**Multimedia Agent Workflow**

[WIP - Might be scrapped]

**Difficulty:** Level 400

Types of Media:

* Subtitles files
* Youtube video

Tool Use

<https://github.com/yt-dlp/yt-dlp>

* Yt-dlp is a generic API to youtube to download videos, convert video to just audio, and possible download transcripts/subtitles
* yt-dlp: single video to audio files
* yt-dlp: playlist to audio files
* Yt-dlp: download subtitles/transcript to text
* Whisper: transcribe japanese audio to japanese text
* Normalize Writing System
  + Take the japanese text and make it use the correct natural writing language
* OCR (with OpenAI) a series of pages for a pdf to raw text
* MapReduce - distill the body of text into the fewest of vocabulary, grammar rules
* Convert the body of text into json structure for import into the main application.