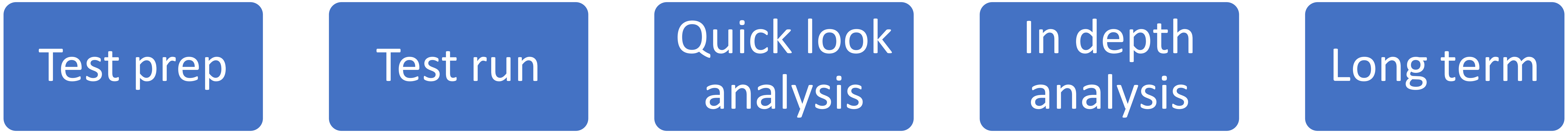
FIREfly
A Prototype Flight Test Data Server

October 2020

The HDF Group

Mike Folk
Interim CEO
The HDF Group

A test data lifecycle



Test prep

Test run

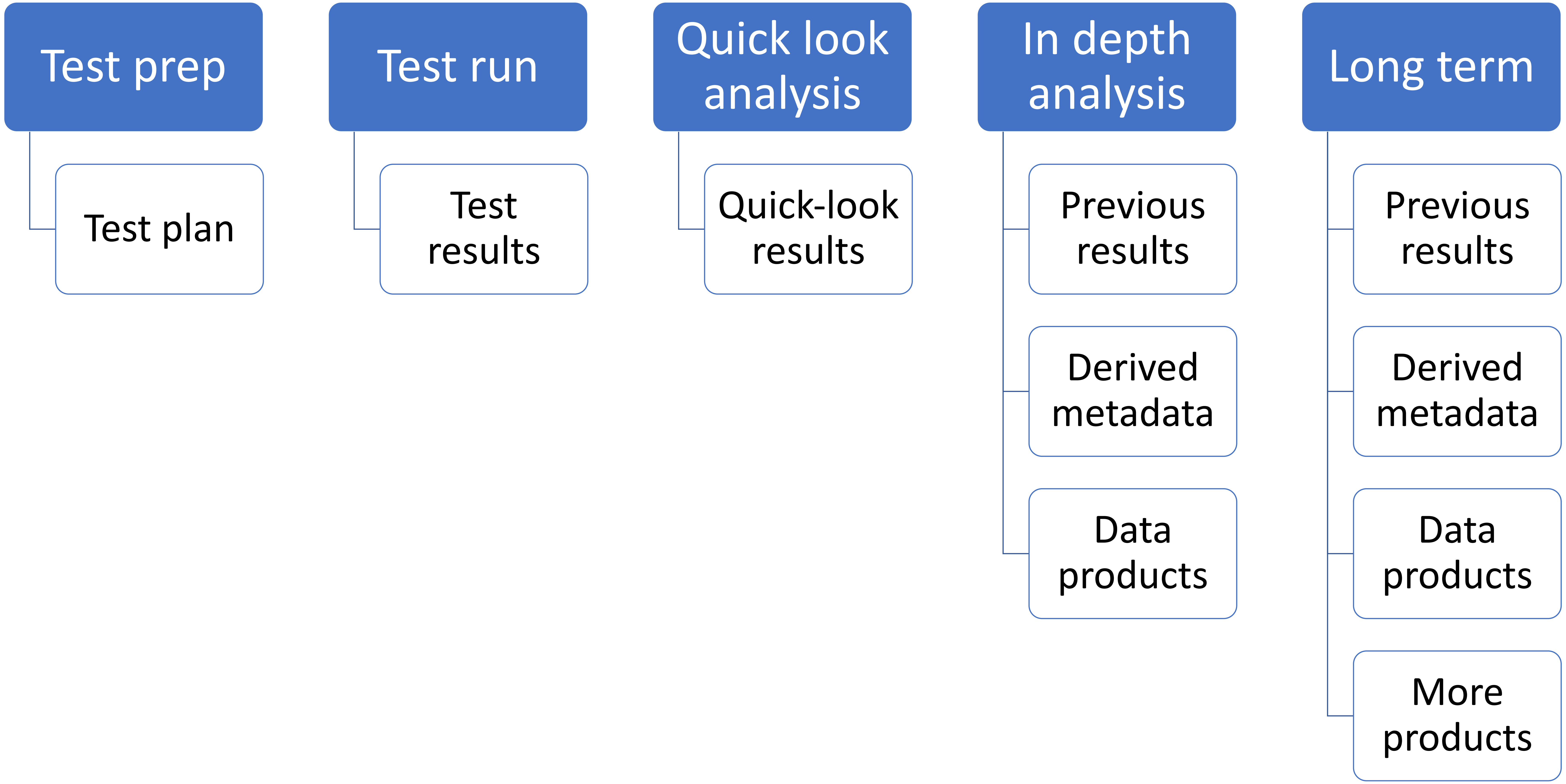
Quick look analysis

In depth analysis

Long term

A test data lifecycle – the data

The HDF Group



Test prep
- Test plan

Test run
- Test results

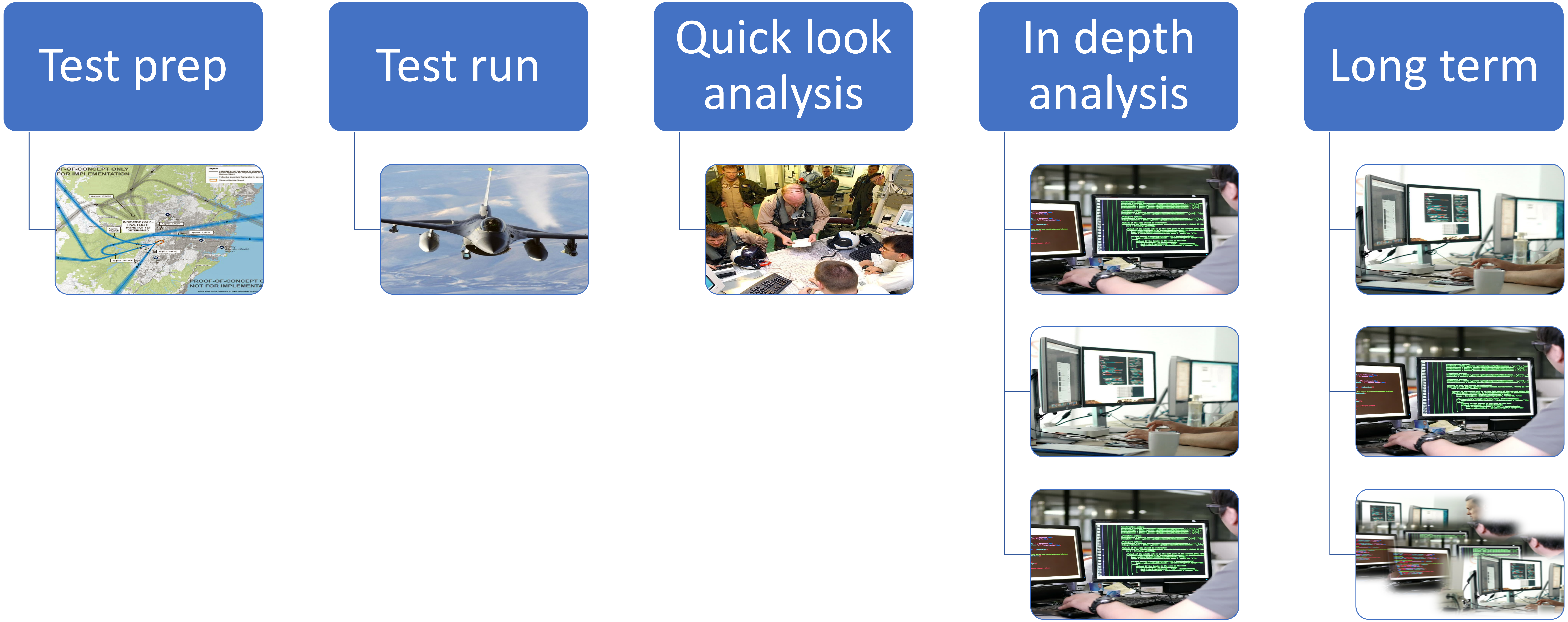
Quick look analysis
- Quick-look results

In depth analysis
- Previous results
- Derived metadata
- Data products

Long term
- Previous results
- Derived metadata
- Data products
- More products

10/14/2020

A test data lifecycle – the people



Test prep

Test run

Quick look analysis

In depth analysis

Long term

Test prep Test run Quick look analysis In depth analysis Archive

Test prep Test run Quick look analysis In depth analysis Archive

Test prep Test run Quick look analysis In depth analysis Archive

Test prep Test run Quick look analysis In depth analysis Archive

This Photo by Unknown Author is licensed under CC BY

www.mapsopensource.com

The HDF Group

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Test prep
Test run
Quick look analysis
In depth analysis
Archive

