SyncSense

The Music Supervisor Agent

Turn prompts into feelings: emotion → score → video, in sync.

Powered by Leonardo

Meet the Team

Emily Donovan

Creative Director & AI Art Lead

A filmmaker, producer, and marketing lead with a background in "grieftech" research, Emily brings a unique perspective to AI-assisted storytelling. She designs cinematic, story-first visuals and sound, building robust style bibles and prompt systems. Emily oversees the entire creative process from concept to final distribution, ensuring emotional coherence, brand alignment, and ethical integrity.

- **Creative Vision**

  Narrative clarity, aesthetic alignment, creative producing, finishing craft, and audience growth.

- **Lead Responsibilities**

  Treatment, style bible, Leonardo look-dev, shot ordering to music grid, editorial finish, social cutdowns, premiere + PR.

Josh Rauvola

AI Engineer & Research Lead

An AI engineer and researcher experienced in shipping production systems in regulated environments. Josh leads the technical development of SyncSense, focusing on emotion modeling, music generation/retrieval, sync logic, and green-compute practices. His expertise translates complex research into reliable, high-impact creative tools that deliver on time.

- **Technical Expertise**

  Robust AI pipelines, emotion curve taxonomy, safety/evaluation discipline, and performance-minded engineering.

- **Lead Responsibilities**

  Emotion extraction/curve modeling, music gen or retrieval, sync engine, evaluation and safety, provenance and compute efficiency.

AI Video Looks Impressive — But Often Feels Flat

The Missing Piece

Current AI video tools create stunning visuals, but they're missing something fundamental: **emotionally aligned sound**. Music and sound design are the invisible threads that weave visual storytelling into an emotional experience.

Existing platforms don't bind emotion → music → video in a controllable, integrated loop. The result? Beautiful visuals that fail to land emotionally, leaving audiences impressed but unmoved.

The Gap

No tools connect emotional intent to musical score to synchronized video generation

The Impact

Gorgeous AI-generated imagery that feels hollow, disconnected, forgettable

The Opportunity

Transform AI video from technical novelty into genuine cinematic storytelling

Why Sound Is Half the Story

50% of Impact

Sound accounts for approximately half of a scene's emotional and cinematic impact — yet it's consistently overlooked in AI video workflows.

Musical Architecture

Tempo, key signature, dynamics, and melodic motifs fundamentally alter how audiences perceive meaning, pacing, and emotional truth.

Unlocking Story

A thoughtfully chosen temp track can reveal story beats, pacing opportunities, and emotional transitions that picture alone cannot communicate.

Every filmmaker knows the transformative power of the right music cue. Horror becomes comedy. Tension becomes release. The same footage tells completely different stories depending on the score beneath it. This is the principle SyncSense brings to AI video generation — intentional emotional design from the very first prompt.

The Power of Sound: One Visual, Three Emotions

The same visual footage can evoke vastly different emotions and narratives based solely on the accompanying sound. Observe how subtle changes in the soundscape can redefine the entire context of a scene.

▶ 00:08
▶ YouTube ⬈
Quiet

▶ 00:09
▶ YouTube ⬈
Funny

▶ 00:09
▶ YouTube ⬈
Scary

Scene 1: No Sound

A visually clear scene of a person walking through a park. Without any accompanying sound, the moment feels flat, clinical, and devoid of emotional resonance, leaving the viewer detached.

Scene 2: Wrong Sound

The same visual, but paired with an utterly inappropriate sound—perhaps a cheerful, cartoonish melody or a bizarre noise. The experience becomes jarring and comical, undermining the original visual intent.

Scene 3: Correct Sound

With an unsettling, suspenseful, or scary soundtrack, the benign park instantly transforms into a place of foreboding. The viewer's perception shifts dramatically, driven purely by the auditory input.

Our Solution: The Music Supervisor Agent

SyncSense acts as an intelligent music supervisor embedded directly into the AI video generation pipeline, orchestrating emotion, sound, and image into a unified creative vision.

Prompt Analysis

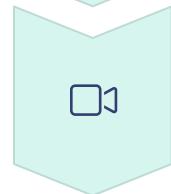Reads creative brief and extracts emotional intent

Emotion Mapping

Generates valence/arousal curve with descriptive tags

Score Selection

Generates or retrieves perfectly matched music cue or generates new music to match

Synced Video

Guides Leonardo to generate visuals aligned to score dynamics

The result is a closed-loop system where emotion drives every creative decision — from the musical architecture to the visual pacing, transitions, and compositional choices. No more beautiful but emotionally flat AI videos. Just genuine, resonant storytelling.

