

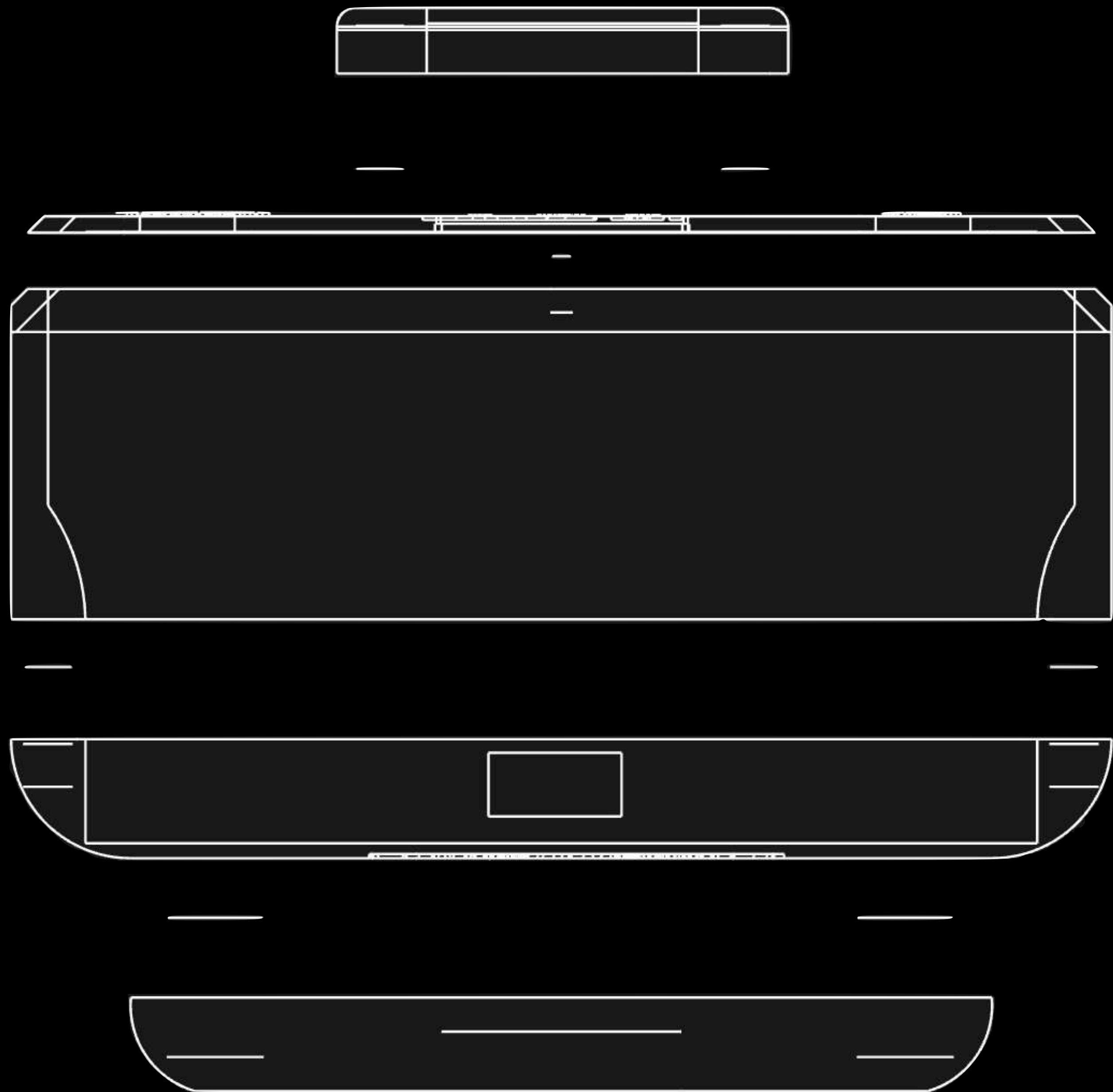
VJ^(AI)

Process Book

Saif Bajnaid

IND-402-02

Embodied Intelligence 2026



- Design Brief
- Research
- Sketches
- Software Process
- Hardware Process
- Finalization

Design Brief

Executive Summary

VJ(AI) is an embodied artificial intelligence performance enhancer for DJs that generates real-time visuals that react to the music being played. It allows DJs to elevate the energy of their set without hiring a VJ or learning complex visual software.

Problem definition

DJs want to create immersive experiences, but immersive visuals require a steep learning curve.

- VJ software requires hours of preparation, programming, and technical knowledge.
- Hiring a VJ is expensive and unrealistic for independent or emerging DJs.
- Running DJ software and VJ software simultaneously is cognitively and technically overwhelming.

Functional Requirements

- **Device must detect audio energy and adjust visuals in real time.**
- **Hardware must operate with minimal components**
- **Visual output must run as a secondary screen / projection instantly.**
- **Able to provide visuals that matches the song being played in the visual style the DJ sets and can choose to create variations from.**

Project Goals

- **Generate visuals that emotionally and rhythmically follow the music.**
- **Build a system that evolves with the DJ's style over time.**
- **Deliver a tool that works on any setup within seconds.**
- **Make complex visual performance feel effortless and accessible**

Success Criteria

- **Setup under 60 seconds** from opening laptop to projecting visuals
- **Positive emotional response:** DJs feel more expressive and crowds feel more immersed.
- **Consistency:** visuals align with the DJ's style across multiple uses.

Research

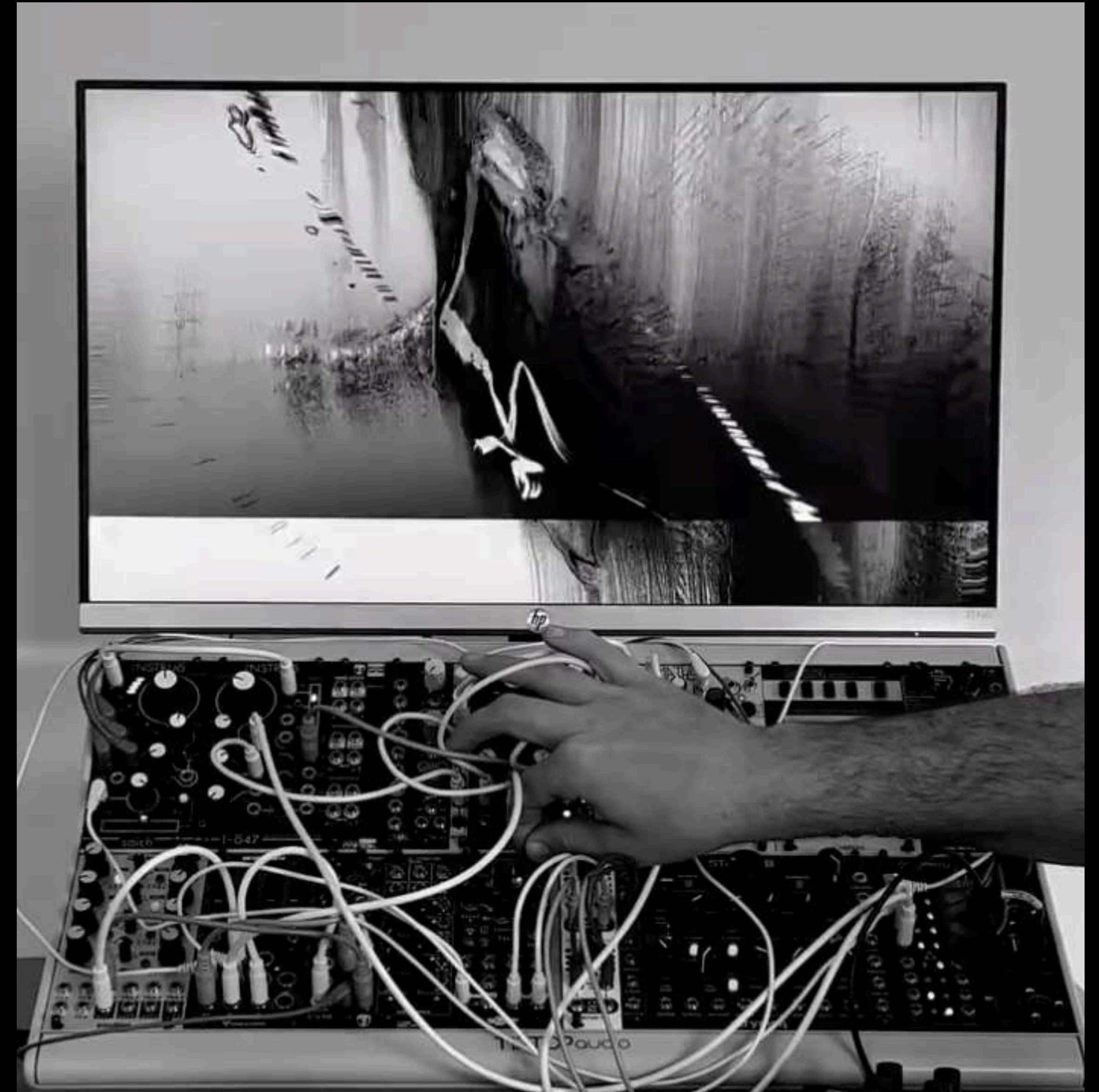
What's the niche?

Independent and performance focused DJs who want to elevate their sets with visuals.



What's the problem?

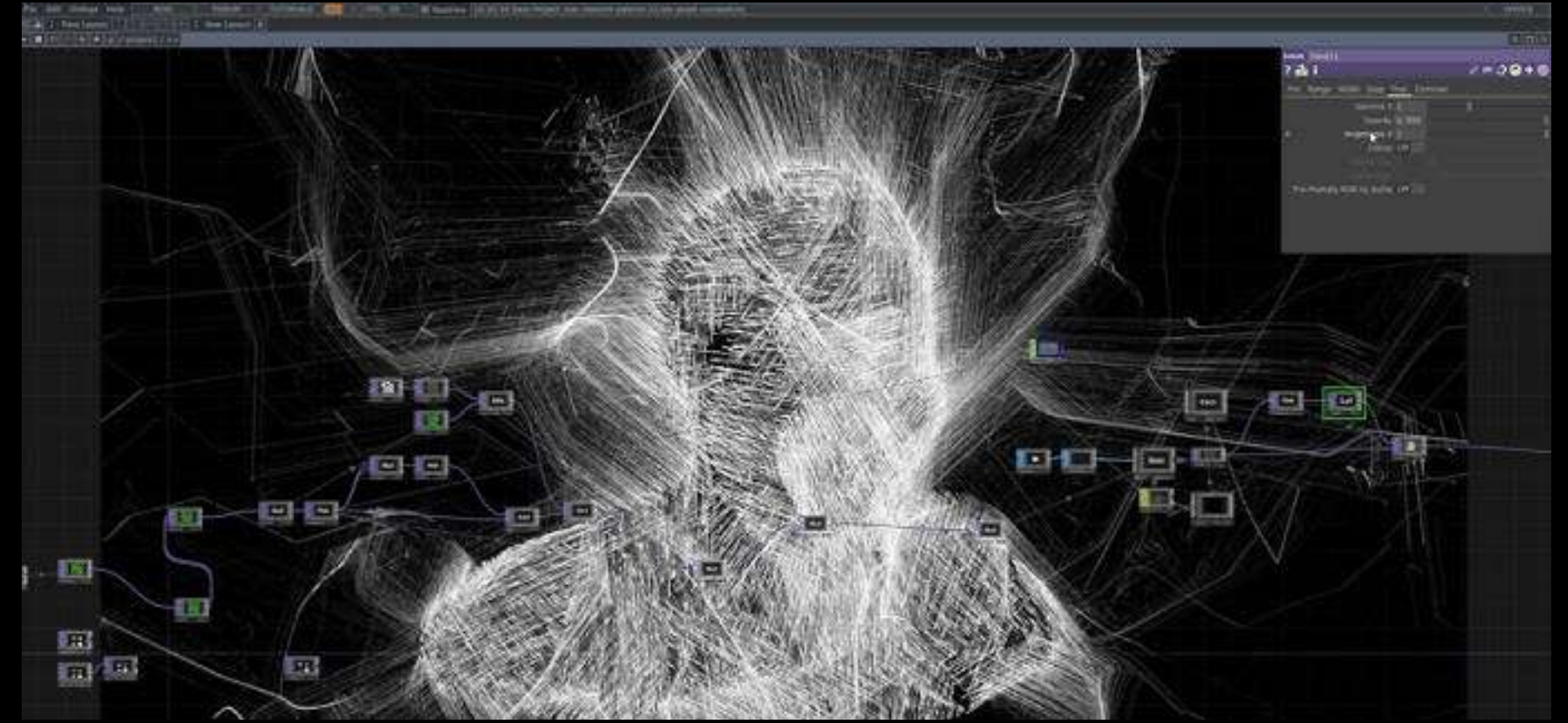
VJ softwares have a steep learning curve and requires hours of preparation and programming



What's the need?

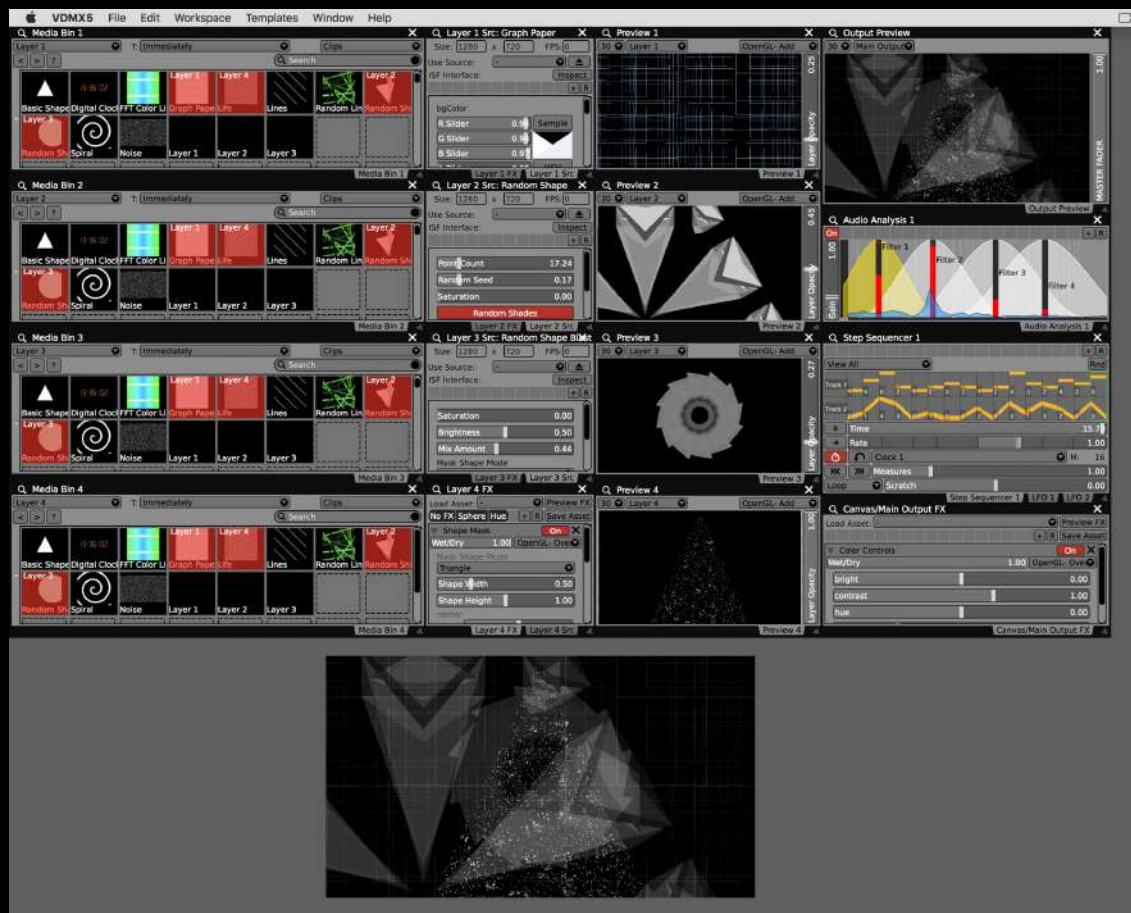
A way to generate visuals for DJ sets to make it more immersive without having to hire a VJ or to constantly monitor a VJ software.





What exists?

