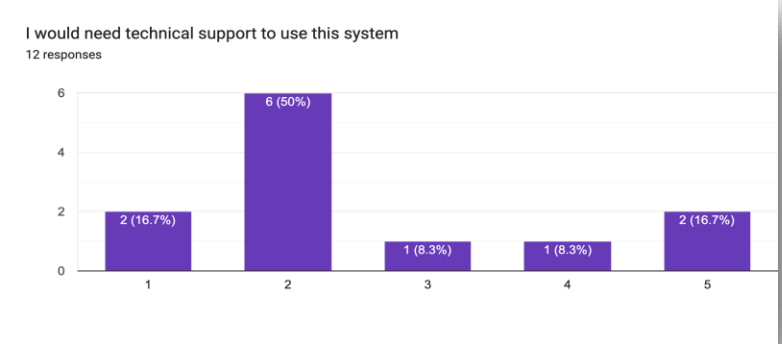
83% (10 of 12 users)
exported the CSV successfully.

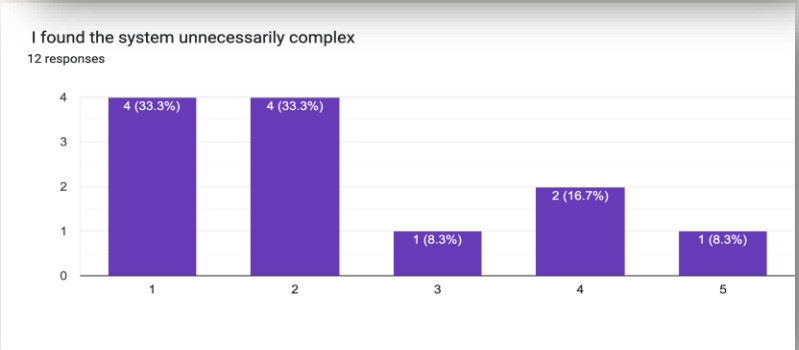
However, **2 users** reported that while the system confirmed the export, the file **"didn't download,"** indicating a potential bug.

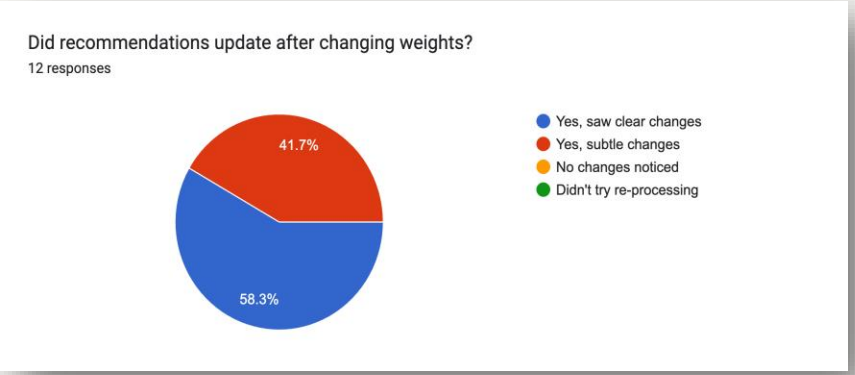


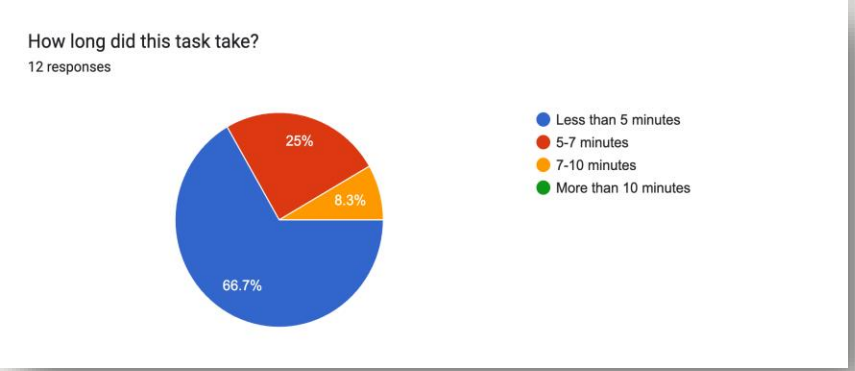
Mean SUS Score: 83.5 ± 15.8

	Expected	Actual	Difference
Task 1:	5 min	2.5 min	-50%
Task 2:	5 min	3.0 min	-40%
Task 3:	7 min	5.2 min	-26%
Task 4:	5 min	3.2 min	-36%









What frustrated you MOST?

