

JIM CRIST-HARIF

✉ jcristharif@gmail.com | 🌐 [jcrist](https://jcrist.com) | 🌐 jcristharif.com | 📍 Saint Paul, MN

SUMMARY

An experienced software engineer and manager with a proven track record of developing OSS data analytics tools that users love.

LANGUAGES

- Proficient: Python | Cython | C
- Capable: SQL | Go | Java

TECHNOLOGIES

- K8s | Apache YARN | HPC Schedulers
- DuckDB | Polars | Dask | Ibis

INTERESTS

- Data Analytics Tooling
- API Design / Developer Experience
- Distributed Systems
- Open Source Community Health

EXPERIENCE

Engineering Manager — Voltron Data

Jun 2022 - Oct 2024

- Led the team developing [ibis](#), an ergonomic dataframe API wrapping 20+ SQL databases.
- Interfaced with Product and Leadership to turn product requirements into engineering tasks.
- Authored Ibis's backend wrapping Voltron Data's distributed GPU engine.
- Created [ibis-ml](#), an ML framework for preprocessing and feature engineering using Ibis.

Senior Software Engineer — Coiled

Aug 2021 - May 2022

- Led an effort to revamp [dask](#)'s [parquet](#) support, improving performance by up to 8x.
- Optimized [dask](#)'s networking layer, reducing RPC latency by 2x.

Senior Software Engineer — Prefect

Apr 2020 - May 2021

- Led development of Prefect's core workflow framework.
- Interfaced actively with users and customers, helping resolve issues and improve user experience.
- Optimized Prefect's [dask](#) integration, allowing for resilient and scalable execution of large workflows.
- Designed and implemented platform features, working across the full software stack (backend, frontend, and core SDK).

Senior Software Engineer — Anaconda

Jun 2015 - Apr 2020

- Core contributor to [dask](#) & [distributed](#), a parallel and distributed compute framework for Python. Wrote much of the original implementation of [dask](#) and [dask-ml](#).
- Created [datashader](#), a high performance plotting library for large data.
- Identified and resolved issues deploying Python in enterprise Hadoop environments. This resulted in the development of several new tools, including a [complete JupyterHub deployment](#).
- Created and led development of [dask-gateway](#), a tool for managing and deploying Dask in multi-user enterprise environments. Supported Kubernetes, Apache YARN, and HPC schedulers as deployment backends. Helped several large organizations deploy [dask-gateway](#) to serve hundreds of users.
- Executed [dask](#) consulting engagements, helping several Fortune 500 companies use [dask](#) in production.

OPEN SOURCE

- [Dask](#) steering council member
- [Ibis](#) steering council member
- Creator and lead maintainer of [msgspec](#), currently the fastest JSON validation library for Python.
- Presented [talks and tutorials](#) at 30+ past conferences including Strata, PyCon, SciPy & PyData.
- Past mentor for [Google Summer of Code](#).

EDUCATION

MSc. Mechanical Engineering (unfinished) — University of Minnesota

Sep 2013 - May 2015

Designed novel state estimation algorithm for linear actuator system.

BSc. Mechanical Engineering — University of Minnesota

Sep 2009 - May 2013