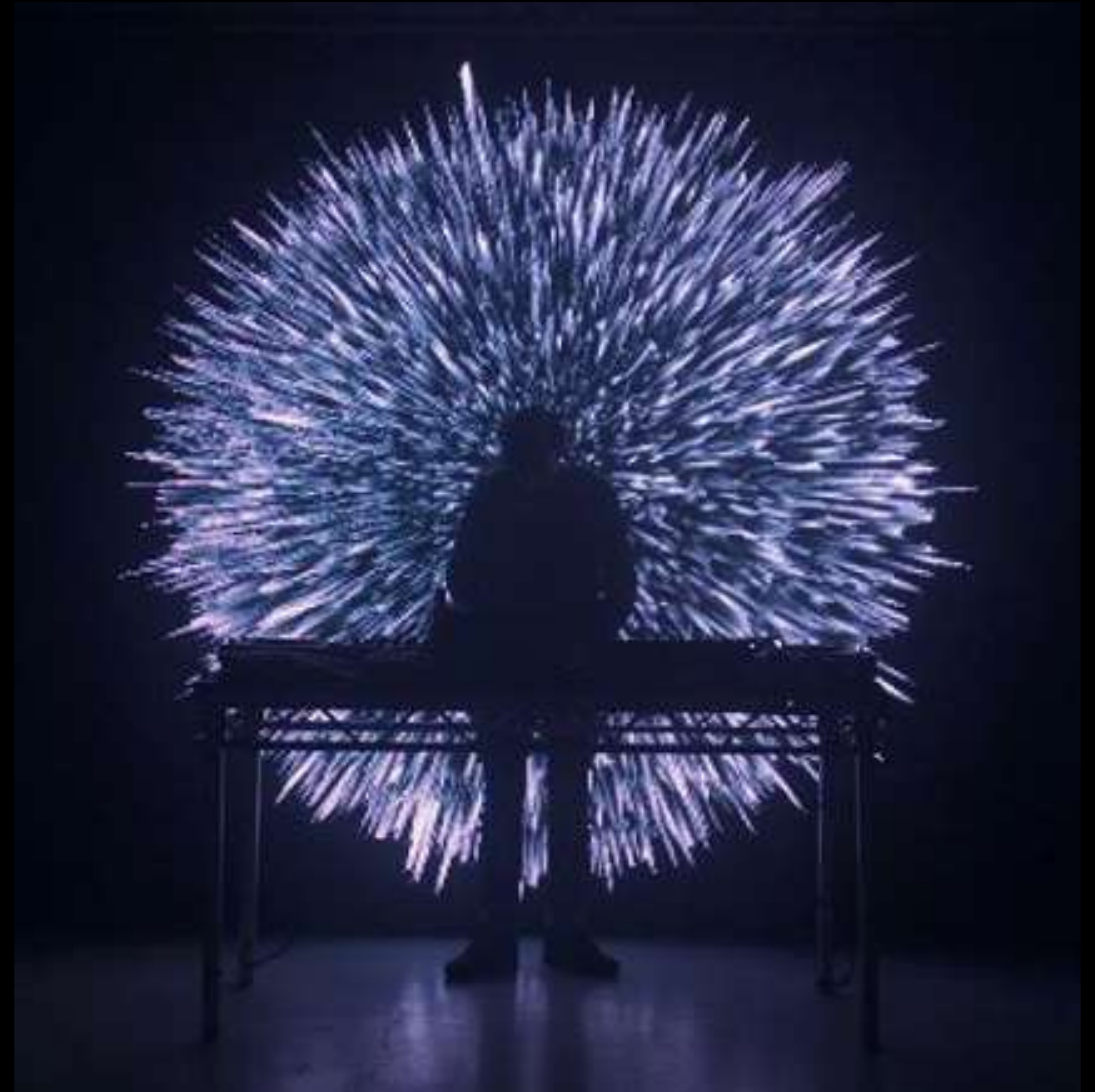
Visual Synthesizers, TouchDesigner, Resolume, VDMX, HeavyM, VJ/MIDI Controllers



Why?

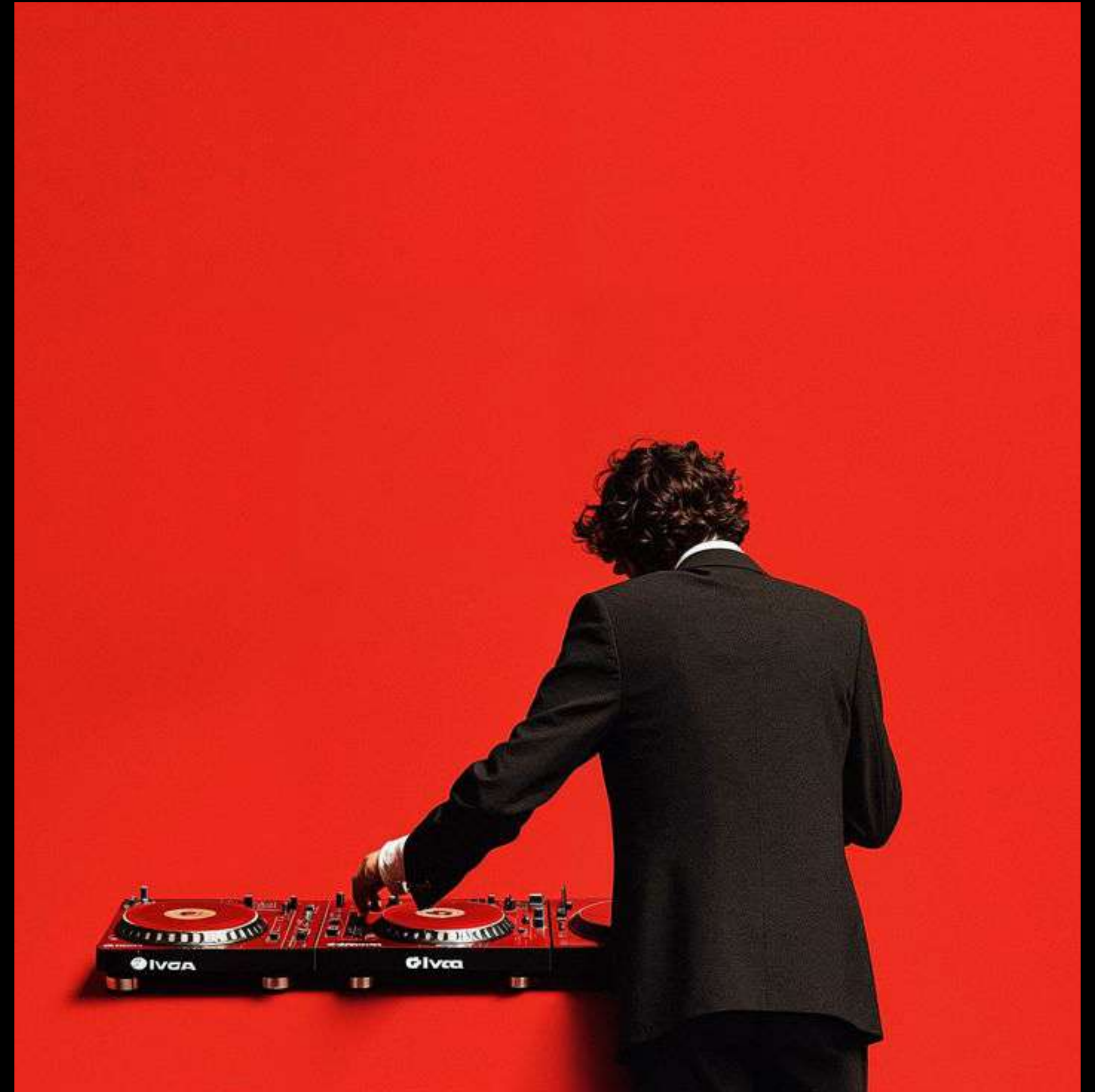
Because it allows the independent DJ to generate professional grade experiences.

It takes the "Brain" of generative AI and the "Body" of a visual synthesizer and merges them into a single and affordable tool to enhance a DJ set.



Questions asked:

- Do you normally wish you had a VJ at your set? If so, what stops you from having one?
- Does the concept of AI generated visuals for your set sound appealing?
- What if there was a tool where it could generate visuals that reacted to your music without the need to learn a VJ software?
- What would make you use a tool like that?



Research

User Persona 1: Tyler, 29, NYC



- **Wishes they had a VJ at their sets but cant afford to hire one.**
- The concept of AI generated visuals that react to the tracks seems intriguing and would love to test it out.
- They would use the proposed tool if it did not require a software learning curve and would not mess up their set.
- Would primarily use the tool at small sets first instead of gigs