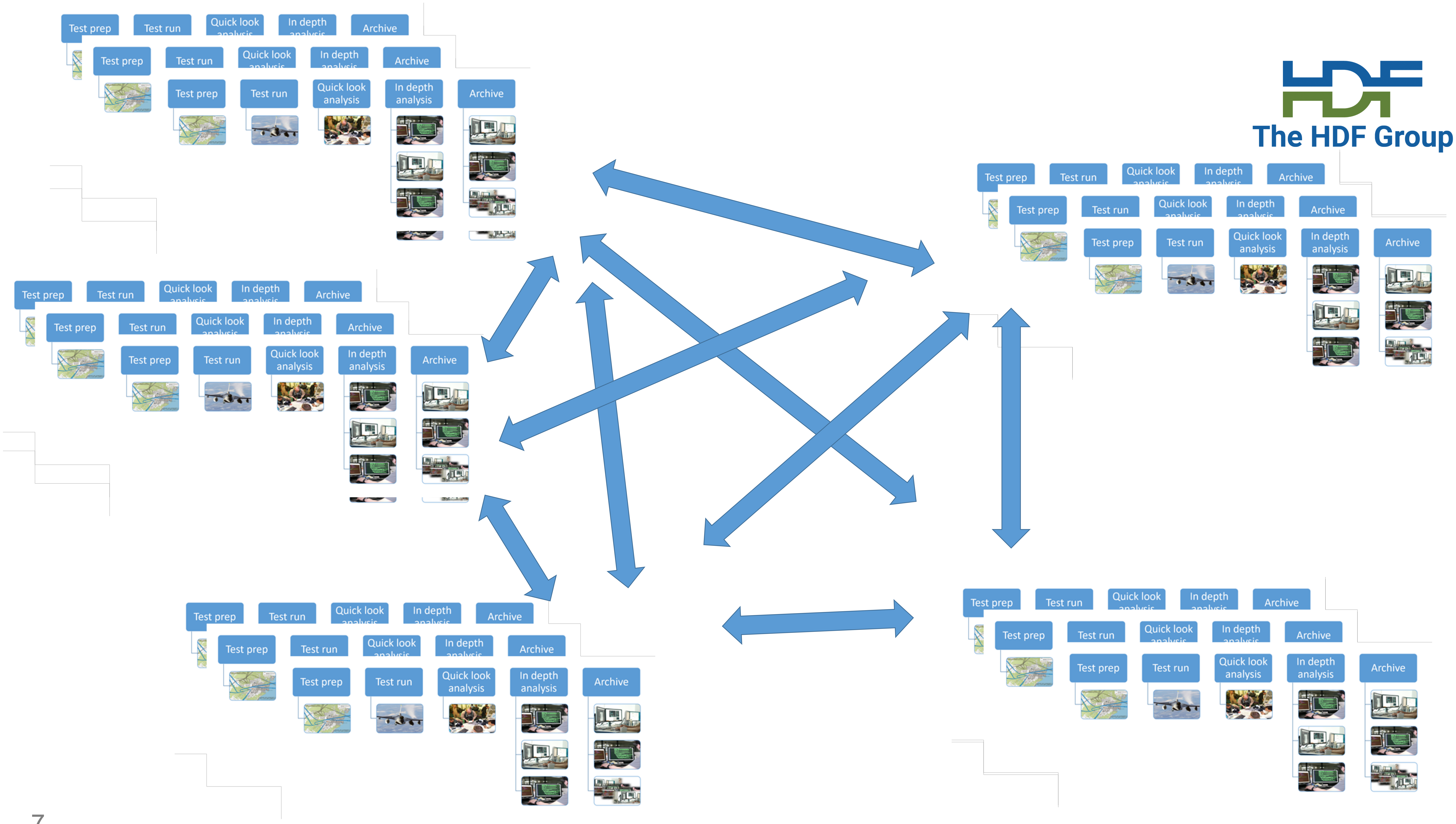
Test prep
Test run
Quick look analysis
In depth analysis
Archive

