



TALK-TO-TUNE

A synthesizer you can talk to

Tyler Fossen · Agentic AI Strategic Roadmap · Capstone Project · August 2026

Tune with your voice, not the dials.



THE USE CASE

A conversational agent living inside a software synthesizer — a virtual instrument in the DAW, the app where music is made. The artist says “warmer,” “grittier,” “make it pulse slowly” — and the agent turns the dials for them — live, mid-session, no tinkering.



THE PROBLEM

Sound design gatekeeps music-making. Artists know the sound they want, but producing it takes years of skill across dozens of interacting controls. Flow breaks. Ideas die in menus.

WHY IT MUST BE AGENTIC

Musical intent is ambiguous and context-dependent. “Warmer” can’t be enumerated into presets or if-then rules — ambiguity is exactly where autonomy earns its place.

AUTONOMY & THE CORE TRADE-OFF

Autonomy vs. control, managed by design: full freedom to explore — every audition is instantly reversible — and zero freedom to commit without human-in-the-loop (HITL) approval.

The musician always owns the cut.

KEY TERMS, IN PLAIN ENGLISH

Agentic — software that pursues a goal by deciding its own steps, not following a script.

Autonomy — how much the agent may do on its own before a person steps in.

HITL (human-in-the-loop) — a person must approve before anything sticks.

DAW — digital audio workstation; the app where music is recorded and produced.

Synth / LFO — a synthesizer generates sound; an LFO (low-frequency oscillator) adds automatic motion — wobble, pulse, sweep — that normally must be programmed by hand.



From preset menus to a musical conversation

TODAY'S MARKET

Automation

Preset browsers, macro knobs — and our prototype today: clickable words mapped to fixed sound changes. Pre-built recall, zero interpretation.



WHERE WE LAUNCH

Augmentation

The agent works beside the artist — interpreting intent, auditioning variations, absorbing the menu-diving.



WHERE THIS GOES

Transformation

Sound design becomes conversation. The artist directs; anyone can shape professional sound by talking.

HYBRID BY DESIGN

A deterministic backbone — the sound engine, locked safe ranges, a hard volume limiter — beneath one bounded autonomous layer.

DATA READY ON DAY ONE

Every input already exists: synth state, the artist's words, and an accept / reject history that compounds into personalization.



One artist in charge, one agent on task



Decision rights

The agent proposes; the artist approves. Commit authority never transfers — one owner of the cut, always human.



People systems

Augmentation that teaches: the agent shows which parameters moved and why, so newcomers learn the craft instead of losing it.



Single-agent architecture

Multi-agent studies put ~37% of failures at agent handoffs, with diminishing returns past ~4 agents. One bounded agent — no orchestra.



Value chain

Built inside VST3 / AU — the universal plugin formats every major DAW accepts. One integration surface; swappable model APIs upstream; plugin marketplaces downstream.



Measured like a product, not a demo

Value quadrant: **Differentiation** — an improved product captured as external revenue, opening a new segment of non-technical musicians.

LEADING KPIs · WEEKLY

- Suggestion acceptance rate
- Iterations to an accepted sound
- Latency — it must feel live
- Human override rate
- API cost per session

LAGGING KPIs · QUARTERLY

- Retention and churn
- Free → paid conversion
- Sounds kept per session
- Satisfaction vs. baseline
- A/B: actual AND perceived speed

≈ 10¢

target API cost per session

\$12 / mo

target pro tier

85%+

target gross margin

GUARDRAIL & KILL THRESHOLD Creative satisfaction must never drop — the Klarna lesson. Kill or rework if acceptance falls below floor or latency above ceiling. Audio-safety violations: zero tolerance.



Safe by construction, not by promise



Bounded action space

The agent can touch synth controls only — no files, no internet, no inbox. There is simply no door for data to leak through.



Deterministic chokepoints

A hard volume limiter and locked safe ranges live outside the model in plain code, protecting ears and speakers. No prompt can talk its way past code.



Guardrails in code, not prompts

Approval is enforced by the commit button's permission model. Every action is logged and one-click reversible — find and fix.



Data, compliance & worst case

Taste history stays per-user; nothing trains without opt-in. Shared presets are treated as data, never instructions. Worst-case failure: a bad-sounding patch, instantly undone.



The synth is built. The agent is next.



STARTED DURING THIS PROGRAM The synthesizer is built and playable — clickable words already change the sound through fixed mappings. The remaining work: replace that deterministic layer with the agent.

PHASE 1 · MONTHS 0–2

Add the agent

Swap fixed word-to-sound mappings for real intent interpretation across the full parameter space. Safety layer first; shadow-mode logging from day one.

PHASE 2 · MONTHS 3–4

Prove the value

Closed beta with 25–50 producers. A/B test speed and satisfaction — actual and perceived. Tune guardrails and kill thresholds.

PHASE 3 · MONTHS 5–6

Ship the product

VST3 / AU release with a freemium tier. Voice input. Personalization from each artist's approval history.

RESOURCES

One builder with AI-assisted development · model API budget under \$300 / month through beta · a producer-community beta cohort.

THE ASK The foundation is built. Greenlight the agent layer — demo in 60 days.

Sources

FRAMEWORKS FROM THE PROGRAM

- Automation → Augmentation → Transformation curve; autonomy & human-in-the-loop (HITL) spectrum — Module 2, Agentic AI Design and Evolution (Prof. Thomas Lee)
- Decision rights, people systems, and value-chain alignment — Module 3, Organizational & Agent Architecture (Prof. Thomas Lee)
- Value generation × capture matrix; leading / lagging KPIs; guardrail metrics and kill thresholds — Module 4, Measuring the Value of Agentic AI (Prof. Zsolt Katona)
- Bounded action space, deterministic chokepoints, and guardrails-in-code — Module 5, AI Governance & Security (Prof. Zsolt Katona)

CASES & STUDIES REFERENCED

- Klarna AI customer-service rollout and partial reversal — public reporting, 2024–2025; discussed in Module 4 (“the Klarna lesson”)
- MAST: Multi-Agent System Failure Taxonomy — UC Berkeley analysis of why multi-agent LLM systems fail (~37% of failures at inter-agent handoffs)
- Multi-agent scaling research on diminishing returns beyond ~4 agents — discussed in Module 3
- METR 2025 randomized trial of AI-assisted developers (actual vs. perceived speed) — referenced in presentation notes