User Persona 2: Fele, 24, Saudi Arabia

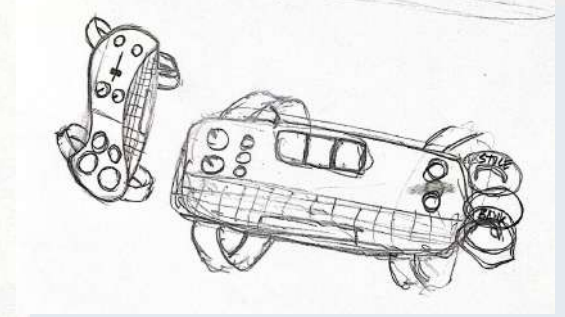
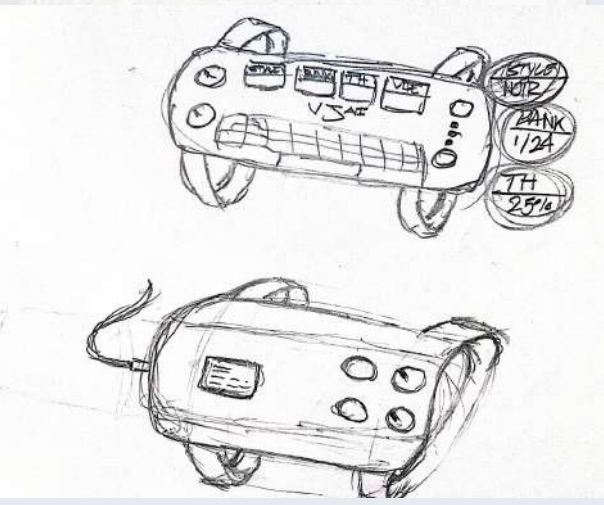
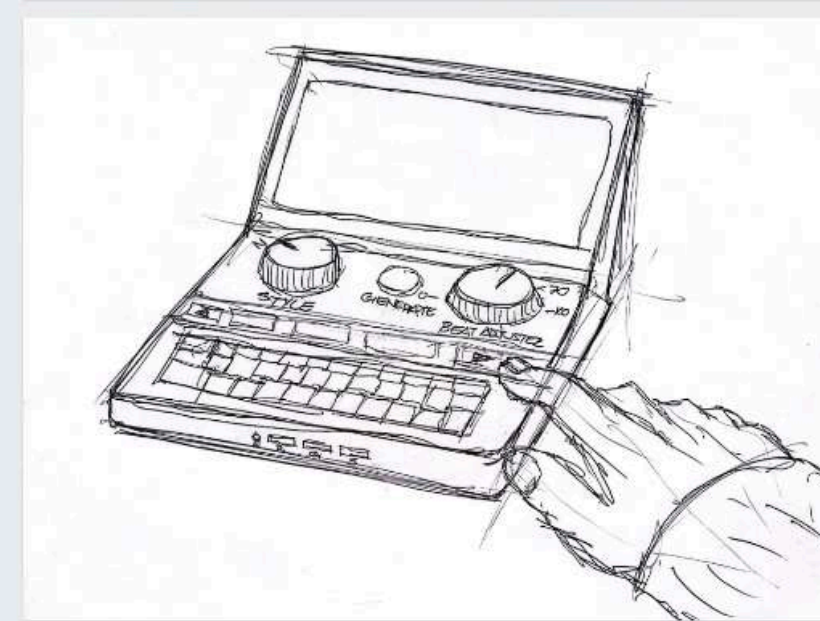
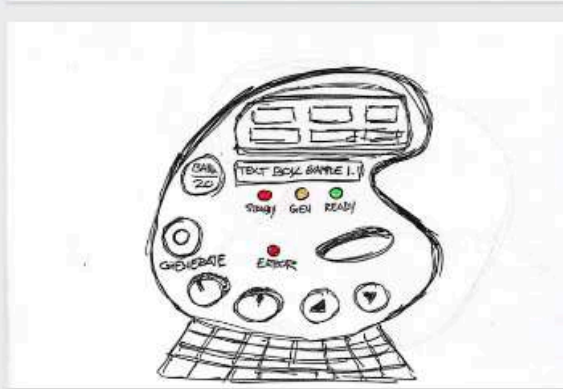
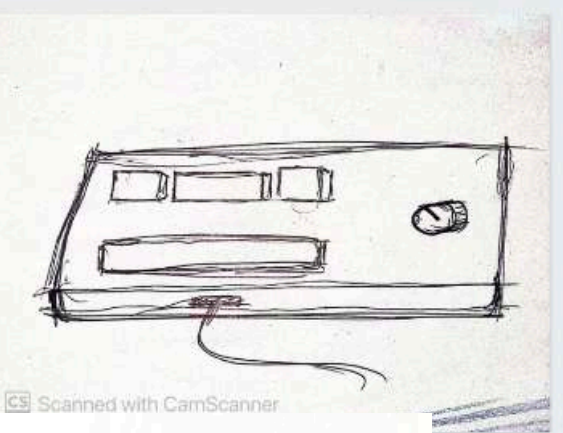
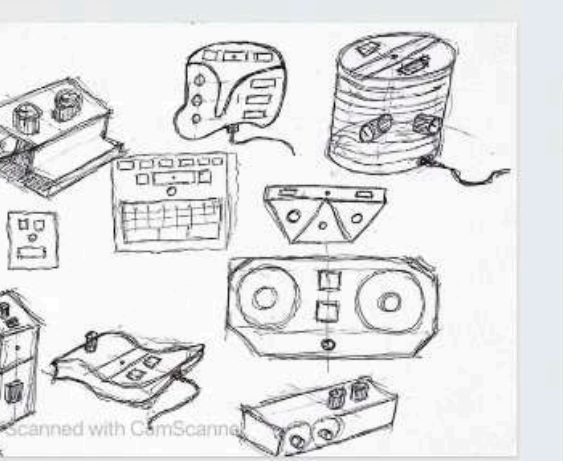
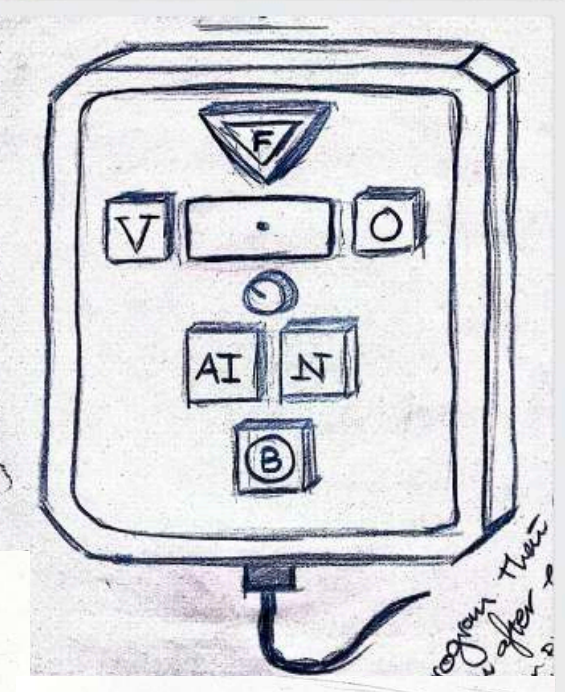
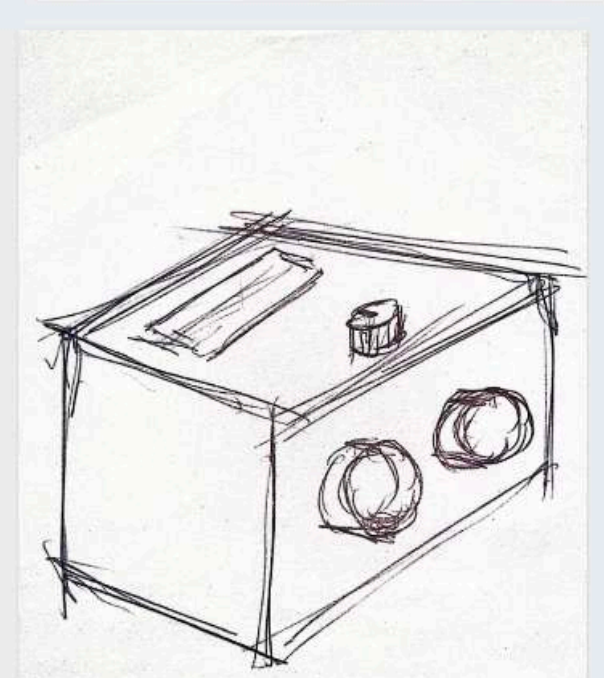
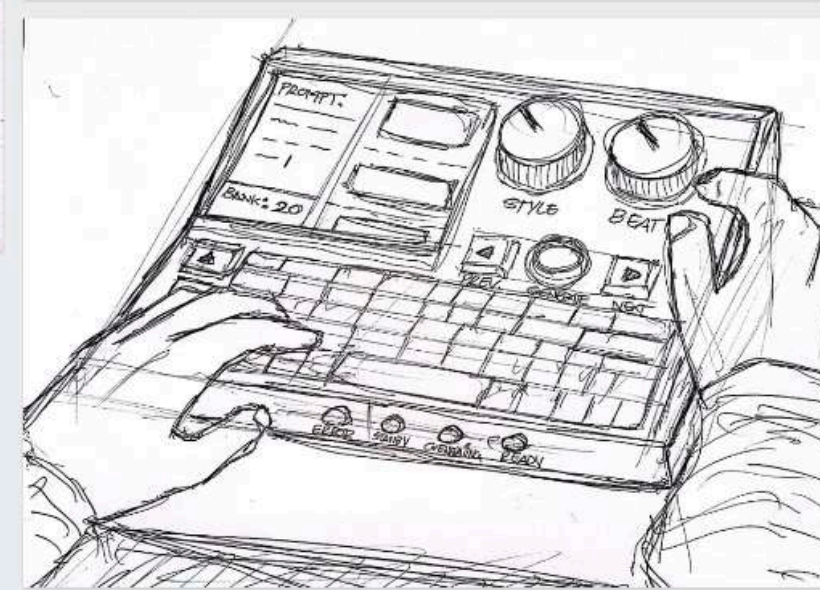
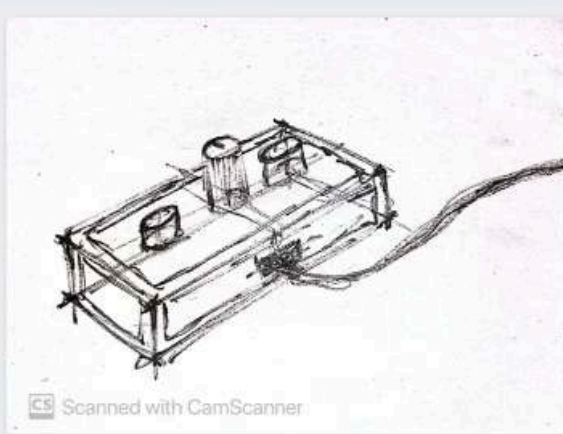
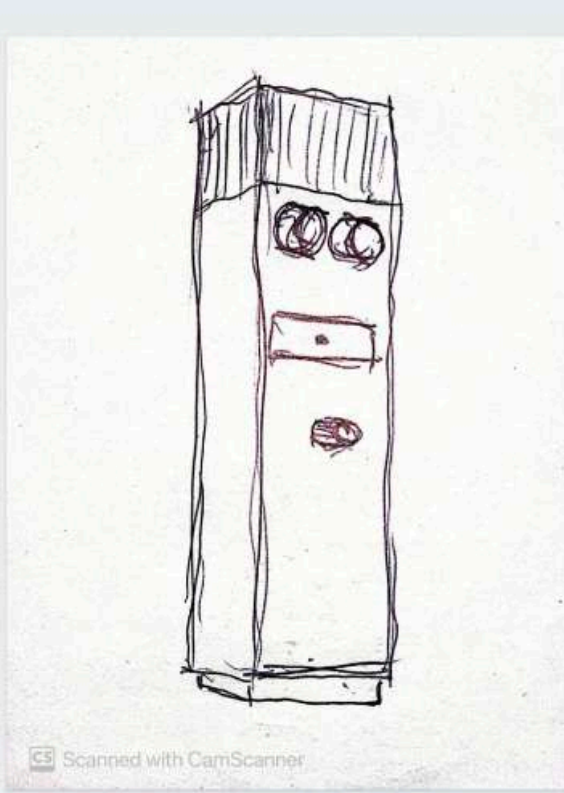
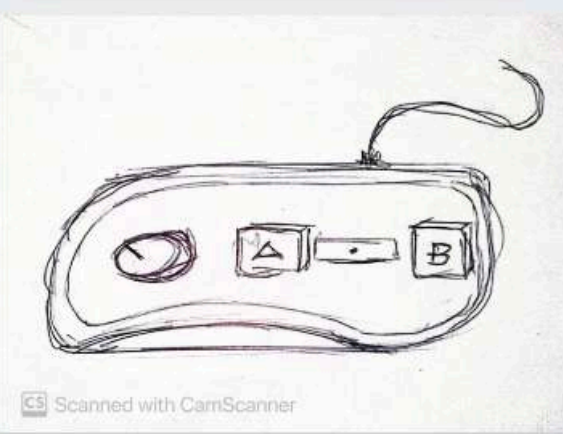
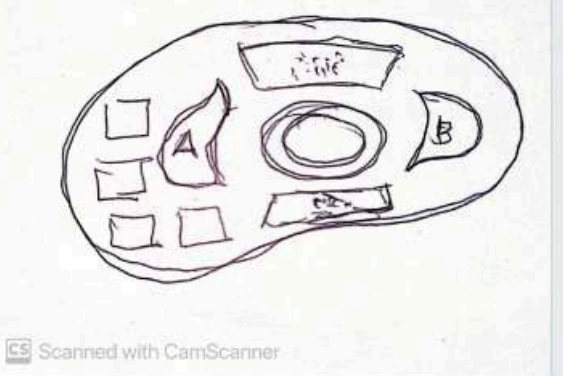
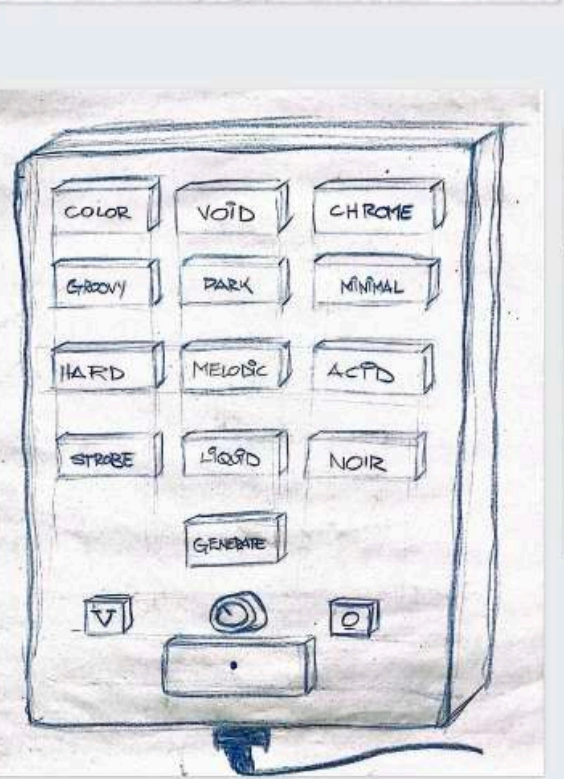
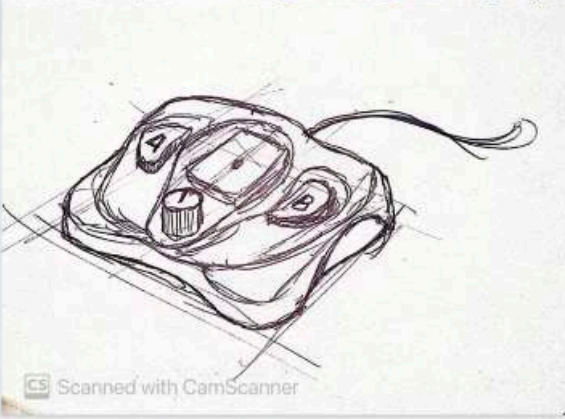
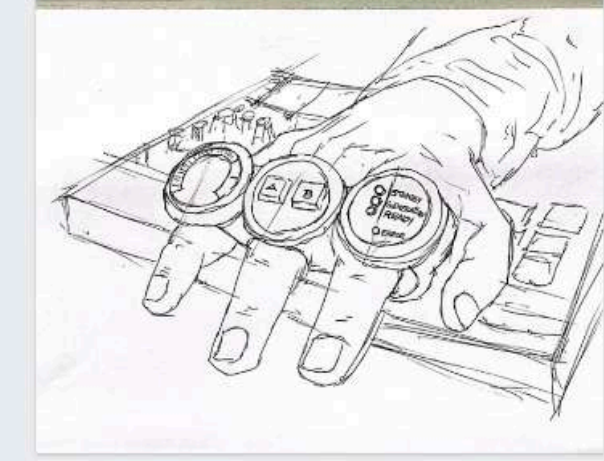
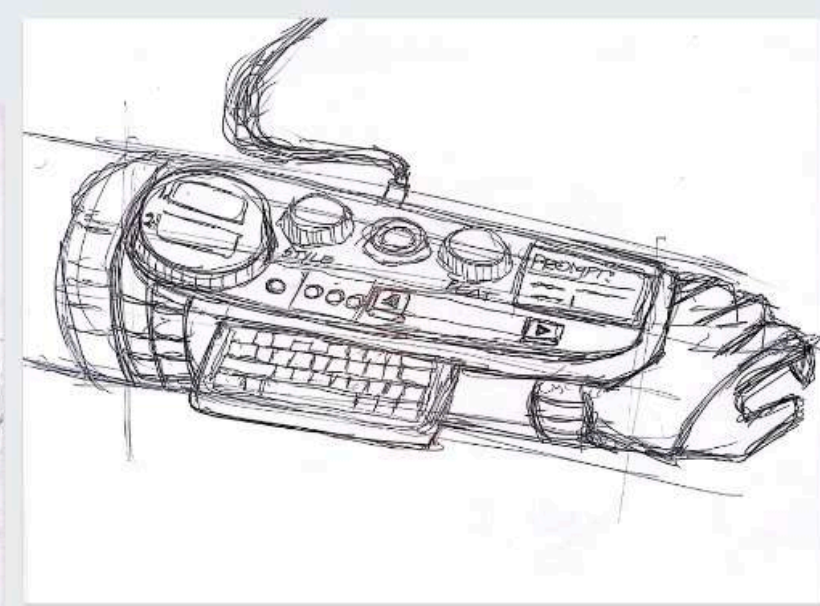
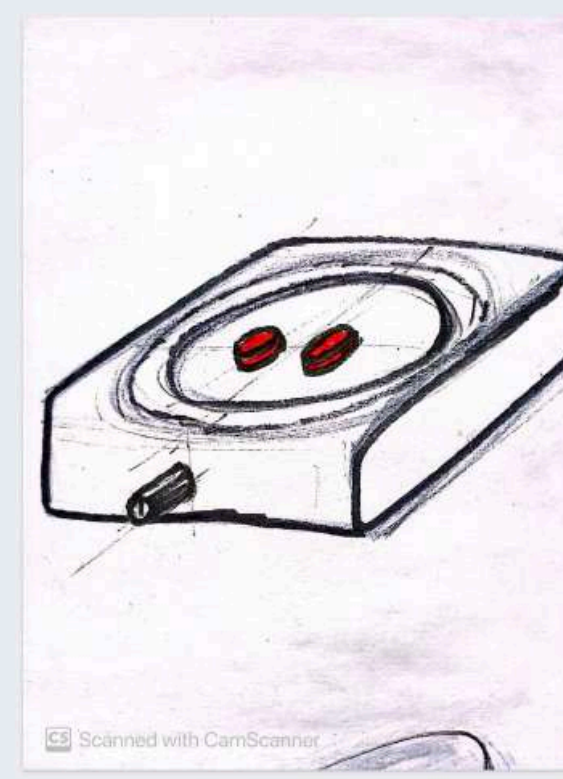
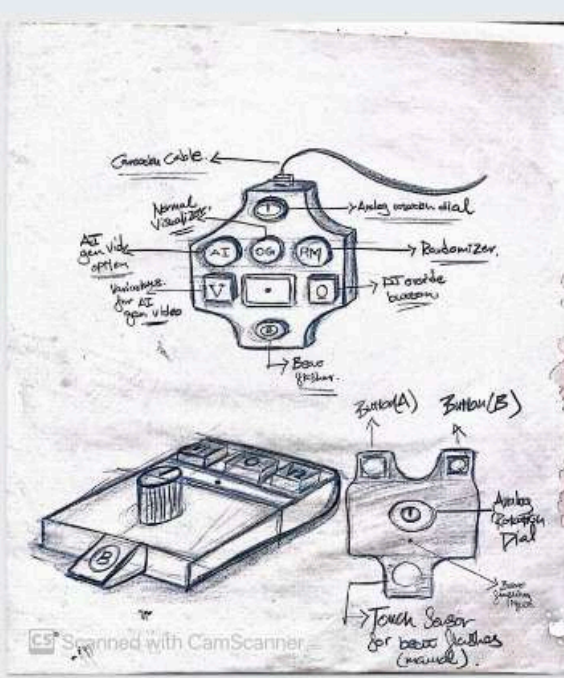
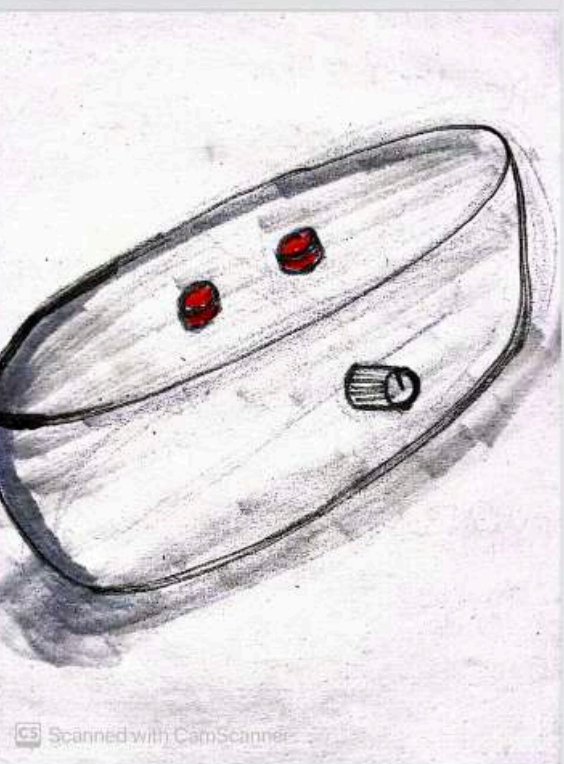
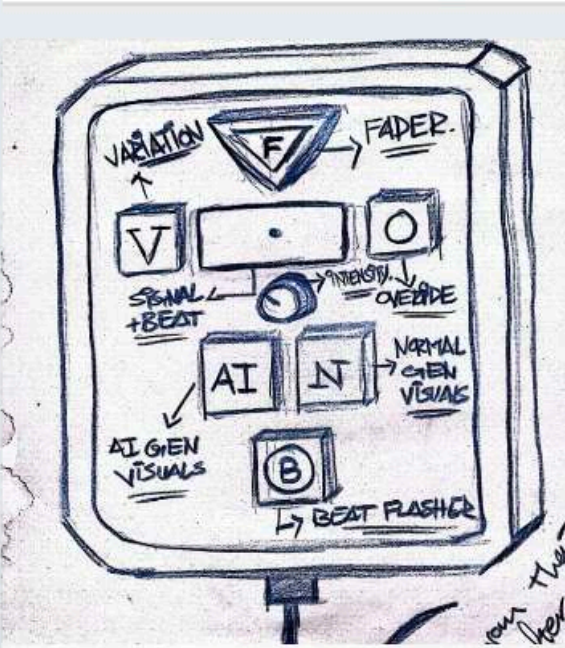
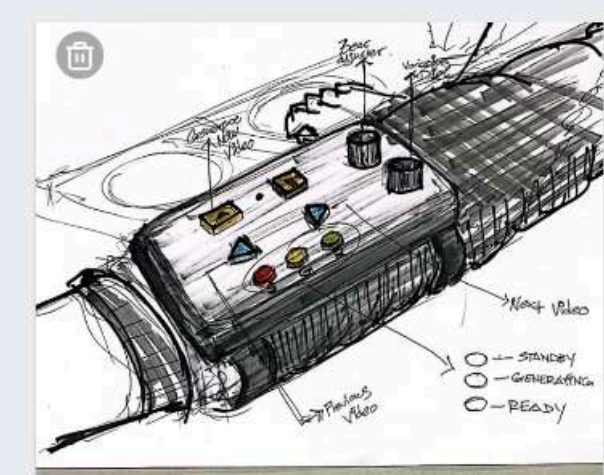
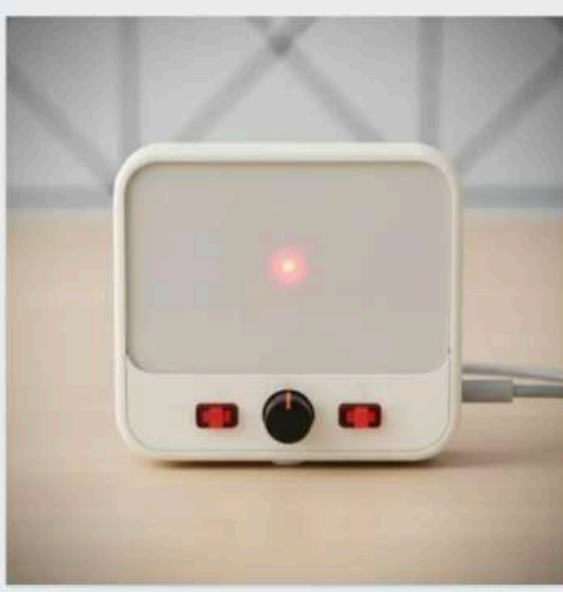
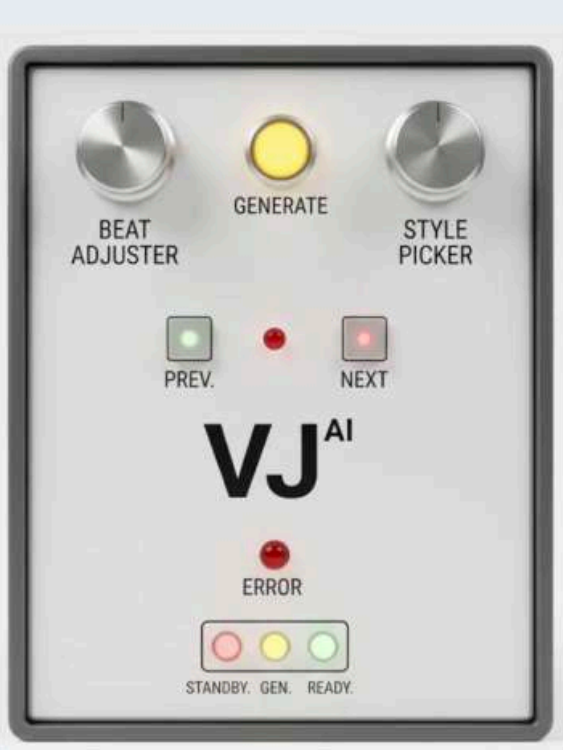
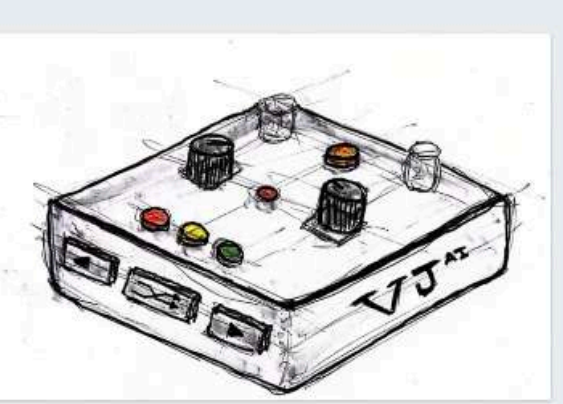
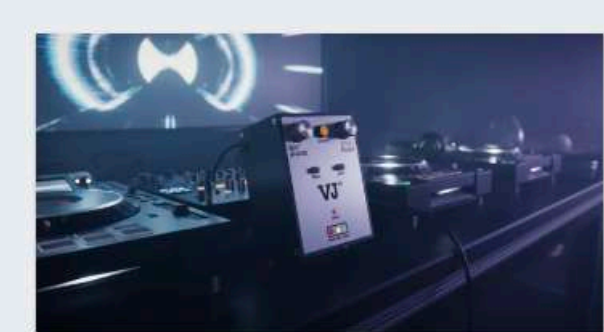