Why Leonardo Is Essential

Leonardo isn't just another tool in the pipeline — it's the visual foundation that makes emotion-driven generation possible. Its unique capabilities enable rapid iteration, stylistic consistency, and precise control over the visual language that complements our musical architecture.

- **Rapid Look Development**

  Generate cohesive visual palettes, recurring motifs, and style consistency across all shots in minutes instead of days

- **Canvas Compositing**

  In-paint and out-paint capabilities allow frame-perfect composites aligned to the musical beat grid for synchronized storytelling

- **Professional Finishing**

  Alchemy and Upscale features deliver clean, broadcast-ready finals optimized for both editorial workflows and social distribution

- **Iterative Testing**

  Batch render scene plates to quickly test multiple emotional curves, allowing rapid creative exploration and refinement

Tech Stack: Lean & Practical

Our production pipeline leverages best-in-class tools across generation, audio post-production, and finishing — all while maintaining lean operations and responsible AI practices.

AI Generation (Audio + Visual)

- **Leonardo:** Primary image generation and Canvas composites
- **ControlNet/IP-Adapter:** Style and character consistency across shots
- **Udio/Suno:** Temporary music cues for rapid prototyping

Audio Production

- **Ableton/Logic Pro:** Music composition and arrangement
- **iZotope RX:** Audio restoration and cleanup
- **Adobe Audition:** Final mix and mastering

Post-Production

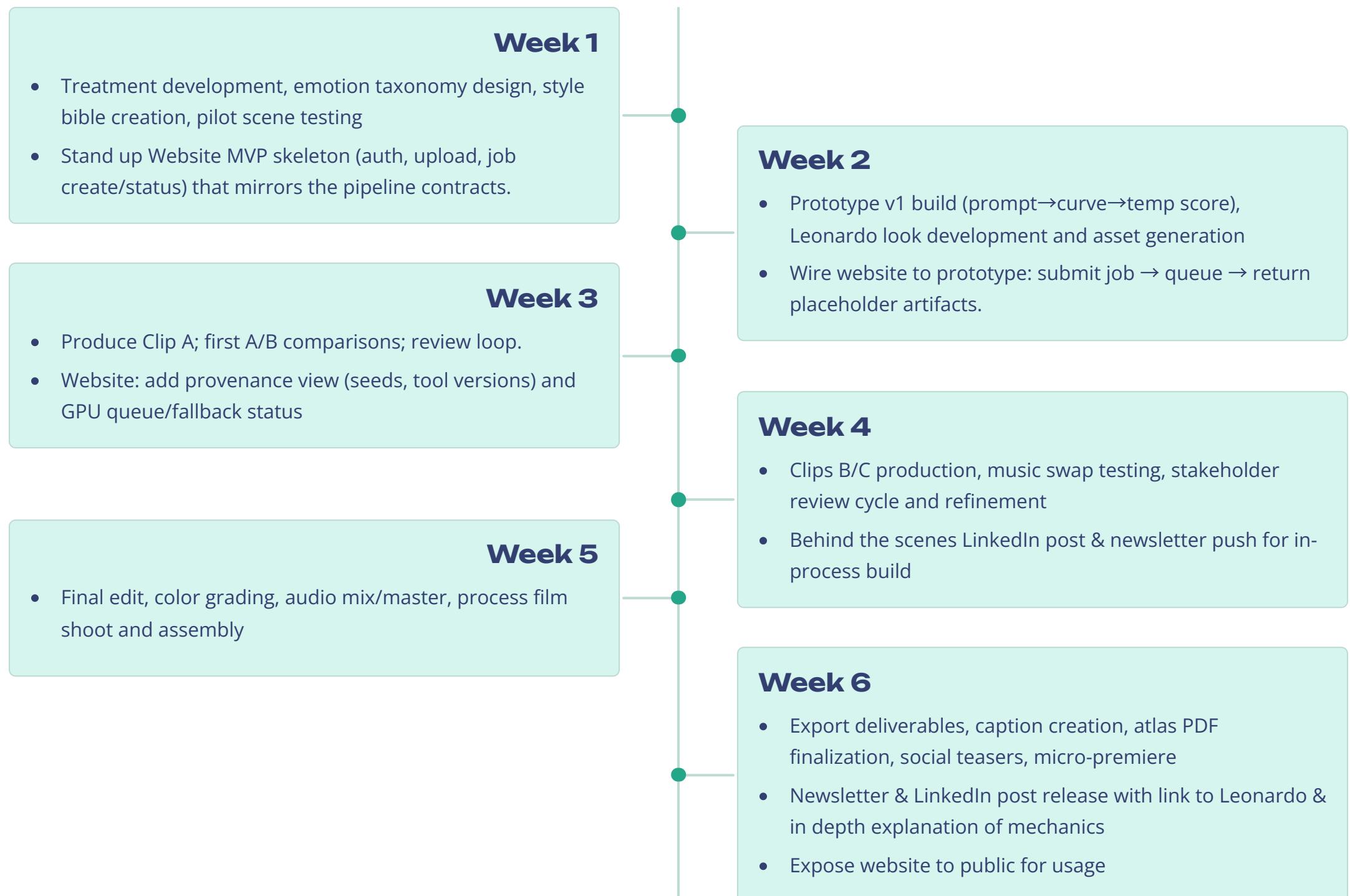- **Premiere Pro:** Editorial and assembly
- **After Effects:** Motion graphics and compositing
- **DaVinci Resolve (ACES):** Color grading workflow
- **Topaz Video AI:** Upscaling and enhancement
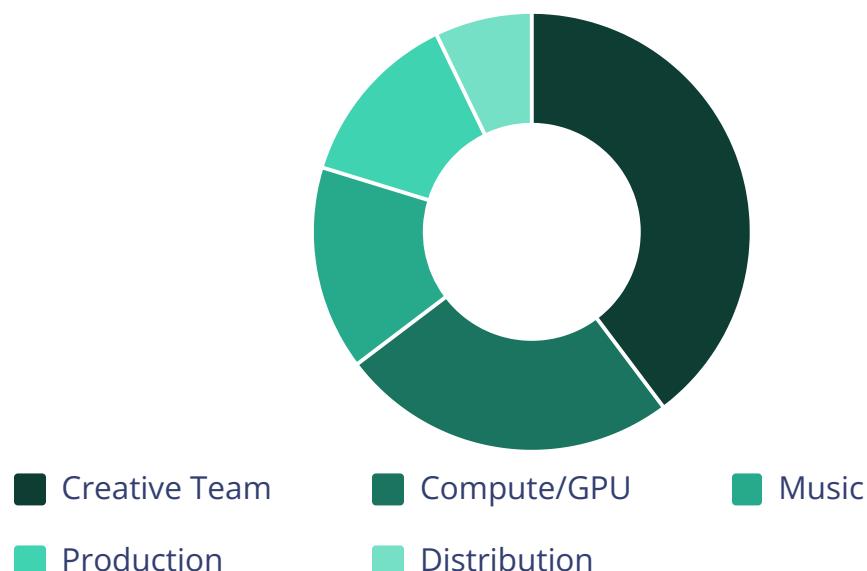- **Frame.io:** Review and collaboration

