Introduction

I've started this project after fiddling around with Notions "Meeting Notes" feature and running the idea of using a Speech to Text tool for our session transcripts by the members of one of my TTRPG groups.

Consensus was that a tool like this would be nice and helpful to have good transcripts and a helpful session summary without taking time out of player's lives to compile concise session notes. For me personally, as somebody with an auditory processing disorder, having a live transcript would also be incredibly helpful to be better able to understand what has been said.

Most group members weren't concerned to use a third-party service, since we wouldn't talk about 'real' or personal stuff, but only a fictional story, concerning fictional characters.

Yet, this solution wouldn't have been ideal and there are real concerns with feeding deeply personal and biometrical information, like a person's voice, to a commercial language model of a company that is based in the USA[1].

That's why I started fiddling around with open-source tools and language models that would run strictly local. With no data transferred to big-tech companies, and this way also not providing information to train commercial models.

My hope is, to create a solution by cobbling together open-source tools and some shell scripts. Ideally, this solution will be easily replicable by everybody who knows their way around a UNIX/Linux command line and has a device available with enough computing-power, RAM and/or GPU.

What TranscriptOMatic most likely won't accomplish anytime soon, is to discern who is speaking.

[1]) And yes, I'm aware of the irony that we already transfer our voices to a US-based service by using Discord voice chat. Doesn't mean one should make it worse.

Created 2026-01-03 22:24:33 UTC by Mela
Updated 2026-02-20 18:55:56 UTC by Mela