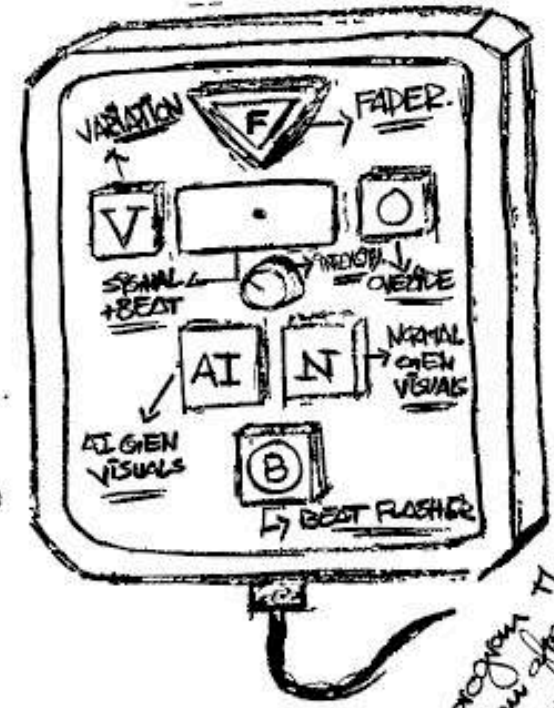
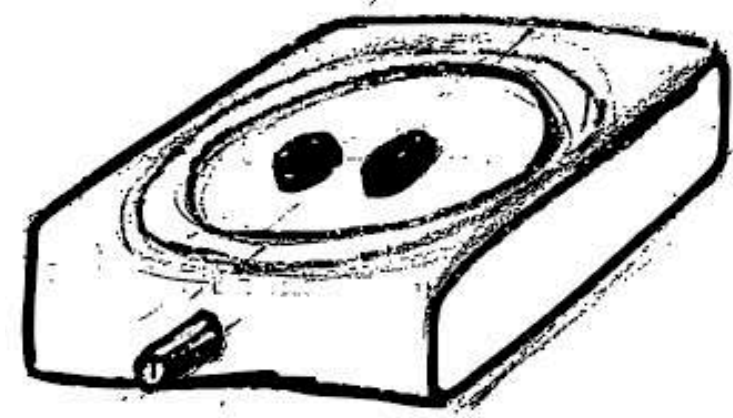
- **Wants visuals within their set without having to use a VJ software.**
- Seemed skeptical about Real-time AI generated videos that react to music due to latency issues
- Would use the tool if made and would ideally love to have a projection system within the tool itself instead of having to hook it up to an existing projector.
- Would use the tool if it was a Plug and play type of solution.

VJ^(AI)

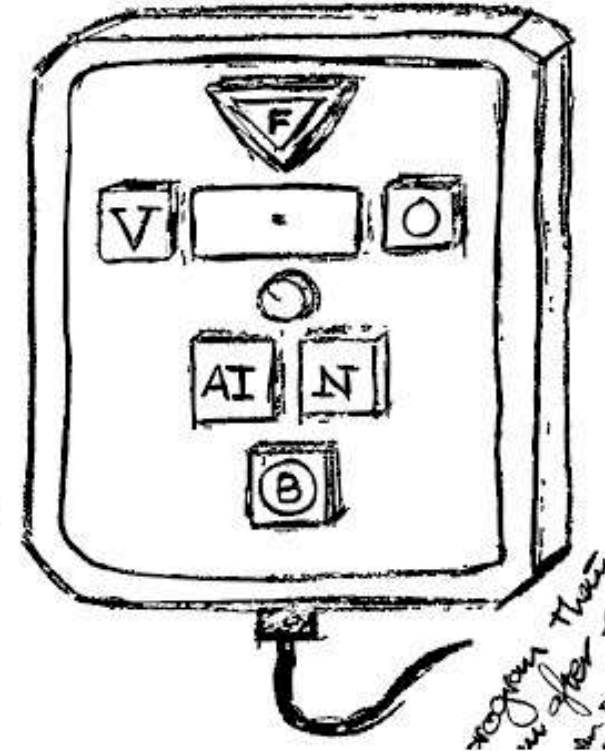
Other Findings: DJ's are also interested in it being wearable and having it also adapt to the booth.

Sketches

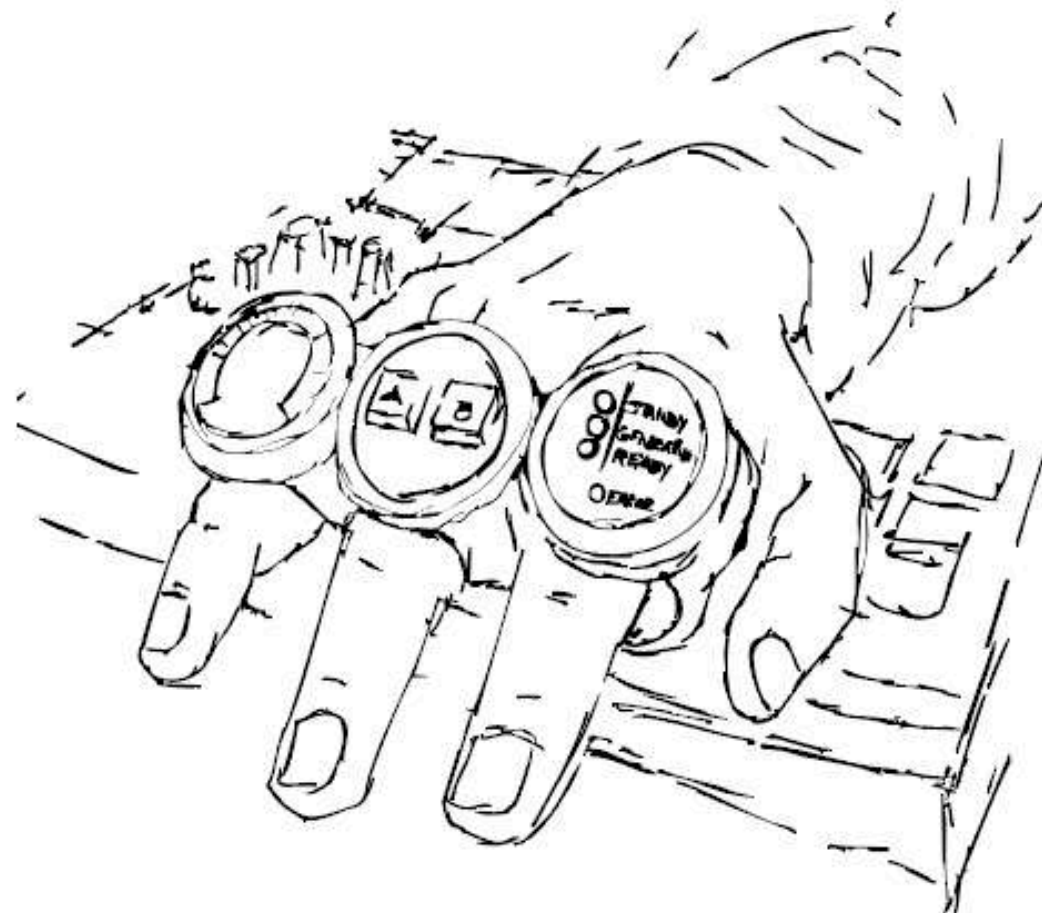
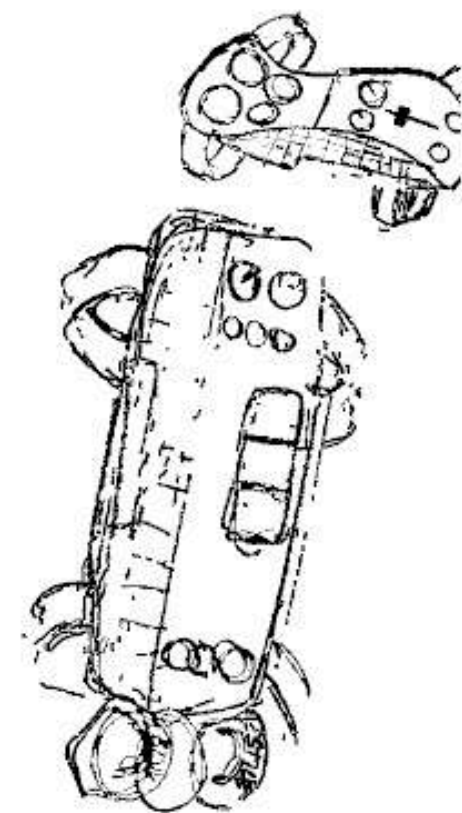
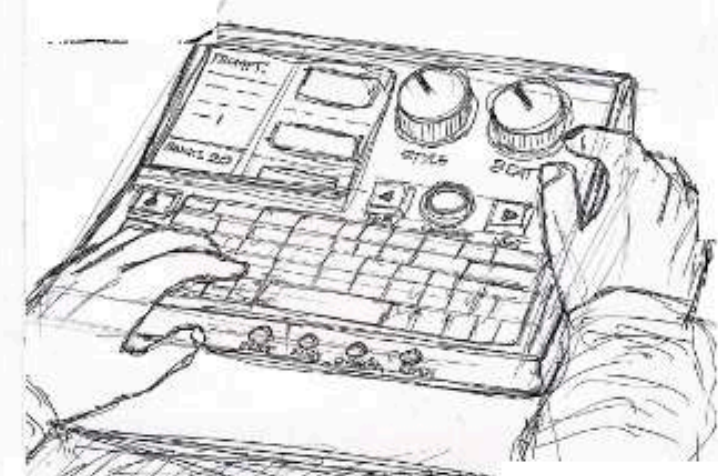
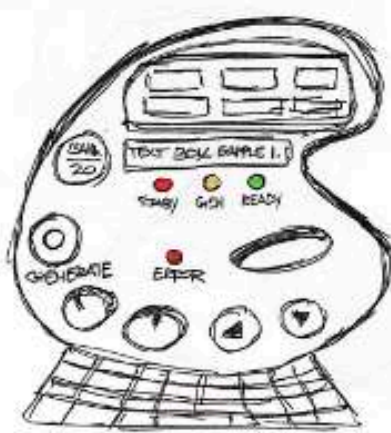
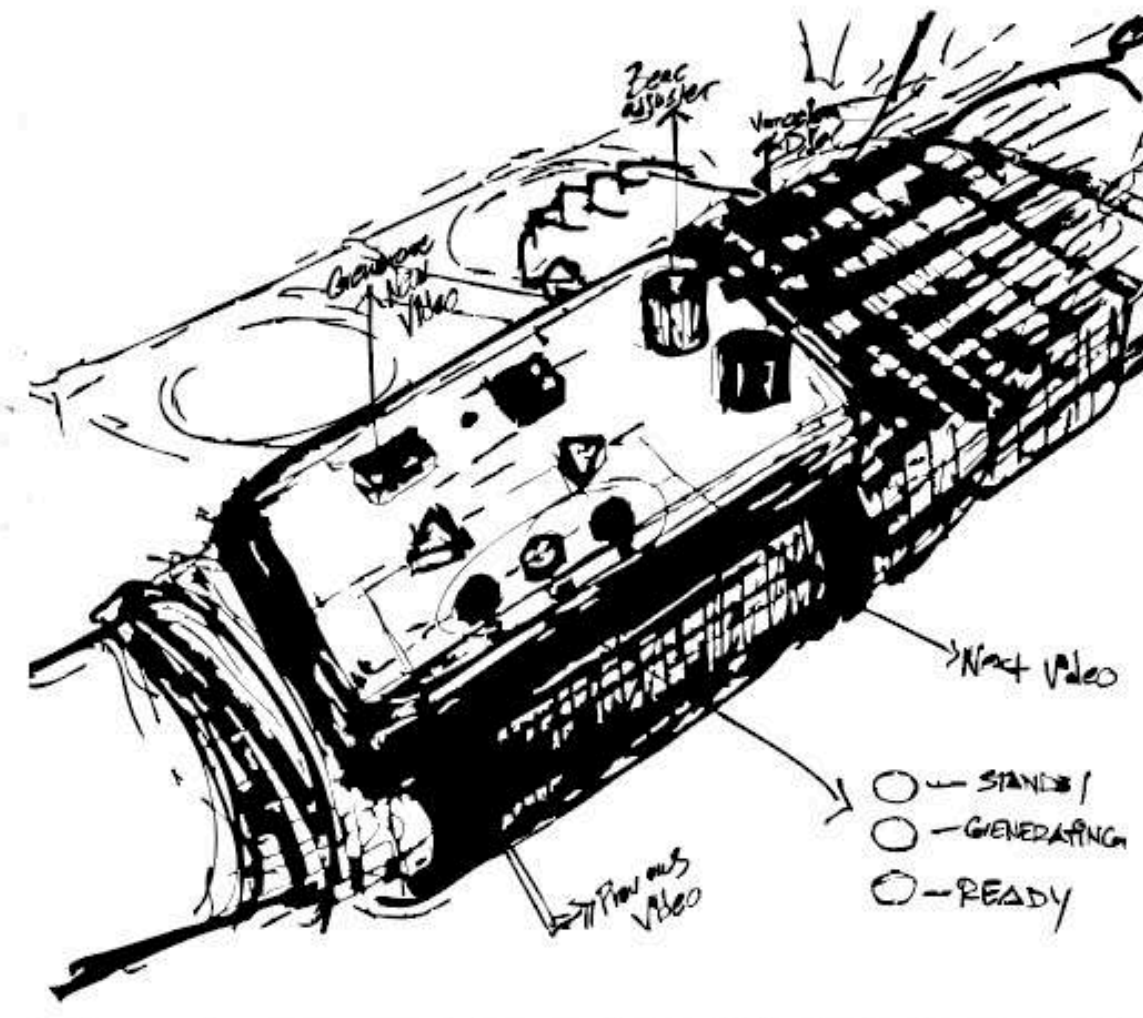
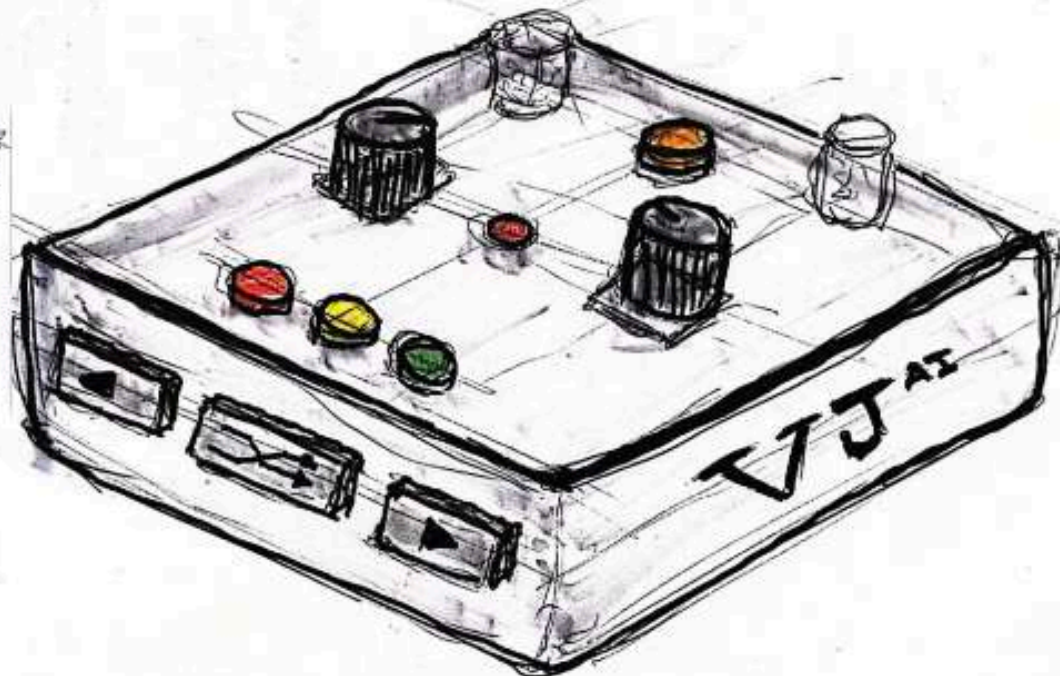
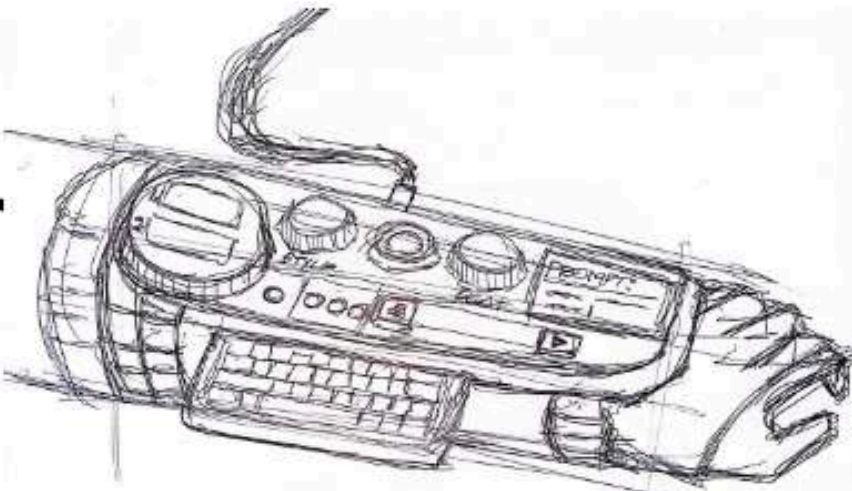
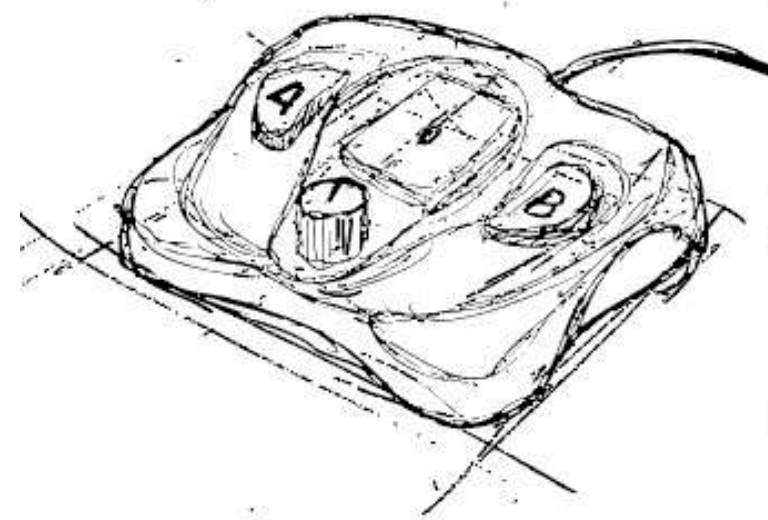
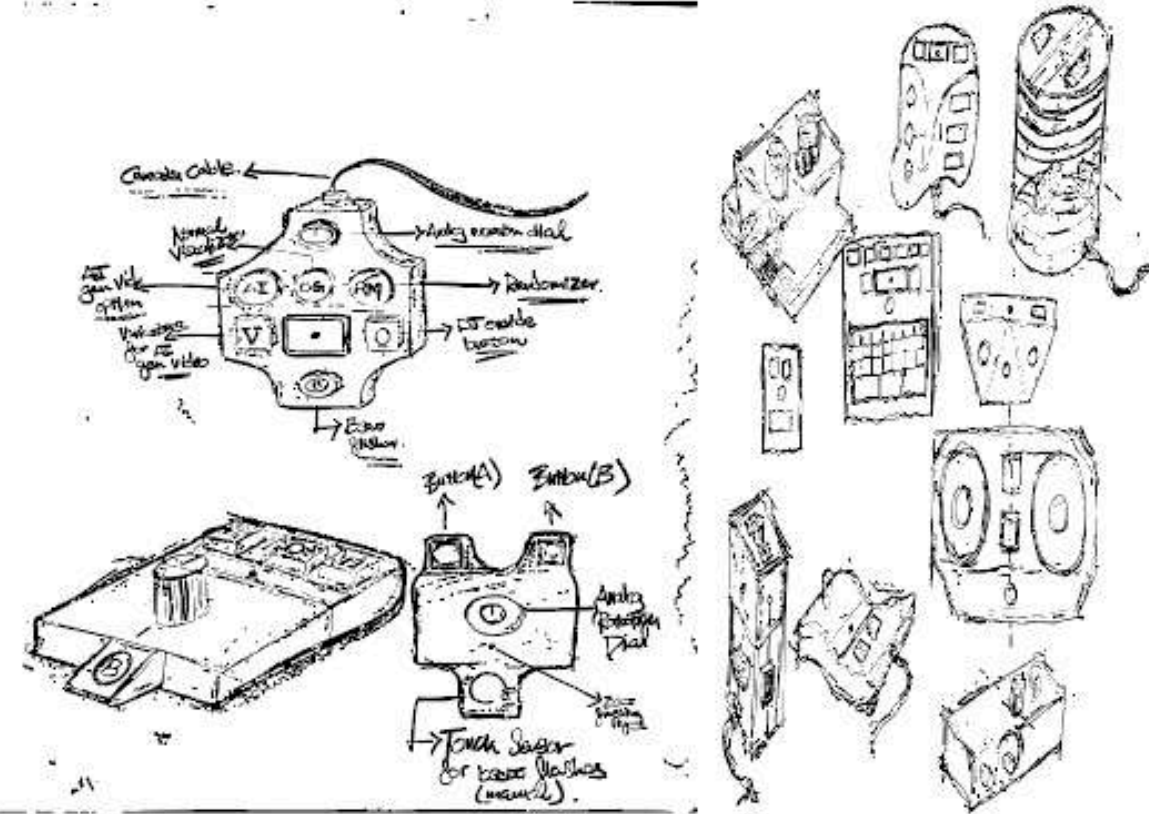
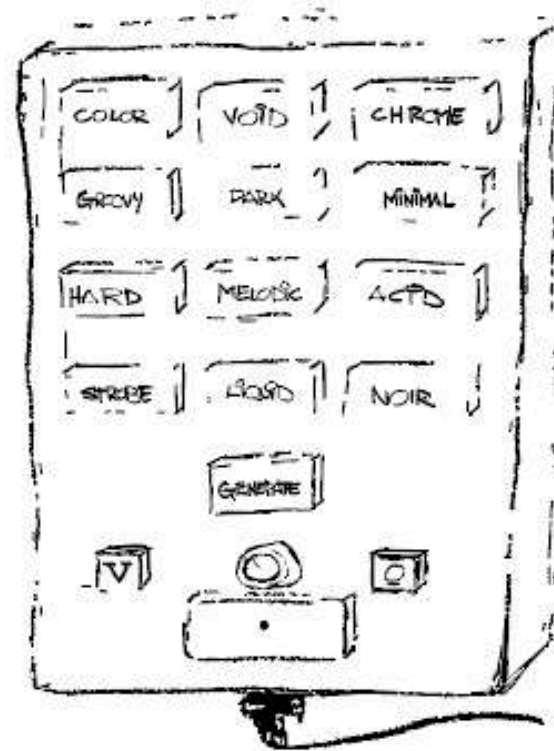