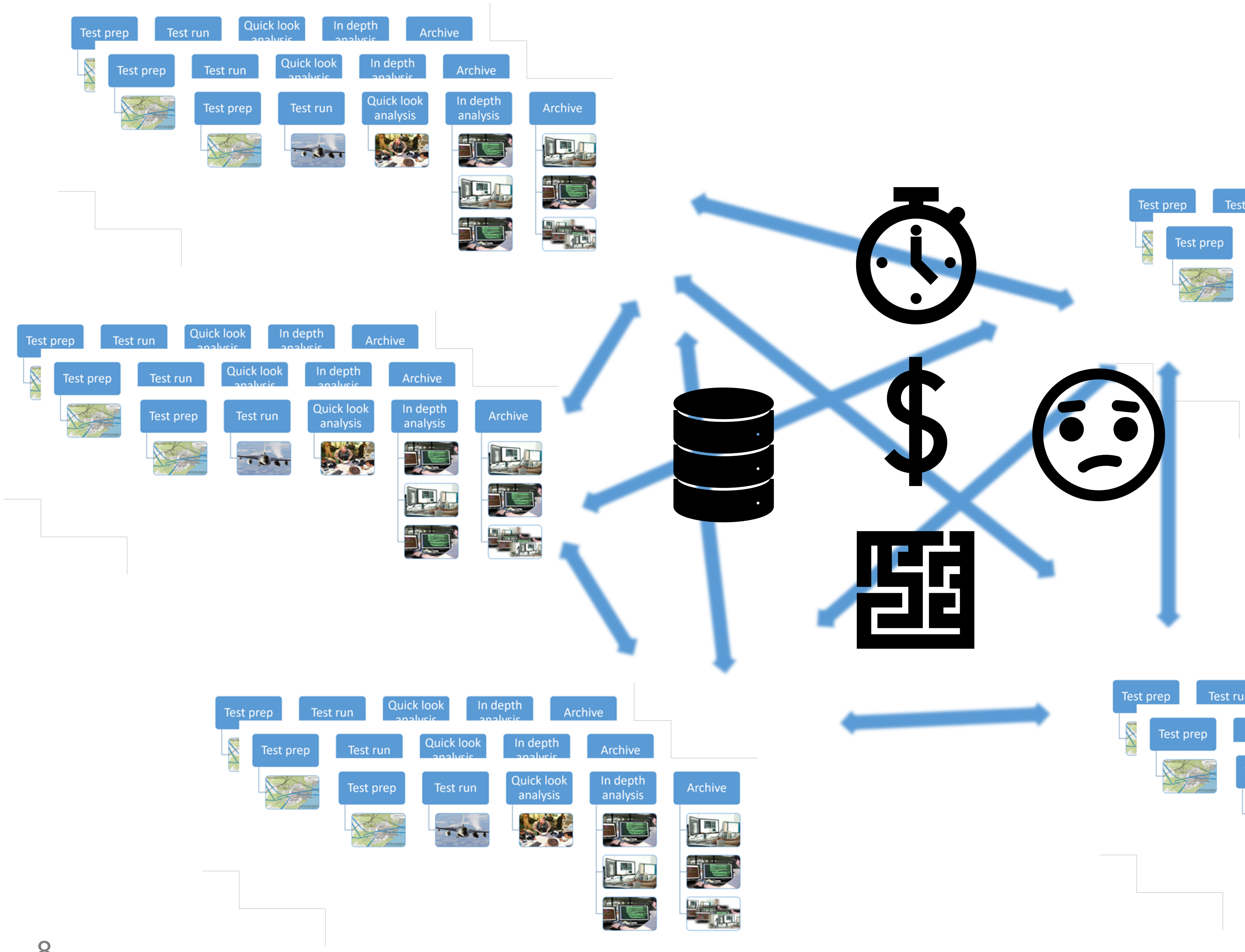
Test prep
Test run
Quick look analysis
In depth analysis
Archive

