

2026 in review

12 changes across 3 quarters

Every episode of My Weird Prompts is written, voiced and assembled by an automated pipeline. This changelog records its major updates — new models, new voices, new stages — and the first episode each one produced. Small prompt tweaks happen weekly and are not listed.

Contents

- Q1 2026** January – March · 1 change
7 January 2026 — The voices come in-house
- Q2 2026** April – June · 6 changes
11 April 2026 — Claude Sonnet 4.6 writes every script
14 April 2026 — Prompts 4.0, and the hosts start remembering
16 April 2026 — DeepSeek takes over
25 April 2026 — DeepSeek V4 Pro
30 April 2026 — Studio-style mastering
20 June 2026 — Character lorebooks
- Q3 2026** July – September · 5 changes
18 July 2026 — Scripts written in one pass again
23 July 2026 — A research agent before every script
1 September 2026 — V5.1: versions that mean something
1 September 2026 — V5.2: primary sources for technical topics
20 September 2026 — DeepSeek V4.1 Flash

Q1 2026

January – March 2026 · 1 change

7 January 2026 · Voices · Pipeline V4

The voices come in-house

Chatterbox now runs on the show's own Modal GPUs instead of a hosted API. Every episode since has been voiced this way.

- Switching from Chatterbox Turbo to Chatterbox Regular cut TTS hallucinations (garbled or invented audio) by about 95%.
- Each host's voice is pre-computed and cached, and an episode's lines are spread across parallel GPU workers. Voicing an episode went from about 36 minutes to about 10.

FIRST EPISODE: [#183 The Multi-Billion Dollar Metal We Leave Underground](#)

Q2 2026

April – June 2026 · 6 changes

11 April 2026 · Script model · Pipeline V5 · Prompts 3.5

Claude Sonnet 4.6 writes every script

Gemini Flash kept writing episodes that ran short, so long-form episodes went to Anthropic's Claude Sonnet 4.6 first (from 8 April). Three days later Sonnet was writing all of them.

- Smaller stages ran on Claude Haiku 4.5.
- Voice memos are now transcribed by OpenAI Whisper.

FIRST EPISODE: [#2158 Claude Managed Agents: Brain Versus Hands](#)

14 April 2026 · Writing · Pipeline V5 · Prompts 4.0

Prompts 4.0, and the hosts start remembering

The script prompt was rewritten as version 4.0. The same day, each host got a written character card and a memory layer that looks up earlier episodes by meaning, so the hosts can refer back to things they've already discussed.

FIRST EPISODE: [#2219 Spec-Driven Life: How AI Planning Beats Project Paralysis](#)

16 April 2026 · Script model · Pipeline V5 · Prompts 4.0

DeepSeek takes over

Scripts moved from Claude Sonnet to DeepSeek V3.2, at about a thirty-sixth of the price per output token. Research grounding also got a dedicated section and a three-round search loop.

FIRST EPISODE: [#2252 Why Lithium-Ion Won \(And What's Next\)](#)

25 April 2026 · Script model · Pipeline V5 · Prompts 4.0

DeepSeek V4 Pro

Scripts moved to DeepSeek V4 Pro, called directly through DeepSeek's own API, with its lighter Flash sibling handling the surrounding stages. V4 Pro wrote close to 3,000 episodes, more than any other model in the show's history.

FIRST EPISODE: [#2403 Choosing Your LLM Eval Framework](#)

30 April 2026 · Voices · Pipeline V5 · Prompts 4.0

Studio-style mastering

Assembled episodes now get a mastering pass: loudness normalisation, fades, room tone between lines, and EQ tuned to each character's voice.

FIRST EPISODE AFTER RELEASE: [#2550 Idempotent Pipelines: Checkpoints, Manifests & Safe Re-Runs](#)

20 June 2026 · Writing · Pipeline V5 · Prompts 4.0

Character lorebooks

Each host now has a lorebook of backstory, and the entries relevant to an episode's topic are pulled into the script prompt. The hosts' personal history comes up when it fits instead of being repeated in every episode.

FIRST EPISODE AFTER RELEASE: [#3776 ZFS Mirroring: Why Your RAID Card Is the Weak Link](#)

Q3 2026

July – September 2026 · 5 changes

18 July 2026 · Writing · Pipeline V5 · Prompts 4.0

Scripts written in one pass again

For a few weeks, scripts were written section by section. That produced episodes that restarted or repeated themselves, so scripts are written in a single pass again.

FIRST EPISODE AFTER RELEASE: [#4380 Why 30 Bases, Not 1 Super-Base in the Middle East?](#)

23 July 2026 · Research · Pipeline V5 · Prompts 4.0

A research agent before every script

Before any writing starts, an agent now researches the topic over several rounds, using web search, open news data (GDELT), PubMed for medical topics, and source code repositories. What it finds becomes the script's research brief.

FIRST EPISODE: [#4444 Testing the Unpredictable: QA for Agentic AI](#)

1 September 2026 · Writing · Pipeline V5.1 · Prompts 4.1

V5.1: versions that mean something

This was the first version bump in nine months, and from here every change that can affect an episode gets one. Prompts 4.1 targets verbal tics that were measured across the whole episode archive, rather than ones that were guessed at.

FIRST EPISODE: [#4966 Local vs Cloud: Running Hugging Face Models](#)

1 September 2026 · Research · Pipeline V5.2 · Prompts 4.1

V5.2: primary sources for technical topics

For technical prompts the research agent now reads arXiv papers and Hacker News discussion, not just whatever a web search ranks highest. It also reports what it looked for and couldn't find, so the hosts don't fill the gap with something that only sounds similar.

FIRST EPISODE: [#4970 Labels vs. Book Covers: Structured Extraction Tradeoffs](#)

20 September 2026 · Script model · Pipeline V5.2 · Prompts 4.1

DeepSeek V4.1 Flash

DeepSeek had started answering V4 Pro requests with V4.1 Flash until V4.1 Pro ships. The pipeline now asks for V4.1 Flash by name and uses it for every stage, because each episode ends by naming the model that wrote it, and that credit should be accurate.

- Images can now be sent with a prompt. Up to five images go straight to the script writer.
- The hosts now receive Daniel's prompt exactly as he wrote it, not a tidied rewrite.
- The model credit at the end of each episode is now read by Hilbert, the show's producer.

FIRST EPISODE: [#5391 The Software Behind Urgent Care Triage](#)