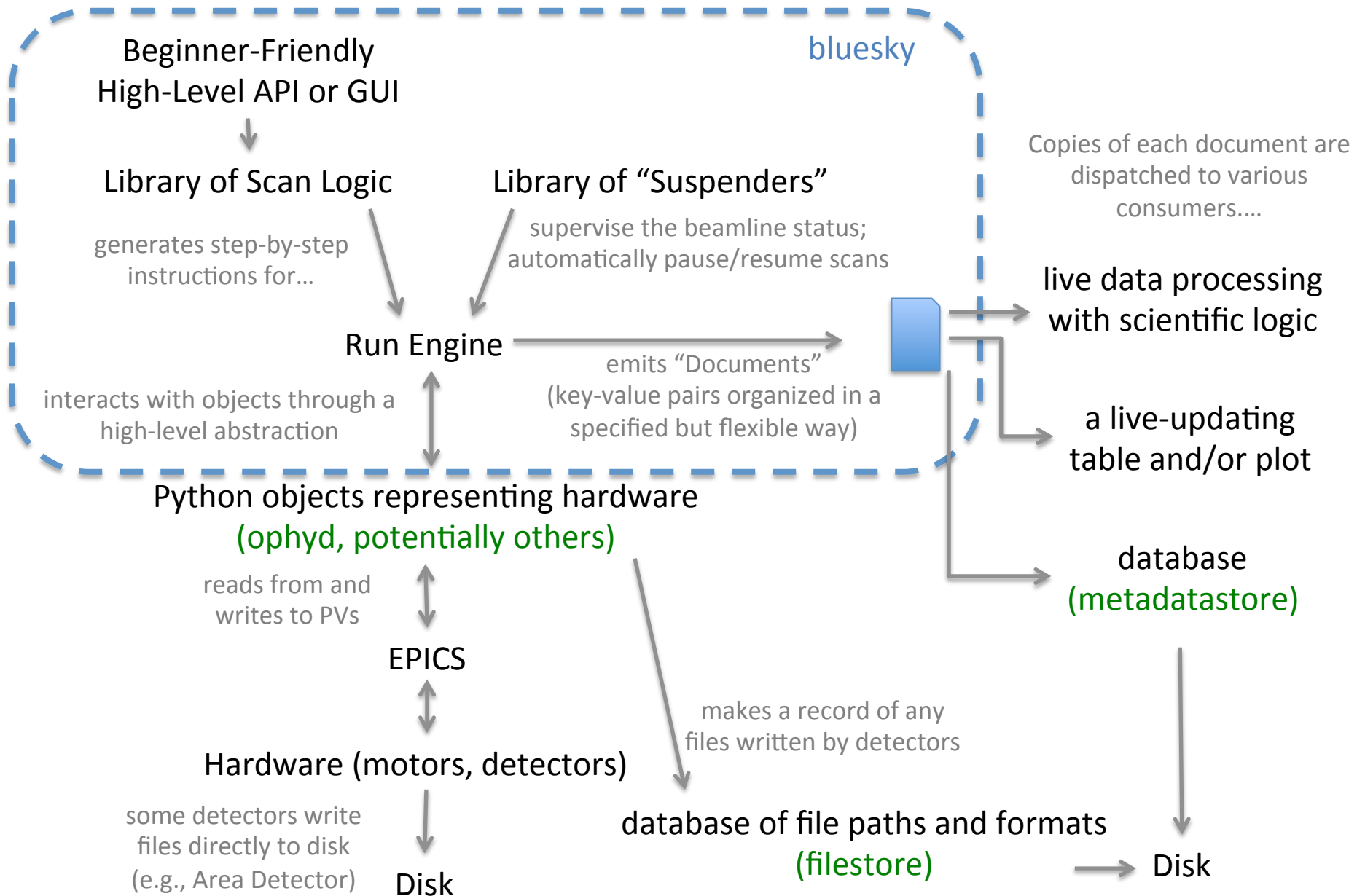
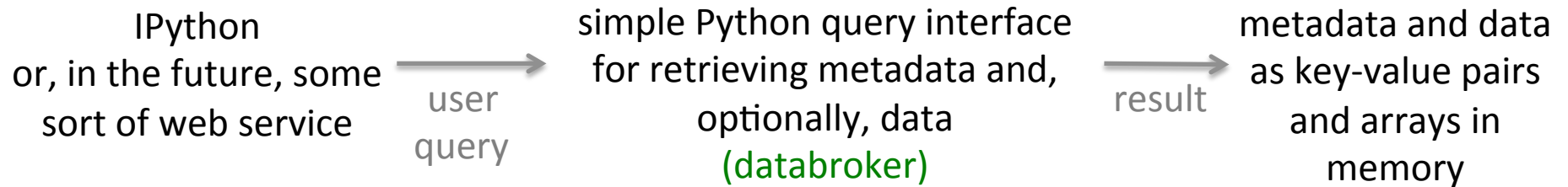


Overview of Data Collection



Overview of Document Model and Data Storage



Under the hood, databroker does all this:

- (1) Get the metadata.
- (2) Load data payload.

metadatastore

- **Run Start:** everything we know at the beginning of the experiment, including open-ended metadata
- **Event Descriptors:** PVs, dimensions, units, open-ended detector-specific metadata
- **Run Stop:** things we only know at the end of the experiment (did it terminate with an error?)
- **Events:** payload of data and references to external files

- (3) Load additional data payload from external files, as needed.

filestore

a database of file paths and file formats with a library of “handlers,” user-configurable code for loading data files into in-memory arrays

+
Disk

- (4) Load archived PV readings during relevant time period.

channel archiver

continuous archival of PV readings at up to 10 Hz

metadata

How will users do analysis at NSLS-II?

Proposals:

**Are you fiercely independent?
Do you have bandwidth to
spare?**

Export your data on site
([suitcase](#)) or download your data
(TBD), install the common
scientific Python packages
([numpy](#), [scipy](#), [pandas](#), [scikit-
image](#), etc.) and our own [scikit-
xray](#), and get to work!

**Do you prefer to drag-and-drop
interfaces to coding?**

All of our scientific logic is also
accessible through VisTrails, a
graphical programming tool
somewhat like LabView.

**Do you hate installing things?
Do you suffer from slow
Internet? Do you prefer to use
NSLS-II compute resources?**

Visit notebook.nsls2.bnl.gov, log
in with your controls account
(soon, we hope, to be provided
when your beam time is
granted), and do your analysis
from any machine, anywhere!
Use our curated examples to get
started.

Jupyter: web-based interactive scientific computing in any language (including Python)



Two Deployments of Jupyter at NSLS-II

- **jupyterhub**: authenticated access (via LDAP) to beamline data with persistent analysis code - *notebook.nsls2.bnl.gov*
- **tmpnb**: public access to demo data and demo notebooks in a personal, temporary “sandbox” that is created fresh and then erased after some time of inactivity - *try.nsls2.bnl.gov*