# PodShot – Product Requirements Document

Authors – Ishita Rastogi, Divya Gupta

## 1. Product Overview & Origin

**PodShot** began as our own little lifesaver: a Chrome extension to capture, automatically transcribe, and distill those "aha" moments when watching videos to learn, without fumbling for a notebook. It extracts and organizes clipped video links, followed by their transcripts and key insights so we can stay immersed in learning. Currently, it's our companion on YouTube, but its heart lies in supporting any video from any platform, making it invaluable for knowledge workers, students, researchers, and content creators.

## 2. Motivation & Early Validation

- **Personal Pain Point:** Keeping up with fast-paced webinars, podcasts, and tutorials meant we were constantly hitting pause to scribble notes or grab screenshots, only to rediscover later that they'd lost their context or been forgotten entirely. That stop-start rhythm only dragged out our viewing time, but without those scattered notes, we'd struggle to recall the very insights we'd worked so hard to capture.
- **Circle-of-Trust Testing:** I rolled PodShot out to five peers, who used it during real study sessions. Feedback was unanimous: "This saves me 10+ minutes per video and keeps my notes in one place."

## 3. Problem Statement

Watching videos to learn and upskill means constantly pausing to take notes or screenshots, only to find later they've lost context or been forgotten. Effective note-taking demands as much focus as uninterrupted learning yet focusing on notes pulls you out of the flow, making the whole process feel clunky and inefficient.

## 4. Vision & Success Criteria

- **Vision:** Make video note-taking feel as effortless as hitting "play" or "pause".
- **MVP Success:**
    - **Instant Context Recall:** Each clip includes a timestamped link, transcript, and AI-generated summary for quick reference.
    - **Speed:** Capture-to-insight under 30 seconds
    - **Clarity:** Summaries that can be rated ≥ 4/5 for usefulness

## 5. Core Features

### i. Clip Capture

- Users can mark the start and end of a video segment while watching, and the extension records the video URL, start time, and end time.

### ii. Automated Transcription

- The extension sends the video segment info to a backend.
- The backend downloads the segment, extracts audio, and transcribes it using OpenAI Whisper (or similar).
- The transcript is returned to the extension and stored locally.

**iii. AI-Powered Insight Generation**

- Users click "Generate Insights!" to send the transcript to the backend.
- The backend uses OpenAI GPT to summarize the transcript and extract key insights.
- Insights are returned to the extension.

**iv. Google Docs Integration**

- For each new video, PodShot auto-generates a Google Doc named after the video, then appends every captured clip to it and saves it in your Drive
- Captured clips' links, their transcripts (and optionally the insights) are appended to a connected Google Doc.
- OAuth2 is used for secure Google Docs access.
- The extension ensures content is appended at the correct location in the document.

**v. Persistent State & Feedback**

- The extension stores the latest transcript and clip data in Chrome's local storage.
- Users receive clear status messages and error alerts throughout the workflow.

**vi. Robust Error Handling**

- Handles backend/API failures, quota issues, and Google Docs API errors gracefully.
- Logs errors and provides actionable feedback to users.

---

## 6. Technical Snapshot

- **Frontend:** Chrome Extension (popup UI + background scripts)
- **Backend:** FastAPI server (handles transcription + summarization)
- **APIs:**
  - OpenAI (Whisper and GPT)
  - Google Docs API
- **Authentication:** OAuth2 for Google Docs, environment variables for API keys
- **Storage:** Chrome local storage for state persistence

---

## 7. User Journey

- User opens a video and launches PodShot.
- Clicks "Start Capturing" and "End Clip" to define a segment.
- User then clicks (or not) "Generate Insights!" to summarize the transcript.
- Opens Google drive and locates the document named after the video.
- Finds a clickable video link before each transcript, allowing the user to start playback from the exact captured start timestamp.

- Will see each section displaying the video link first, followed by its corresponding transcript, and then the generated insights.

---

## 8. Technical Workflow

The technical workflow of PodShot is designed to ensure smooth, modular processing across frontend capture, backend processing, and final content delivery. We have developed an n8n-based automation workflow that will further streamline these processes and reduce bugs or manual errors, though it has not yet been integrated into the live automation flow.

**• Current Workflow:**

- Step 1: Chrome extension captures video URL, start/end timestamps.
- Step 2: Frontend sends capture details to backend FastAPI server.
- Step 3: Backend downloads the video segment, extracts audio, runs transcription via OpenAI Whisper (or similar).
- Step 4: Backend generates a unique document ID and checks if it already exists in the Drive, if it does, it decides to append it, otherwise generates a new doc.
- Step 5: Backend processes transcript through OpenAI GPT to generate summarized insights.
- Step 6: Final transcript + insights are sent back to the frontend and written to the associated Google Doc using the Google Docs API.
- Step 7: Chrome local storage maintains session state, error logs, and user feedback.

