**Project Name:** Dagger File Engine

**Description:** A desktop file browser with enhanced lookup capabilities through both pattern occurrence search and vector similarity search using NLP or other media inputs.

**License:** Mozilla Public License 2.0

**Technologies:**

* Tauri for desktop GUI
* SolidJS framework + TypeScript for WebView
* TailwindCSS for UI styling
* Qdrant for storing and looking up vector embeddings
* CLIP (Contrastive Language-Image Pretraining) for image embeddings
* CLAP (Contrastive Language-Audio Pretraining) for audio embeddings
* all-MiniLM-v6 for sentence embeddings
* Tantivy for text occurrence search
* ONNX runtime (ORT crate) for ML model bindings
* notify crate for file state change watches
* walkdir crate for file navigation

**Features:**

* String Occurrence Lookup
  + The Dagger File Engine provides a search bar that allows users to search for files by given text patterns that exactly match file contents.
* Vector Similarity Search
  + The Dagger File Engine provides search options for a given set of media types:
    - Image similarity lookup by text queries or drawing pad, using OpenAI’s CLIP (Contrastive Language-Image Pretraining) for image multimodal embedding.
    - Audio similarity lookup by text queries, microphone input, or audio, using OpenAI’s CLAP (Contrastive Language-Audio Pretraining) for multimodal embedding.
    - Text similarity lookup by text queries or file input, using all-MiniLM-v6 for semantic embedding.
  + All dense vector embeddings are stored in a local Qdrant database for similarity search at runtime.
* File State Mutation Strategy
* File Browsing and Navigation