Program These
Run after



Program These
Run after



Software Process

The Goal

To develop a software that talks to the Microbit hardware and shows the visuals that are being transmitted.

What was explored:

Google Veo
Runway ML
Sora AI
Kling AI
Artlist API
FAL API (Wan 2.2)

The Decision

To use FAL's API to generate videos using Wan 2.2

This was due to Google Veo being too slow and costing too much per generation (\$7), Runway burned through credits fast as well, Sora and Kling did not deliver the style in mind, Artist subscription did not win me over.

Inspiration of video styles

Actual style developed



33



34



35



YZAVoku™ Podómetro. [Archive](#) [Film](#) [Image](#) [Video-Loop](#) [Info & Contact](#)



39



40



41



Prompts: (4 styles)

NOIR: (Typed prompt here) Extreme high contrast black and white, 35mm film pushed to ISO 3200, deep chiaroscuro with harsh backlight, heavy grain and film texture, slow creeping handheld camera, paranoid oppressive mood, explosive kinetic energy, camera always in motion, fast movement, staccato visual rhythm, never a static frame, seamlessly looping

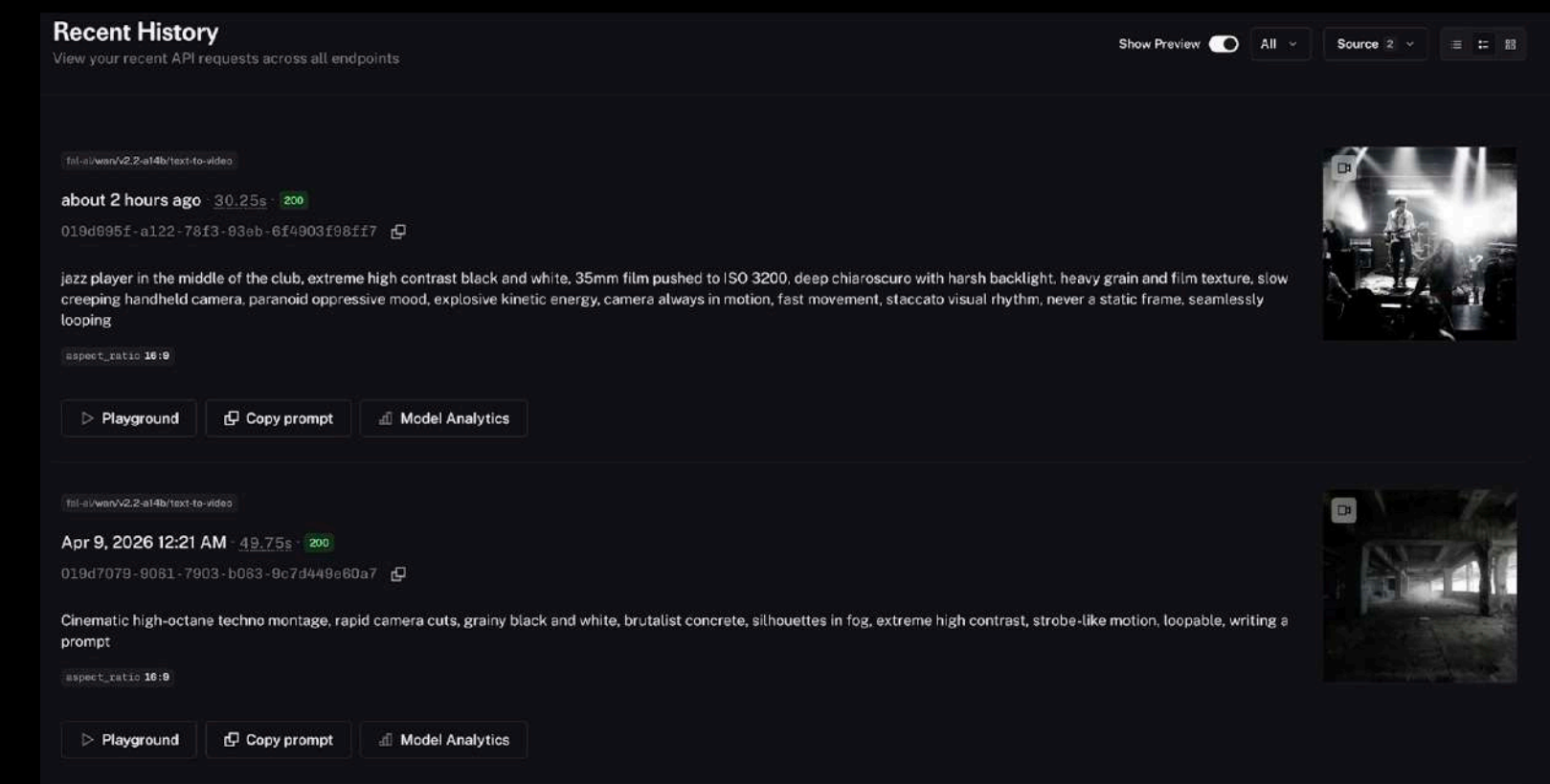
RED: (Typed prompt here) Lit by a single deep crimson practical light source against near-total darkness, blood red illumination with pitch black shadow falloff, extreme chiaroscuro, 35mm grain, slow hypnotic camera drift

BLUE: (Typed prompt here) Shot in deep cobalt blue sodium light, cold desaturated 3am color temperature, raw documentary handheld camera movement, pushed film grain, silent and isolating mood

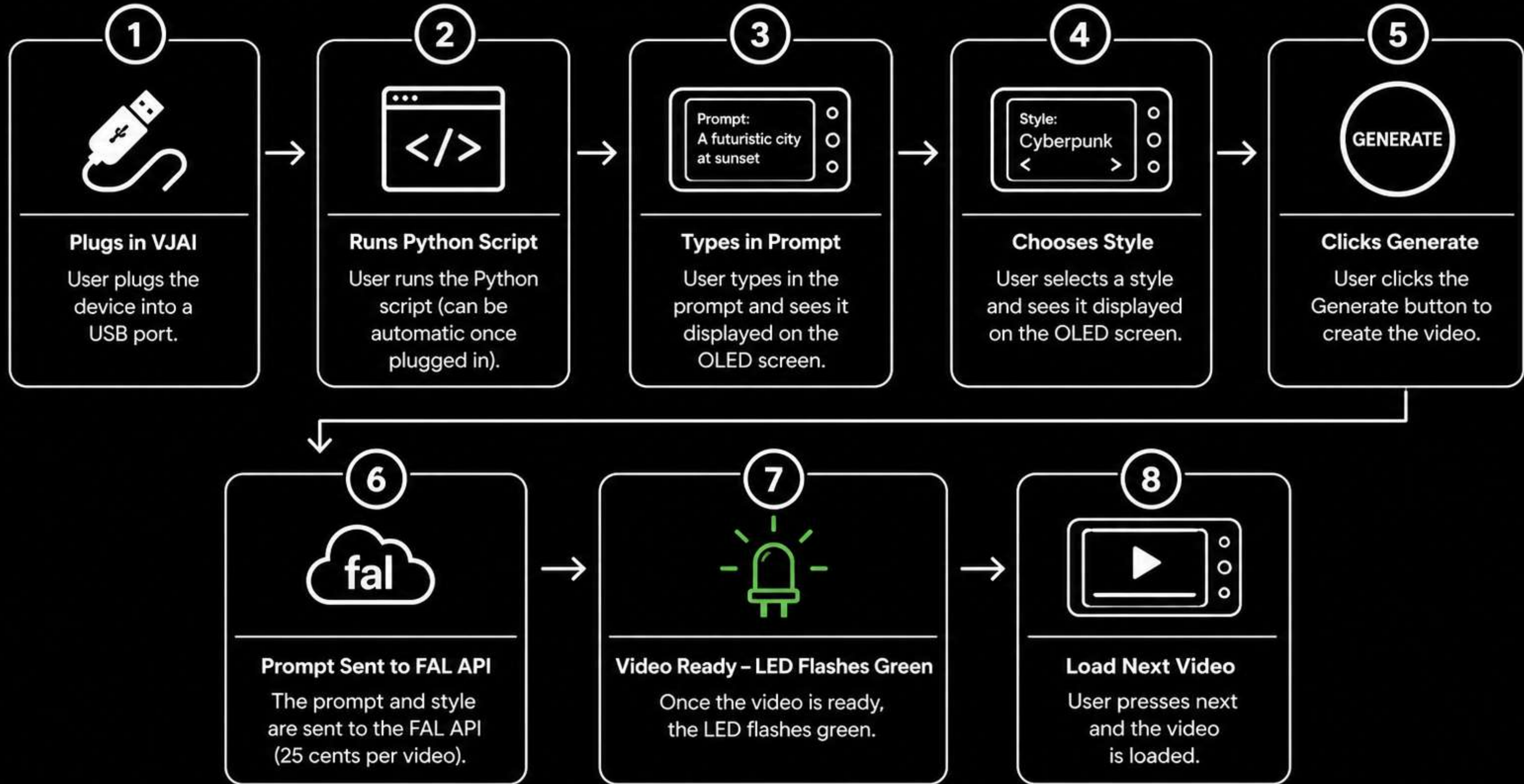
GLITCH: (Typed prompt here) Stroboscopic multiple exposure technique, overlapping ghost frames and temporal fragmentation, film burn and damage artifacts, violent strobe timing, experimental underground film aesthetic, disorienting push-pull camera, heavy grain