**• Planned n8n Integration :**

- We have developed an n8n automation workflow that can orchestrate and monitor each step of the backend process, including:
    - Automatic job queuing and retry for transcription and summarization tasks.
    - Monitoring API quota limits and dynamically routing tasks if primary services fail.
    - Logging detailed error states into a central dashboard for faster debugging.
    - Notifying users (via extension or email) when a process completes, fails, or is delayed.
- Integration Impact (Post-Implementation):
    - Reduced manual oversight and error handling at the backend.
    - Enhanced visibility into system performance and task completion.
    - Fewer errors arising in other steps when a single step is modified or enhanced.
    - Faster identification and resolution of API or service-level issues.
    - Improved user experience through more consistent and reliable automation.

This future integration with n8n will act as a safeguard layer, ensuring that as we scale PodShot's functionality (including support for additional video platforms), the automation remains robust, resilient, and maintainable.

---

## 9. PodShot Competitive Analysis & Market Research (2025)

Direct Competitors (Highly Relevant):

| Competitor | What They Do | Strengths | Weaknesses | Pricing |
|---|---|---|---|---|
| **Snipd (Main)** | AI-powered podcast and YouTube clipper; transcribes and bookmarks-highlights with timestamps; syncs notes to Notion, Readwise | Excellent clip capturing, timestamped highlights, podcast and YouTube support, Mobile-friendly | Primarily podcast-focused; only applicable on videos hosted on the app or YouTube videos | Free plan available; Premium at $4.99/month |
| **Eightify** | AI summaries of YouTube videos (Chrome/Safari, Mobile) | Handles long videos, multilingual support, mobile apps | Bullet-point summaries only, paid after trial | Free trial, then $6.99/month |
| **NoteGPT** | Summarizes YouTube videos and provides mindmaps and Q&A | Deep learning options, mobile apps | Limited free usage | Free up to 15 summaries/month; Paid from $2.99/month |
| **TubeOnAI** | Summarizes YouTube videos and articles, organizes dashboards | High volume capacity, cross-content support | Higher priced | Free up to 60 minutes/month; Paid starts at $19/month |
| **ReClipped** | Manual video clipping and annotation | Timestamped notes, exports to Notion/Obsidian | No AI summarization; manual effort required | Free |
| **YT Notes** | One-click AI-generated summaries for YouTube lectures | Lightweight, no login required | 5 free summaries/day limit | Free (option to tip) |
| **ZippyNotes** | Real-time timestamped notes & AI summaries | Organized note system, tagging features | Basic feature set | Free |

## 10. Market Research Insights

### Target Audience

- Students using YouTube, online lectures, MOOCs

- Researchers and academics capturing conference or lecture insights
- Professionals watching recorded webinars or training sessions
- Content creators transforming video content

**Key Trends**

- Video-based learning is the dominant format in online education.
- Growing demand for instant, reliable, structured AI takeaways.
- AI summarization tools are rapidly becoming standard.
- Seamless integration into workspaces (Google Docs, Notion) is critical.

**Opportunity Gaps**

- Most tools either capture a clip or summarize the whole video, but few integrate both seamlessly.
- Few offer real-time timestamped clip capture, transcription, and AI insights in one flow.
- Strong mobile solutions are still rare and fragmented.

---

## 11. How PodShot is Different

| Aspect | PodShot | Competitors |
|---|---|---|
| Clip-based capture | User-defined start/stop timestamps on videos | Most of them offer full-video summarization |
| Automatic transcription | Each clip individually transcribed | Typically full video transcription only |
| AI Insight Generation | Summarized insights for each captured clip forming notes | Summaries often for entire videos |
| Google Docs Integration | Automatic document creation and organized storage | Manual copy/export or limited integrations |
| Fast Workflow | Clip-to-Doc completed in under 30 seconds and 2-3 clicks | Slower or multi-step manual workflows |
| Cross-platform Scalability | Expanding to Vimeo, Spotify, and webinars | Primarily YouTube-focused tools |
| Mobile Capture Vision | Mobile app planned for video learning capture | Limited mobile support among competitors |

## 12. Success Metrics

- Reduced time to capture and summarize video content.
- Accurate transcripts and insights generated and appended to correct Google Docs.
- User satisfaction and retention.
- Low error rate and high reliability.

---

## 13. Next Steps or Future Enhancements

- Working functionality on additional video platforms like Spotify and Pendo.

- Enhanced UI (Smarter, more organized documents with data automatically structured into points, tables and other easy-to-read formats).
- A mobile friendly solution to work for video streaming apps as well.