MedGemma Demo Video Script

Enso Atlas Demo Video Script

Duration: 3 minutes (maximum) **Format:** Screen recording with voiceover
Recording notes: Highlight key interactions with cursor movements

---

0:00-0:20 | Hook and Problem Statement

Visual

- Open on a hospital pathology lab setting (stock footage or static image)
- Transition to WSI thumbnail grid

Voiceover

> "Every year, over 20,000 women in the United States are diagnosed with ovarian cancer. For many, the critical question becomes: will this treatment work for me?
>
> Today, predicting response to therapies like bevacizumab requires extensive molecular testing, takes weeks, and often provides uncertain answers.
>
> What if we could look at a standard pathology slide and predict treatment response in minutes, with explainable evidence?"

Key Points

- 20,000+ annual ovarian cancer diagnoses (US)
- Current challenge: treatment response prediction is slow, expensive, uncertain
- Pathology AI offers a faster, evidence-based alternative

0:20-0:50 | Solution Overview

Visual

- Show Enso Atlas landing page / main interface

- Pan across the clean clinical UI
- Briefly highlight the three main panels: Viewer, Evidence, Report

Voiceover

> "Introducing Enso Atlas: an on-premise pathology evidence engine for treatment response insight.
>
> Three key differentiators:
>
> First, evidence-based predictions. Every result comes with visual heatmaps showing exactly which tissue regions drove the prediction, plus similar cases from our reference cohort.
>
> Second, local-first architecture. All inference runs on-premise. No patient data ever leaves your hospital network.
>
> Third, structured clinical reports generated by MedGemma, Google's medical AI model, providing tumor board-ready summaries."

Key Points

- Introduce "Enso Atlas" name clearly
- Evidence-based: heatmaps + similar cases
- Local-first: HIPAA-friendly, no PHI exfiltration
- MedGemma: structured, clinician-ready reports

0:50-1:40 | Live Demo - Analysis Workflow

Visual

- Screen recording of the actual application
- Show each step with deliberate cursor movements

Voiceover

> "Let me show you how it works with a real case from the TCGA ovarian cancer cohort.
>
> [Click on slide selector] Here we have a diagnostic H&E whole-slide image. Using OpenSeadragon, we can navigate the slide just like a digital microscope - zoom in on regions of interest, pan across the tissue.
>
> [Click Analyze button] When we run analysis, Enso Atlas processes the slide through our pipeline. First, we tile the slide into patches and generate embeddings using Google's Path Foundation model - a vision transformer trained on millions of pathology images.
>
> [Show loading, then results appear] The prediction appears with a confidence score. In this case, we see a high likelihood of bevacizumab response.

[Hover over heatmap] But here is where it gets interesting. This heatmap overlay shows attention weights - the regions our model focused on most. Brighter areas contributed more to the prediction. This is not a black box.

[Scroll to similar cases panel] We also retrieve the most similar patches from our reference cohort of 288 slides from 78 patients. Clinicians can see: this tissue pattern has been seen before, and here is how those patients responded."

Key Points

- Load TCGA ovarian cancer slide
- Demonstrate WSI navigation (zoom, pan)
- Run analysis - show progress
- Display prediction score with confidence interval
- Highlight evidence heatmap - attention visualization
- Show similar cases panel with retrieval results

Timing Breakdown

- 0:50-1:00: Load slide, navigate WSI
- 1:00-1:10: Click analyze, explain pipeline
- 1:10-1:25: Show prediction result
- 1:25-1:40: Heatmap and similar cases

1:40-2:20 | Report Generation

Visual

- Click Generate Report button
- Show report loading/generation
- Scroll through report sections
- Demonstrate print/export

Voiceover

"For tumor board presentations and clinical documentation, Enso Atlas generates structured reports.

[Click Generate Report] With one click, we call MedGemma - Google's 4-billion parameter medical language model - to synthesize a clinical summary.

[Report appears] The report includes structured sections: patient context, prediction summary, evidence description, similar case analysis, and clinical considerations.

[Scroll through sections] Notice the structured format. This is designed for integration into electronic health records and tumor board discussions.