The way it works:

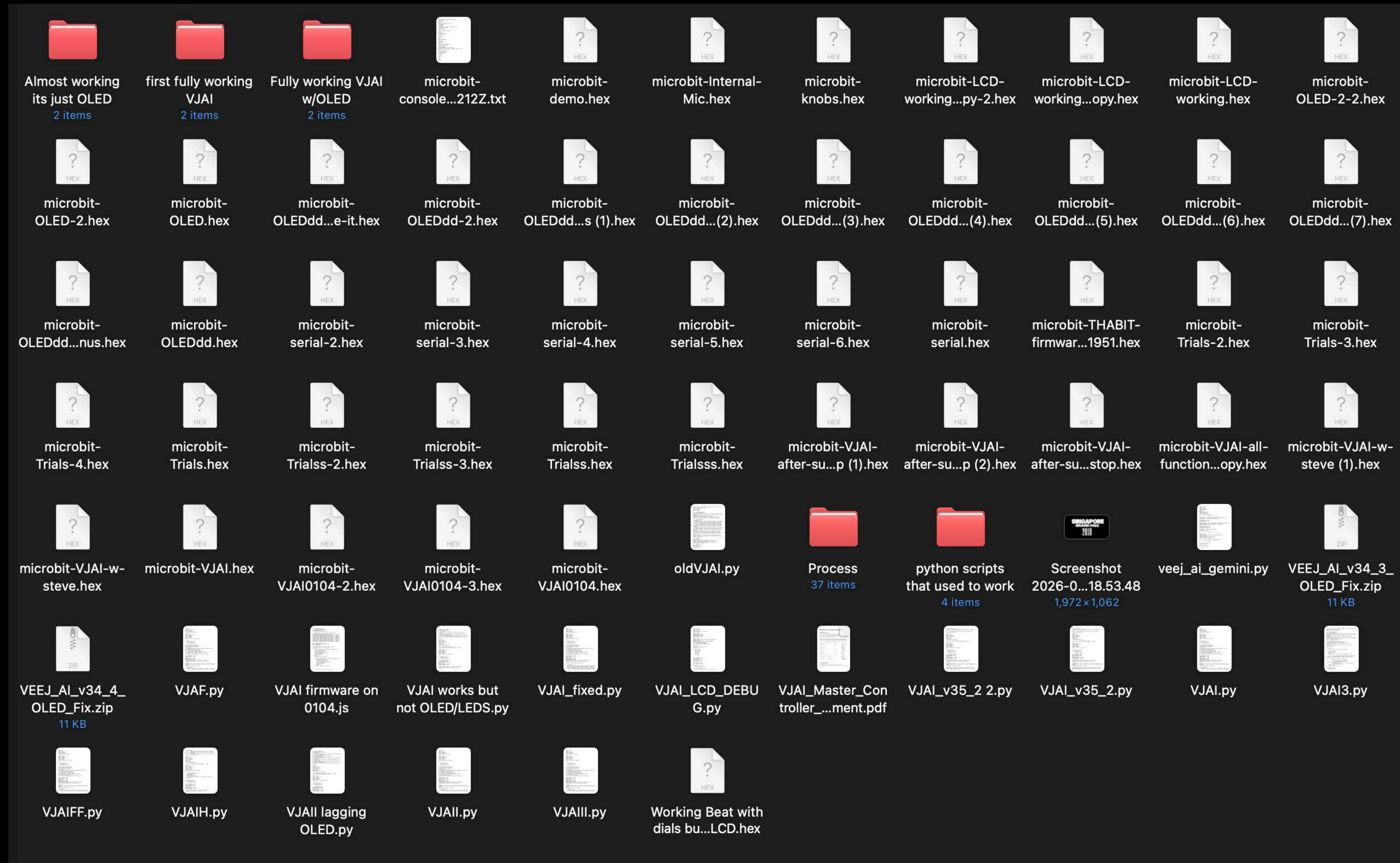
- User runs the python script (can be automatic once plugged)
- Types in the prompt, sees it in OLED
- Chooses the style, sees it on OLED
- Clicks Generate button
- Prompt is sent to FAL API (25 cents per video)
- Once video is ready LED Flashes Green
- User presses next and video is loaded.



User Journey



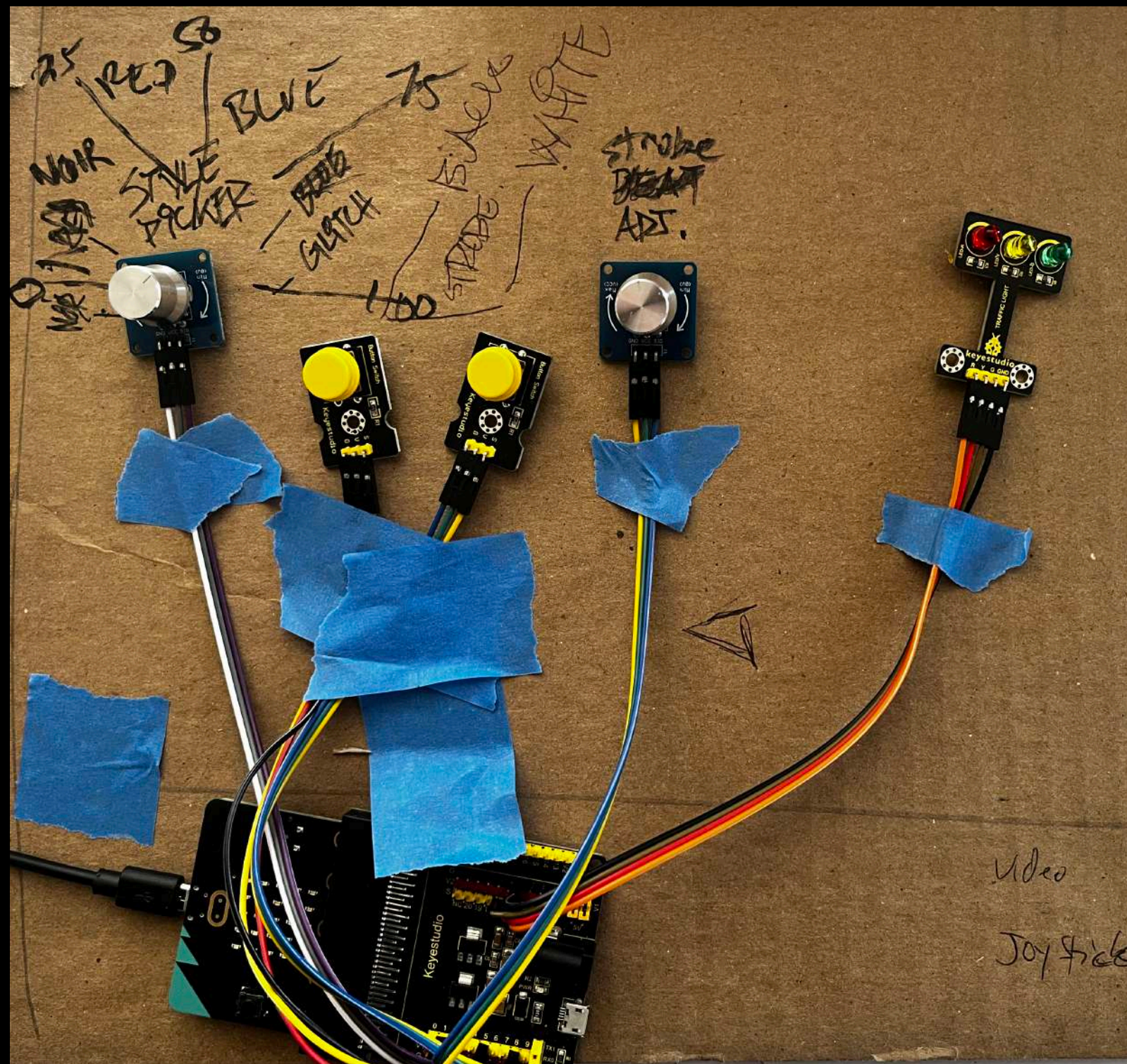
Trials and error



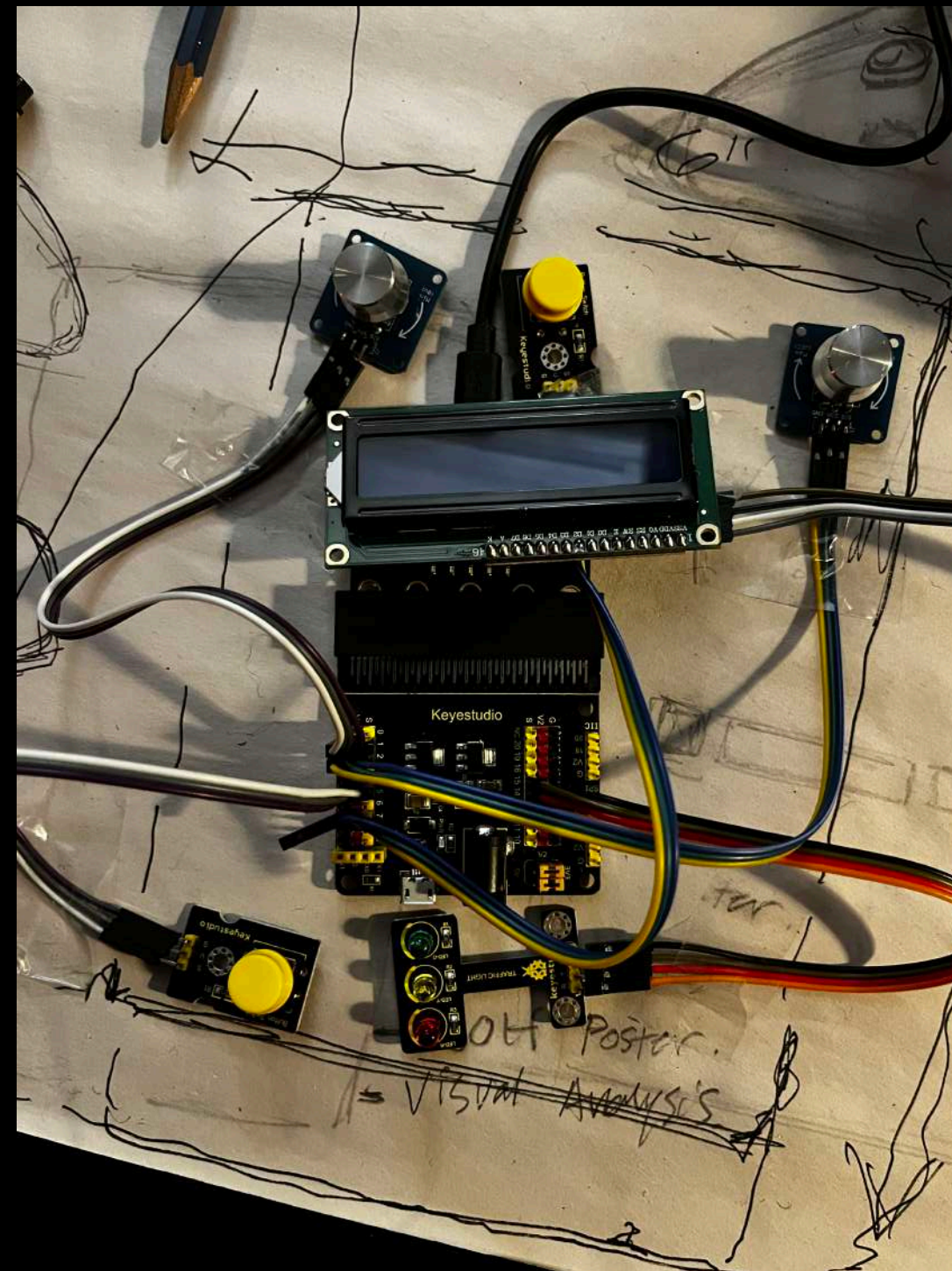
The amount of files to figure out the best way it works and to debug issues along the way. Python files are the script and the hex files are the firmware for the micro-bit.

Hardware Process

Hardware Development



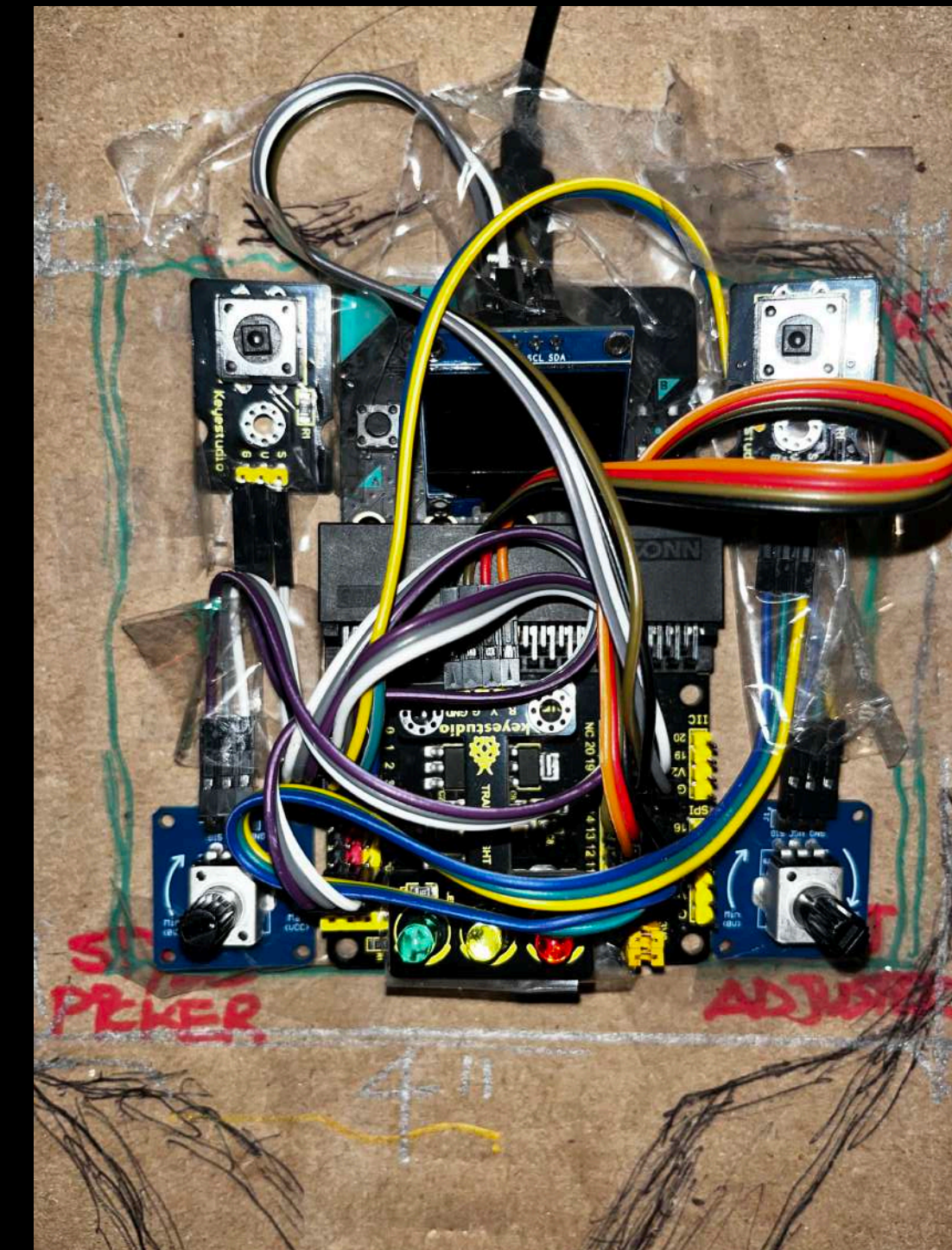
Laying out the components and labelling each function to see how a user would interact.



Deciding to add an LCD screen to show status + feedback instead of laptop screen.



