

<https://github.com/vectara/vectara-ingest/pull/43>

This is the [Arguflow](#) team's submission to the "best contribution to vectara-ingest" part of the hackathon. We really admire the work that has gone into this repository, and want to start a trend of making it compatible with more varied services.

This PR adds support for [Arguflow](#) to the crawler such that users are able to add documents/chunks to Arguflow, Vectara, or both.

Internally, we were motivated to add this support so that we can stand up more Arguflow demos by using the crawlers, however, we are PR'ing it because we think it can also offer Vectara users value.

Arguflow has support for a few things that Vectara does not which users may desire:

- [self-host option for E2E data ownership](#)
- [support for arbitrary LLM's and text-embedding models](#)
- [RAG options beyond summarization \(both fully managed and search-before-generate\)](#)
- [enhanced duplicate detection beyond UPSERT](#)

Arguflow also ships with a [default search and chat UI](#) in addition to its [OpenAPI spec](#) which makes it a bit easier to get going with a deployment you can share and reverse-engineer to build your own applications.

This repository is great, and we really admire [@ofermend](#)'s work on it especially. Excited to enhance it and bring it more into the lens of the open source AI world! We also edited the documentation in all the right places, let us know if may have missed a few spots.

Happy to address any review comments or change requests in a timely manner.