Operations

- Comprehensive prompt library with seed/version control
- Full provenance logging for transparency
- Green-compute tracking for responsible AI usage

Timeline: 6 Weeks to Launch

Week 1
- Treatment development, emotion taxonomy design, style bible creation, pilot scene testing
- Stand up Website MVP skeleton (auth, upload, job create/status) that mirrors the pipeline contracts.

Week 2
- Prototype v1 build (prompt→curve→temp score), Leonardo look development and asset generation
- Wire website to prototype: submit job → queue → return placeholder artifacts.

Week 3
- Produce Clip A; first A/B comparisons; review loop.
- Website: add provenance view (seeds, tool versions) and GPU queue/fallback status

Week 4
- Clips B/C production, music swap testing, stakeholder review cycle and refinement
- Behind the scenes LinkedIn post & newsletter push for in-process build

Week 5
- Final edit, color grading, audio mix/master, process film shoot and assembly

Week 6
- Export deliverables, caption creation, atlas PDF finalization, social teasers, micro-premiere
- Newsletter & LinkedIn post release with link to Leonardo & in depth explanation of mechanics
- Expose website to public for usage

Budget: $10,000 USD

- Creative Team
- Compute/GPU
- Music
- Production
- Distribution

Risk Mitigation
- **Licensing:** Default to original/generated audio or cleared library content
- **Consistency:** Rigorous seed control, style bible adherence, IP-Adapter workflows
- **Compute:** Budgeted GPU allocation with render queue management and lower-res fallback passes
- **Scope:** Three tightly defined clips; locked picture before polish phase begins

Call to Action

Fund SyncSense to prove that feeling is a feature. In six weeks, we'll deliver a working prototype, three compelling demonstration shorts, comprehensive process documentation, and a public atlas that empowers the entire creator community. This isn't just another AI demo — it's the bridge between impressive technology and genuine emotional storytelling.

Together, let's transform AI video from "cool" to *compelling*.