This PDF is a description of my most recent projects

[Archivy](#) - [Source Code](#) - [Hacker News](#)

Archivy is an open source knowledge management project.

I built Archivy in summer 2020 because I wanted a community-driven, open source, extensible note-taking system that could assimilate web content. Archivy software has now been downloaded hundreds of thousands of time on [Docker](#) and the [Python Package Index](#), and is also available on the AUR and Nixpkgs.

Archivy is both a command line interface that allows you to run your local knowledge base on your computer, and a web application you can access from anywhere. It has developed into a popular project due to its appealing design, extensive organization features and its potential for web web archival. I've been adding [many features](#) and plugins, notably [archivy-hn](#) / [archivy-pocket](#) to download and save online content automatically and [archivy-static-site-gen](#), which allows you to turn your Archivy instance into a read only public website.

As founder and maintainer of the project, I built the frontend interface, set up the hosting and distribution mechanisms, and engineered the backend logic that allows people to learn through Archivy. I also get contributions from developers who want to see their own ideas implemented in Archivy, and have to take into account the needs of contributors and users to lead the community in the best direction possible,

Archivy is written in Python with the Flask web framework. I use vanilla javascript and CSS for the frontend design and interactivity. Archivy documentation is hosted on GitHub Pages.

[AdiosCorona](#)

Note: Our organization's source code is not public.

AdiosCorona is a non-profit organization created at the start of the pandemic, led by a group of international scientists. It reviews recent academic literature and then distributes clear, illustrated information on the virus to people around the world.

I'm the lead developer of the organization, and I built the website from scratch in 2020. I designed the frontend and engineered a backend for the scientists to add articles. AdiosCorona is built on NetlifyCMS and Jekyll, with a custom theme I made in Figma and then HTML/CSS after discussions with designers. It's hosted on Netlify and DNS is handled through OVH.

We've had ~1.5M unique visitors since 2020 and [have been mentioned by international media and institutions](#), including the French Health Ministry and The Conversation.

Espial

Espial ("discovery") is a new project I've been working on, that will be launched and open-sourced soon. Espial seeks to enhance and augment content discovery and automatic organization in knowledge management software, using Natural Language Processing.

Through my work on knowledge management software with Archivy, I noticed most programs in this space burden their users with organization, as they end up spending too much time tagging and linking content. None of them actually recommend automatically detected hierarchies and links between your notes. I've been developing this program to solve that issue, using Natural Language Processing.

Espial is a backend API built in Python. I use the Spacy library to process text. From this processed text, I generate a graph of interlinked concepts and notes with the networkx graph library, which is then rendered with D3.js. Espial uses word embeddings and Named Entity Recognition to extract this web of "concepts" from the notes. This data structure ends up collecting many uninteresting concepts, so my algorithm applies statistical heuristics like term frequency, inverse document frequency, and document similarity to remove them, only keeping those which actually revealed relevant links between notes.

Espial isn't just the graph represented in the video, it's an entire API interface to analyze and process text in your knowledge base. It can tell you which documents could be relevantly linked together and which detected concepts could interest you. Its semantic search interface also allows you to input large amounts of content or articles, and it then searches for related ideas in your own knowledge using trained word embeddings.

This open-ended project has been an incredible learning experience in modern Natural Language Processing. Espial is effectively helping me discover new connections and patterns between my notes and articles. I'm excited to launch it soon.

This project was first started at the Summer Program in Experiental and Applied Reasoning, under the mentorship of [Gavin Leech](), AI PhD at Bristol University.