> [Click export/print] Reports can be exported as PDF or printed directly for clinical use."

Key Points

- One-click report generation
- MedGemma 4B generates the summary
- Structured sections for clinical workflow
- EHR-ready format
- Print and export capability

Timing Breakdown

- 1:40-1:50: Click Generate Report, brief explanation
- 1:50-2:05: Show report sections
- 2:05-2:20: Export/print demonstration

2:20-2:50 | Technical Highlights

Visual

- Architecture diagram overlay (from README)
- Show terminal/logs briefly (optional)
- Hardware shot (DGX Spark or server rack image)

Voiceover

> "Under the hood, Enso Atlas combines state-of-the-art foundation models.
>
> For embeddings, we use Google's Path Foundation - a vision transformer producing 384-dimensional representations trained on diverse pathology data. This gives us robust, transferable features across tissue types and staining protocols.
>
> For classification, we use CLAM - a gated attention multiple instance learning architecture that handles the gigapixel scale of whole-slide images while providing interpretable attention weights.
>
> For retrieval, FAISS enables sub-second similarity search across our entire reference cohort.
>
> For reporting, MedGemma 1.5 4B generates structured clinical text with built-in safety guardrails.
>
> Critically, everything runs locally. We deploy on NVIDIA DGX Spark for edge inference - no cloud, no PHI leaving the hospital. A typical slide processes in under two minutes."

Key Points

- Path Foundation: 384-dim embeddings, robust feature extraction
- CLAM MIL: attention-based classification, interpretable
- FAISS: fast similarity search
- MedGemma 1.5 4B: structured reports with safety guardrails
- Local inference: no PHI leaves the hospital
- Edge-ready: runs on DGX Spark
- Processing time: under 2 minutes per slide

Timing Breakdown

- 2:20-2:30: Foundation models overview
- 2:30-2:40: Pipeline components
- 2:40-2:50: Local/edge deployment, timing

2:50-3:00 | Closing

Visual

- Return to main interface
- Team attribution slide
- Call to action / contact info

Voiceover

> "Enso Atlas brings evidence-based, explainable AI to pathology treatment response prediction.
>
> Built for the MedGemma hackathon by Hanson Wen.
>
> The future of precision oncology is local, fast, and transparent."

Key Points

- Reiterate impact: evidence-based, explainable, local
- Attribution: team name, hackathon context
- Closing statement

Production Notes

Recording Setup

- Resolution: 1920x1080 minimum, prefer 4K
- Frame rate: 30 fps

• Audio: Clear voiceover, no background music during demo sections

Screen Recording Tips

• Use a clean browser window (hide bookmarks bar, extensions)
• Pre-load the demo slide to avoid loading delays
• Practice the workflow 2-3 times before recording
• Move cursor deliberately - viewers need to follow

Voiceover Options

1. Human voice (preferred for authenticity)
2. TTS with natural-sounding voice (ElevenLabs or similar)
3. On-screen text with background music (no narration)

Editing Checklist

☐

Trim dead air and loading times

☐

Add subtle zoom effects on UI elements being discussed

☐

Include cursor highlight/spotlight effect

☐

Add section title cards for transitions

☐

Verify total runtime does not exceed 3:00

Assets Needed

☐

Enso Atlas frontend running locally

☐

Pre-processed TCGA slide with embeddings

☐

Architecture diagram (use README version or create cleaner graphic)

☐

Team/attribution slide

Transcript Summary

Full narration word count: approximately 650 words Speaking rate: 150 words/minute = 4.3 minutes of narration

Action required: Trim narration by approximately 30% to fit 3-minute constraint. Prioritize the live demo section; technical details can be condensed.

Condensed Version Option

If timing is tight, use this condensed script for the technical section (2:20-2:50):

> "Technically, we combine Path Foundation for embeddings, CLAM for attention-based classification, and MedGemma for report generation. Everything runs locally on NVIDIA DGX Spark - no cloud, no data leaving the hospital. Processing takes under two minutes per slide."

This reduces the technical section from 30 seconds of dense content to 15 seconds of key highlights.