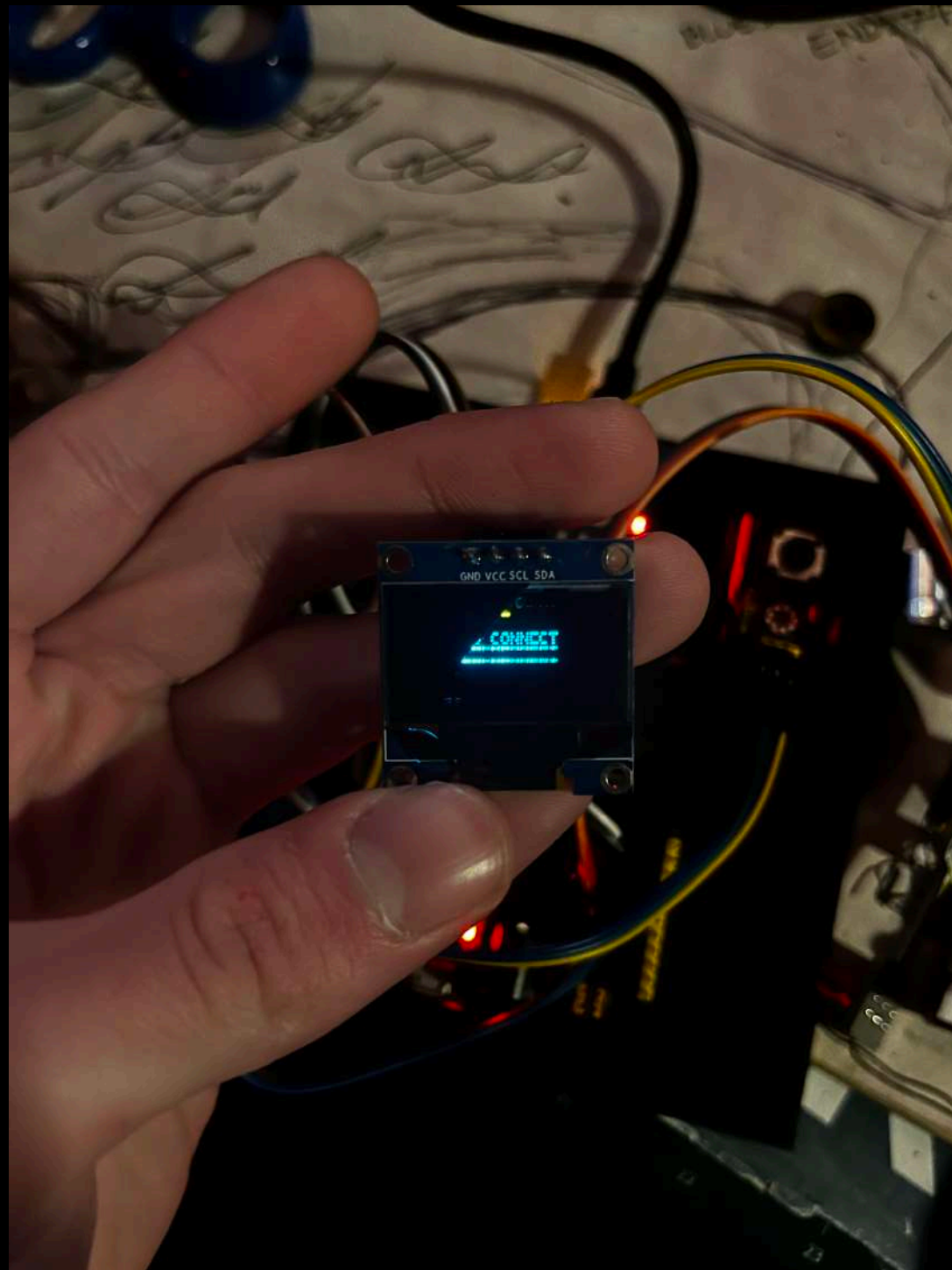
Adding an OLED screen as the LCD screen was too limited in terms of things to display.



Final arrangement to design an outer shell around to make it feel real.

Some Mishaps along the way

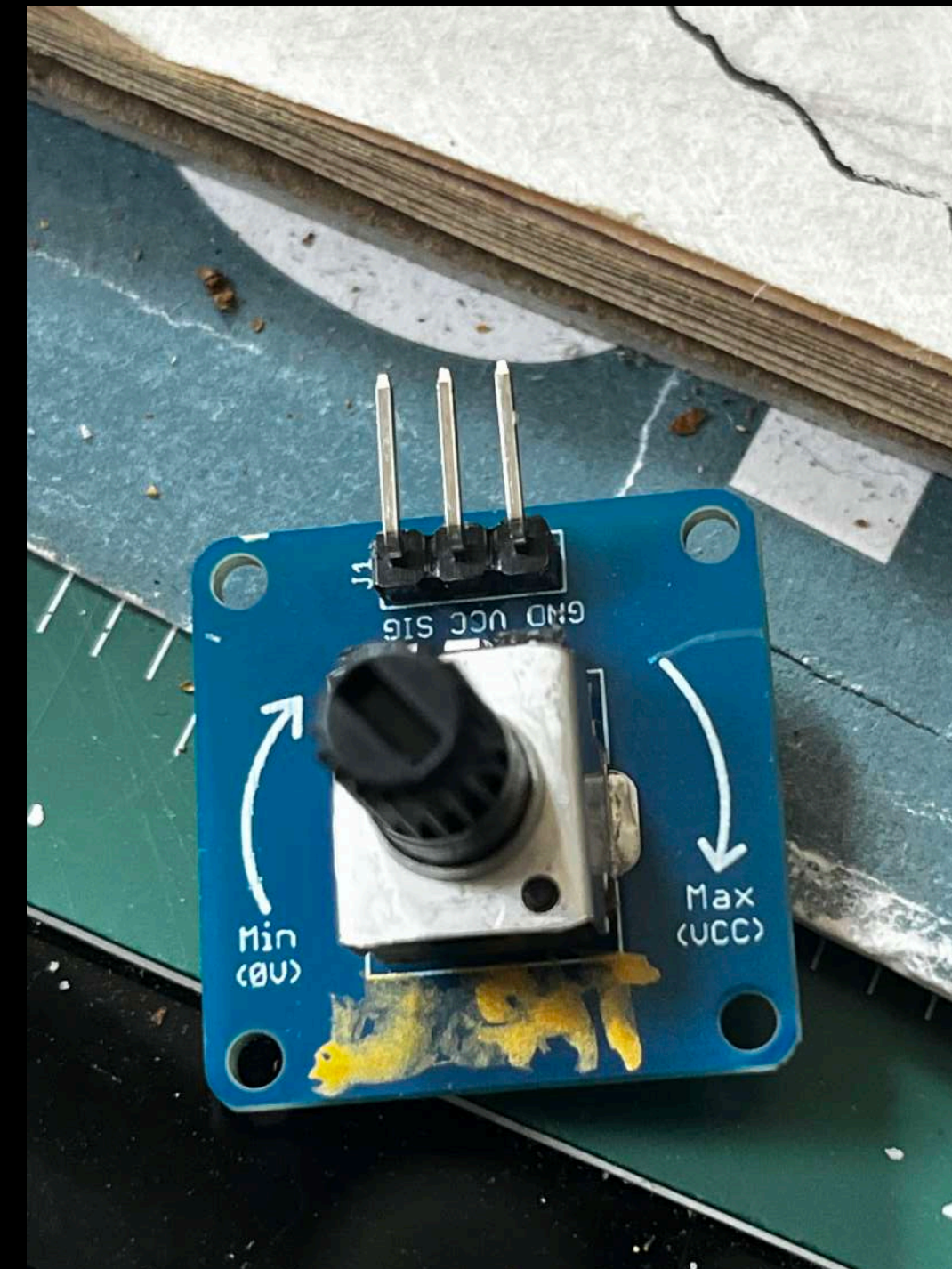
noun.
An unpredictable outcome that is unfortunate.



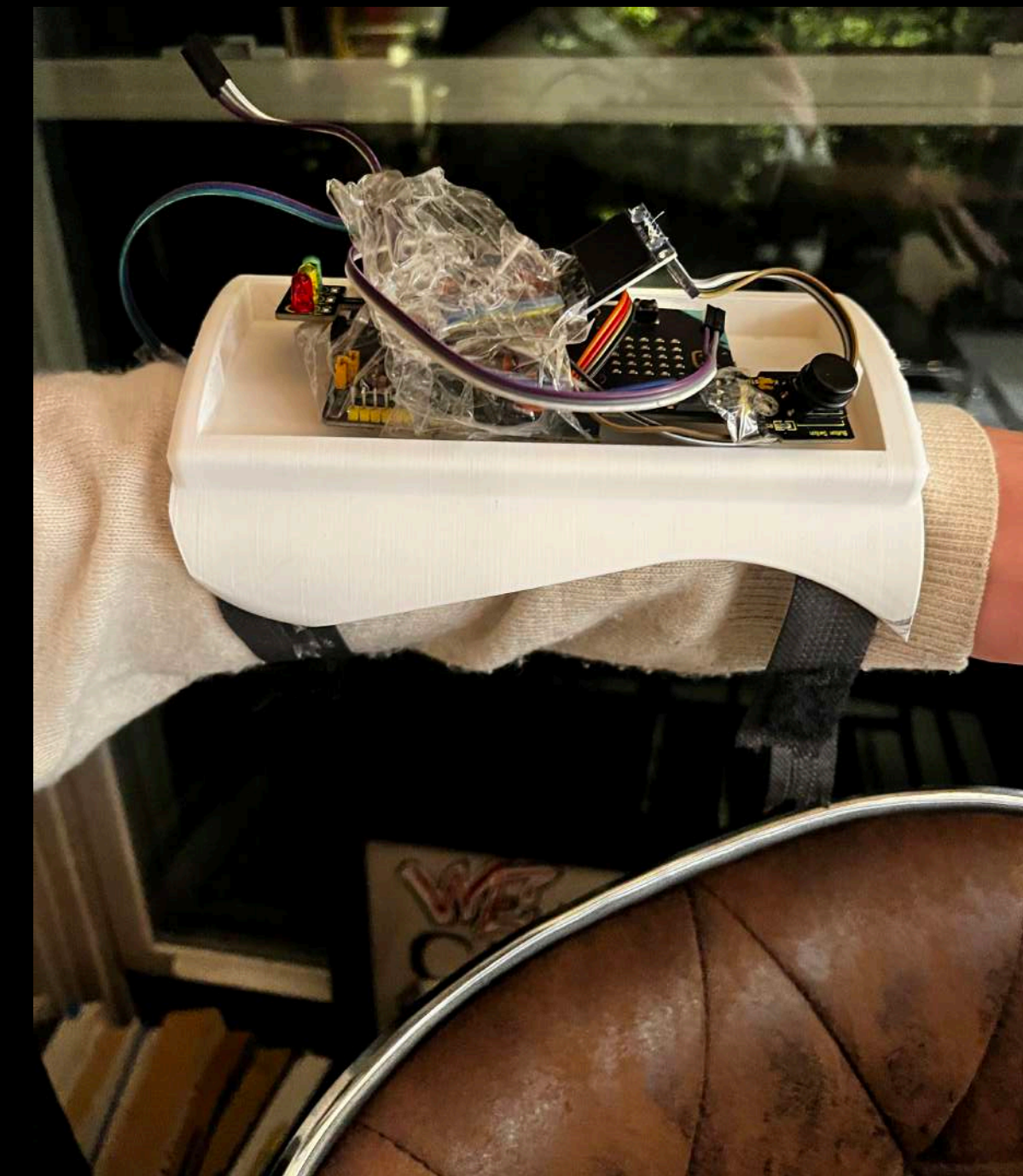
Broken OLED Screen.



Many iterations to get the tolerances right.



Burnt Rotary encoder making the whole solution not work.

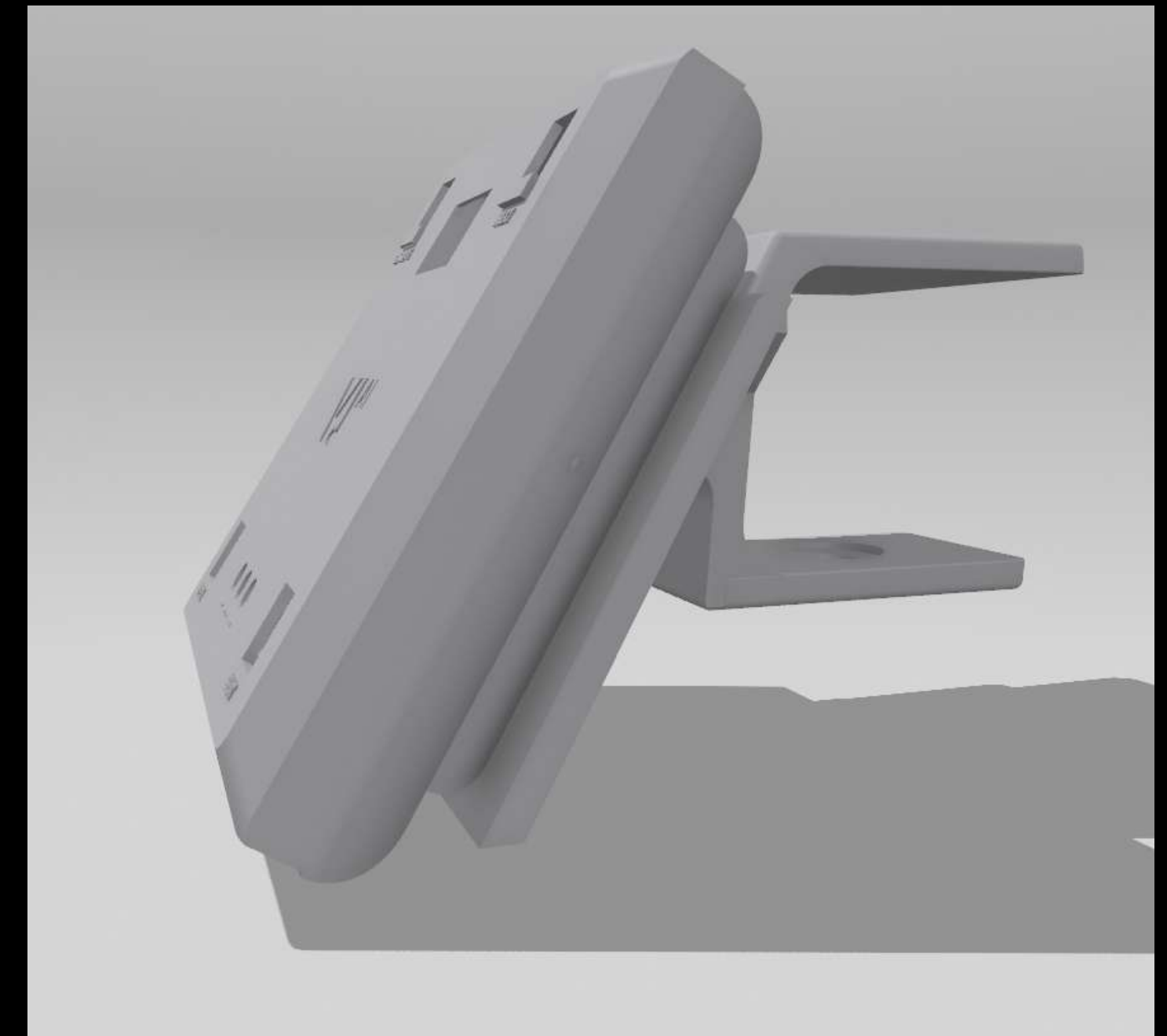
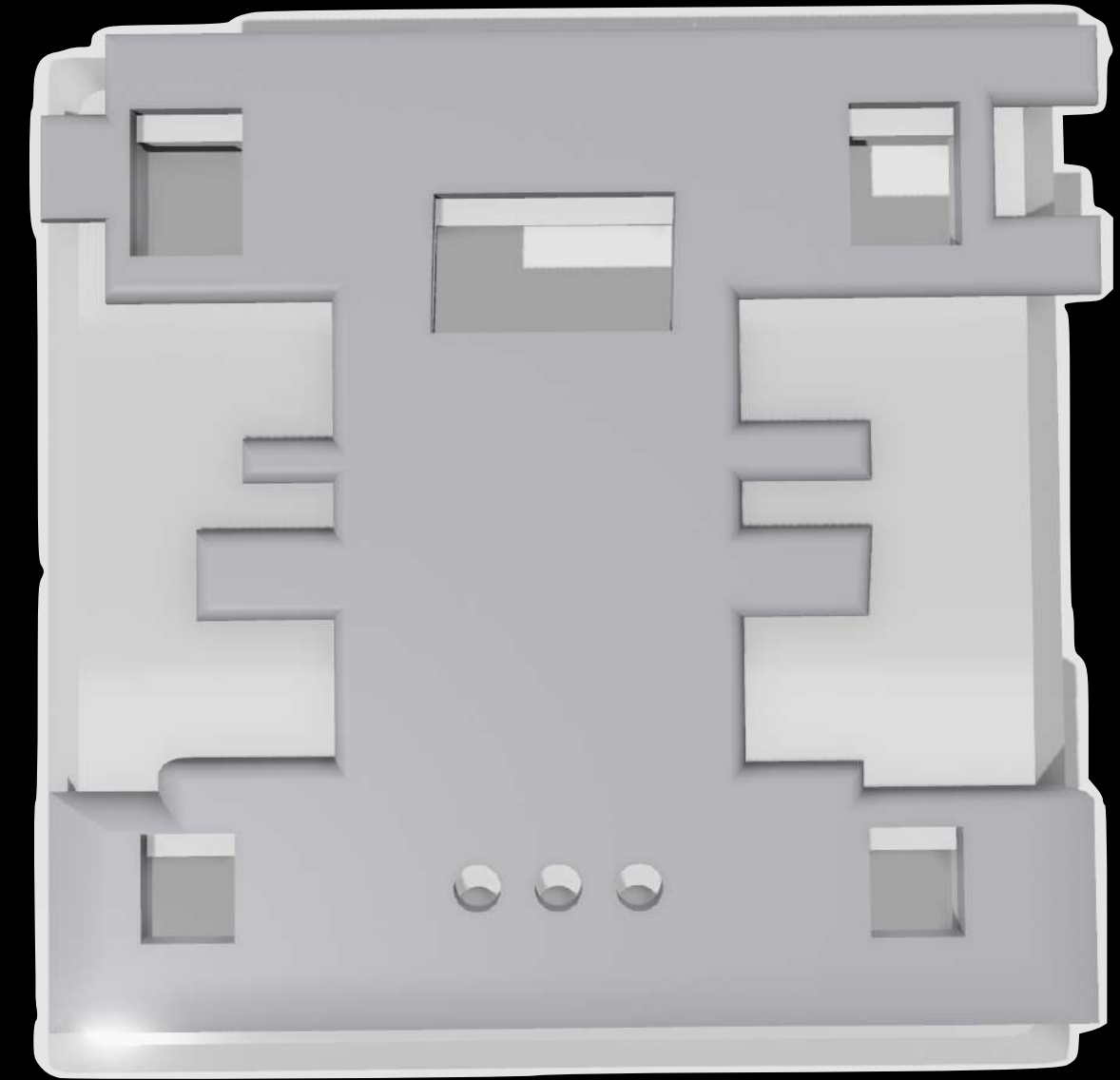


Exploring other forms.

