

Data Access Kickstart

A simple way to access and manage your data sources with Python

Save your data scientists the time and headache of worrying about data sources, access, and formatting. Our Data Access Kickstart delivers a consistent and simple way for Python users to consume all types of data without requiring customized code.

In a short 100-hour engagement, our experts set your team up with Intake, an open-source data ingest/loading library that enables data scientists to access any type of data set by searching a catalog. We document the setup of Intake and train your team on how to use it so they can get started immediately.

Requirements

To qualify for this Kickstart project, you must have the following:

- Data containers supported by Intake (Pandas, Numpy, Dask, Xarray, etc.)
- Existing data available in (or convertible to) one of the data formats supported by Intake plug-in
- Existing runnable notebooks or scripts demonstrating how to access and use samples of applicable data types

Easier Data Management in a set of four deliverables:



Replace an existing set of runnable notebooks or scripts with custom data-access Python code with references to entries in an Intake catalog



Interface to Intake-supported data container (Pandas, Numpy, Dask, Xarray, etc.)



Final report summarizing the data management plan and how to use and extend the data catalog(s)



Instruction on how to work with data catalogs in your organization

Get Started Today

1

Confirm that Data Access is a good fit for your organization.

2

Contact an Anaconda Kickstart specialist to discuss your project.

3

Start your Data Access Kickstart and manage data with Python.