1 response

This form

If you could change ONE thing, what would it be?

1 response

The name of the app

Any other comments or suggestions?

1 response

Please try to make feedback forms less boring

The Big Picture: ScrumBot Scores High on Usability and Value

Beyond individual tasks, we asked users about their overall perception of the system. The feedback confirms that ScrumBot is seen as easy to use, well-integrated, and valuable.

I found the system unnecessarily complex: Average **2.3 / 5** (Low is good)

Limitations, Risks & Ethics

STUDY LIMITATIONS

- Small sample (N=12)
- Homogeneous participants (students 22-28)
- Artificial environment (demo data)
- Self-reported measures

SYSTEM RISKS

- AI hallucinations possible
- Potential bias in recommendations
- Privacy concerns with transcripts
- Over-reliance on AI decisions
- Free API created token bottleneck

ETHICAL SAFEGUARDS

- Informed consent obtained
- Anonymous data collection
- Transparent AI justifications
- Human override capability (Human in the loop)
- Advisory role (not prescriptive)

Synthesized Findings: Validated Strengths & Actionable Opportunities

The study provides a clear mandate: double down on the powerful and well-received AI core, and focus our design and engineering efforts on refining the key user interactions identified in the journey.

What We Should Protect & Amplify

- **Core AI Performance:** Story extraction is fast, accurate, and highly rated.
- **AI Explainability:** Justifications and score components are clear and build significant user trust.
- **Overall Simplicity:** The system is perceived as easy to learn and not cumbersome.
- **Task Visibility:** Progress indicators during processing are effective and appreciated.

Where We Should Focus & Refine

- **UI Feedback Loop:** The #1 priority is providing stronger visual feedback when users adjust recommendation weights.
- **Assignment Workflow:** The multi-assignment process needs to be more discoverable and intuitive.
- **Minor Bugs:** The export download failure and an issue where onboarding didn't appear for one user need to be addressed.
- **Onboarding Adoption:** Two users skipped the tutorial, suggesting we could explore ways to better integrate it into the initial experience.

Recommendations: Focus on UI Refinement to Match the Core AI's Strength

Based on the user study, we recommend prioritizing targeted UI and UX enhancements in the next development cycle.

1. **Immediately Prioritize UI Feedback for Sliders.**
   Action: Implement more prominent visual cues (e.g., highlights, value changes, animations) to confirm that recommendation scores have updated after a user adjusts a weight. This is our highest-impact, highest-priority fix.

2. **Redesign the Multi-Assignment Workflow.**
   Action: Conduct a design sprint to re-evaluate the UX for assigning multiple stories. The goal is to improve discoverability and reduce user friction.

3. **Address Technical Bugs.**
   Action: Create engineering tickets to investigate and resolve the "export didn't download" bug and the inconsistent appearance of the onboarding tutorial.

4. **Defer Changes to the AI Model.**
   Action: No immediate changes are needed for the AI extraction or recommendation logic. The underlying models and justifications tested exceptionally well with users.

Acknowledgments & Contributions (For Grading - Not Presented)

INDIVIDUAL RESPONSIBILITIES

- Om Vyas – AI/RAG implementation; Backend API development; User study design & analysis
- Nakul Vasani – Frontend UI/UX development; State management (Zustand); CSV export & demo data

TOOLS & AI RESOURCES

- Development: Groq API (Llama 3.3 70B); Next.js, React, TypeScript; Vercel AI SDK, shadcn/ui
- Study & Analysis: Google Forms & Sheets; Claude for study protocol
- Disclosure: AI tools used for boilerplate code, documentation, debugging; All content reviewed by team; Core algorithms designed by team