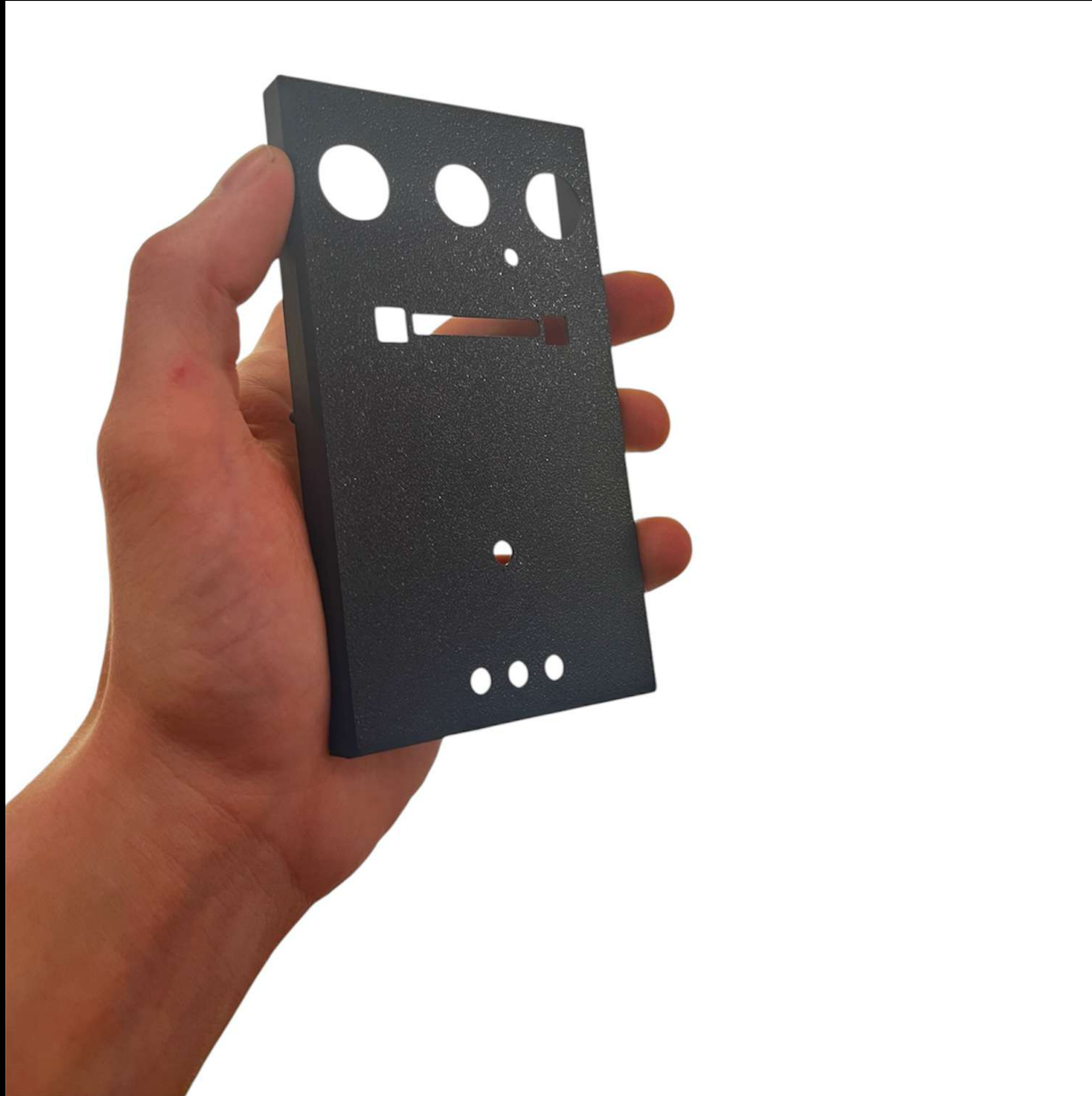
Exploring different forms



Clamping mechanism so it sits outside of the booth.

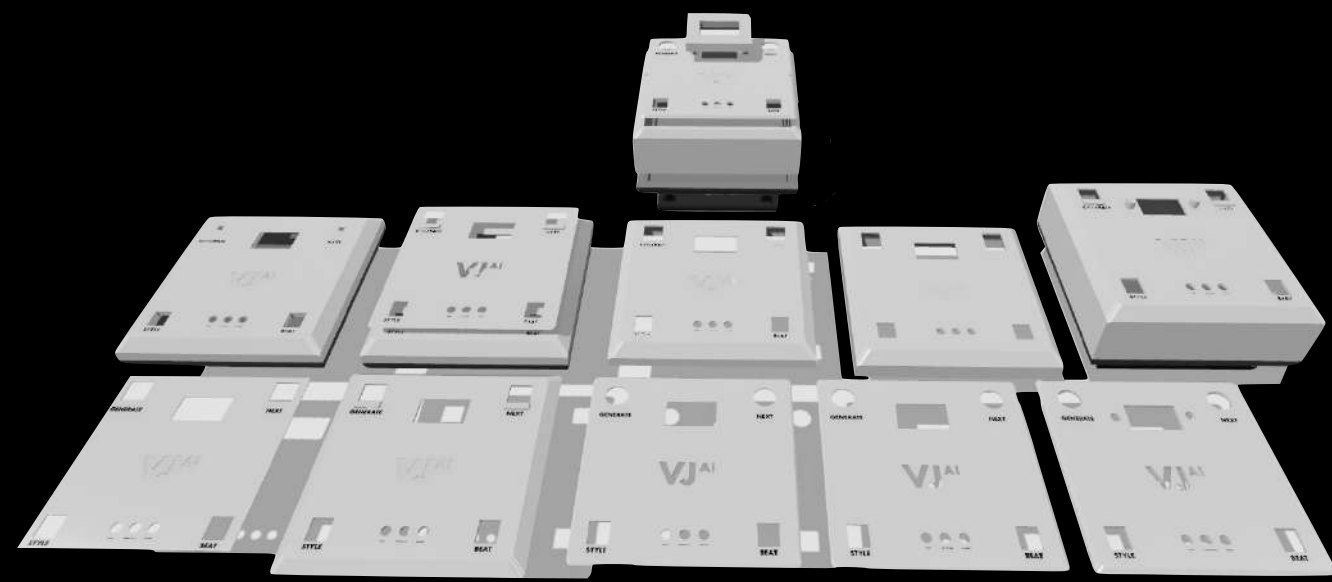


1st attempt

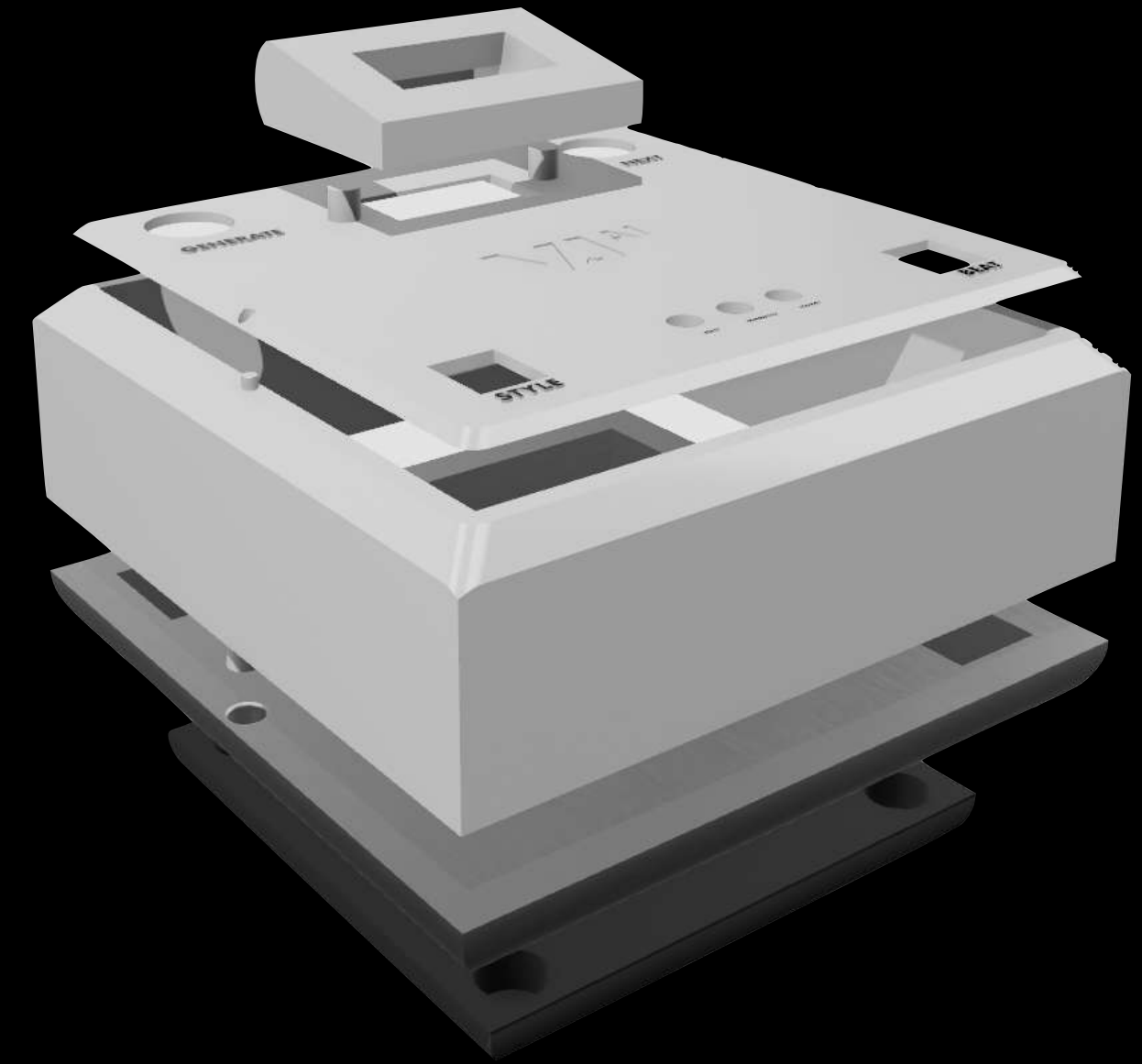


After some sketching, I wanted to 3D print a cover to see what fits and what doesn't. This did not fit at all.

Solutions



**Modular Booth Box
(Functional Solution)**



**Arm-Strap
(Speculative Solution)**



The amount of files that went to print



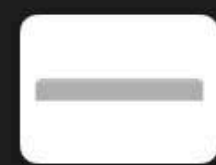
head new.stl



split body top
new.stl



split body bottom
new.stl



head copy.stl



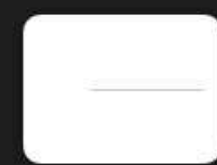
VJAI Arm
2 items



VJAI perfect fit
5 items



thicker top.stl



top w holes.stl



bottom split.stl



top split.stl



head exp.stl



top layer with
connect.stl



cover.stl



head.stl



VJAI toppp.stl



Clip'
4 items



arm VJ.stl



bottom no hole.stl



tester for vj.stl



bottom w hole
correct.stl



VJAI 2 top perf
tol.stl



VJAI bottom.stl



Tester final.stl



VJAI top failed.stl



tester for mini
bed.stl



VJAI 22.stl



inner lining.stl



22.stl



vjai bottom w
slot.stl



LED
sketches.shapr



VJAI bottom.stl



VJAI top 2.stl



VJAI bottom w
details.stl



MicroBit%20refre
nce.stp



button trial.stl



BBC_BOTTOM.stl



BBC_TOP.stl



VJAI top.stl



vjai top new.stl

AI Visualizations (Vizcom, Gemini, ChatGPT)



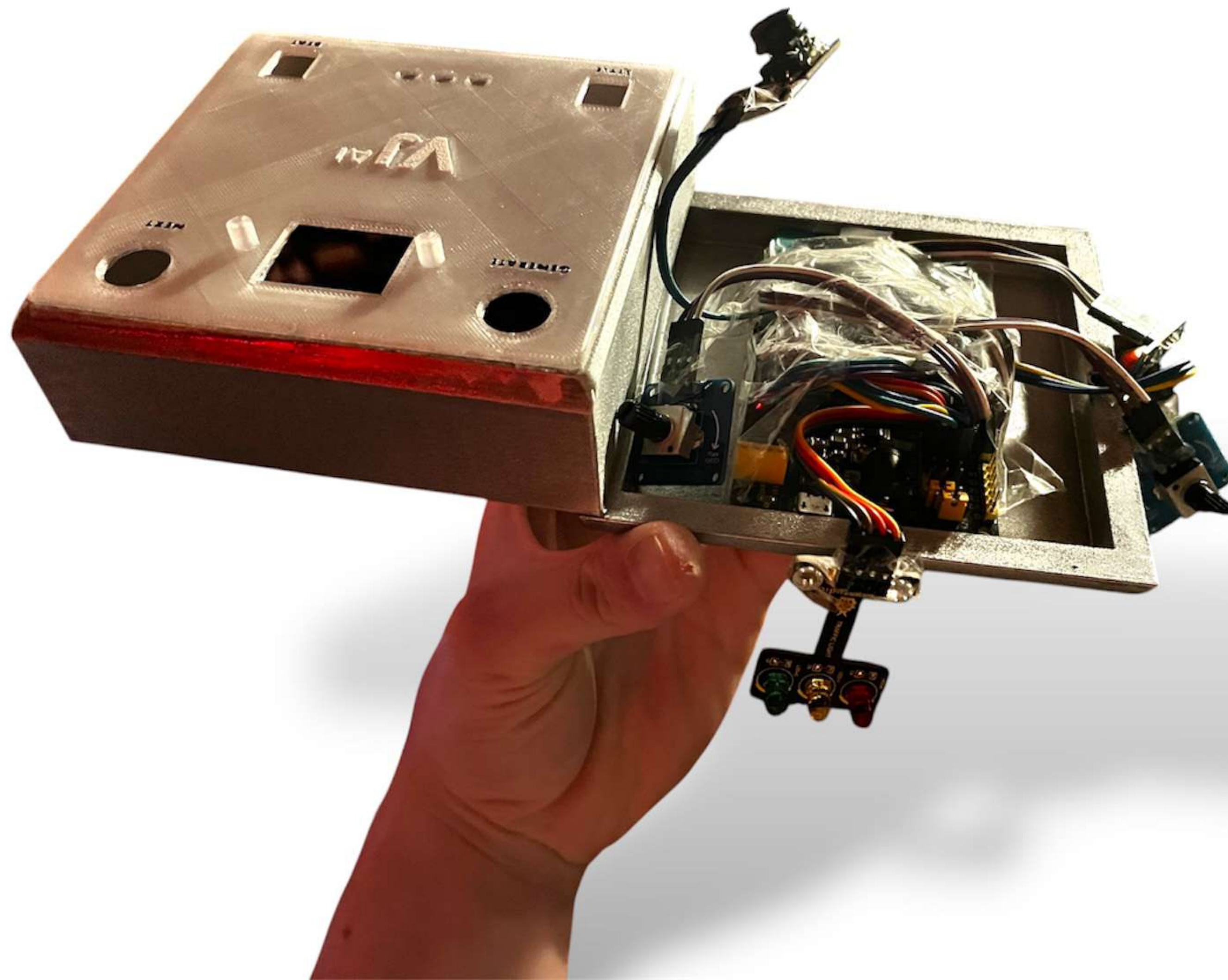
Finalization

1st prototype (almost perfect)



This was fully functional but I disliked the case for the OLED screen and the visible side snap fit mechanism and the overall color as well.

Inside View



2nd & Final prototype (perfect)



What changed here was that I decided to spray paint it to look more finished as well as changing the casing for the OLED to be more flush and snap fit mechanism to be fully concealed in addition to making the top with transparent filament and adding the VJAI letters.

Next Steps

Things to consider:

- **Via bluetooth and not wire**
- **Built in projector**
- **Wearable hardware**
- **Buttons of styles instead of dial**
- **Multiple screens that show Style + Videos**
- **Consider using ESP32**
- **Plug-in to existing VJ softwares**
- **Syncs to lyrics of the music not only reacts to it; Spotify lyrics, etc.**



**Keep in touch &
Check out my other work!**

Saif Bajnaid
[saif@m3nv.com](mailto:satif@m3nv.com)
www.bajnaid.nyc



VJ^(AI)