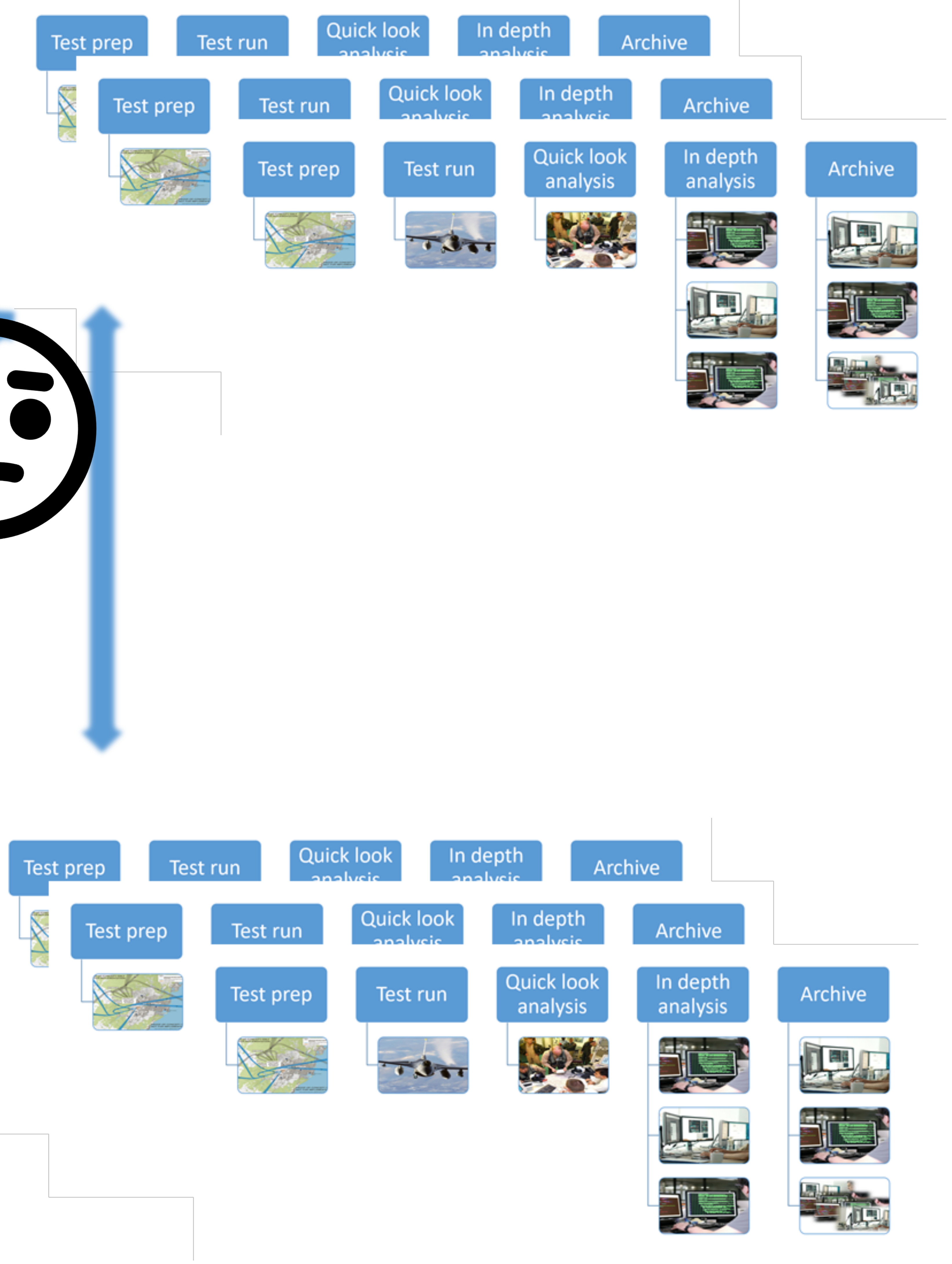
This Photo by Unknown Author is licensed under CC BY

www.mapsopensource.com

The HDF Group




Test prep
Test run
Quick look analysis
In depth analysis
Archive
Test prep
Test run
Quick look analysis
In depth analysis
Archive
Test prep
Test run
Quick look analysis
In depth analysis
Archive
Test prep
Test run
Quick look analysis
In depth analysis
Archive


Test prep
Test run
Quick look analysis
In depth analysis
Archive
Test prep
Test run
Quick look analysis
In depth analysis
Archive

Air Force current state

The HDF Group

…much of this data resides in repositories that

- are isolated from each other
- with limited data search and compute capabilities within them
- and limited data access between them

This in turn

- impedes data discovery
- slows down data analysis turnaround times
- impedes sharing of results
- and constrains larger scale data analyses

Objectives – Develop an…

Open-source web-based data server

That provides a simple, robust capability to

- Ingest datasets from various formats,
- Perform queries to find datasets of interest,
- View data of interest through web user interface,
- Return subsets of the data for client-side analysis,
- Upload Python algorithms for analysis

And be hostable in a Cloud environment

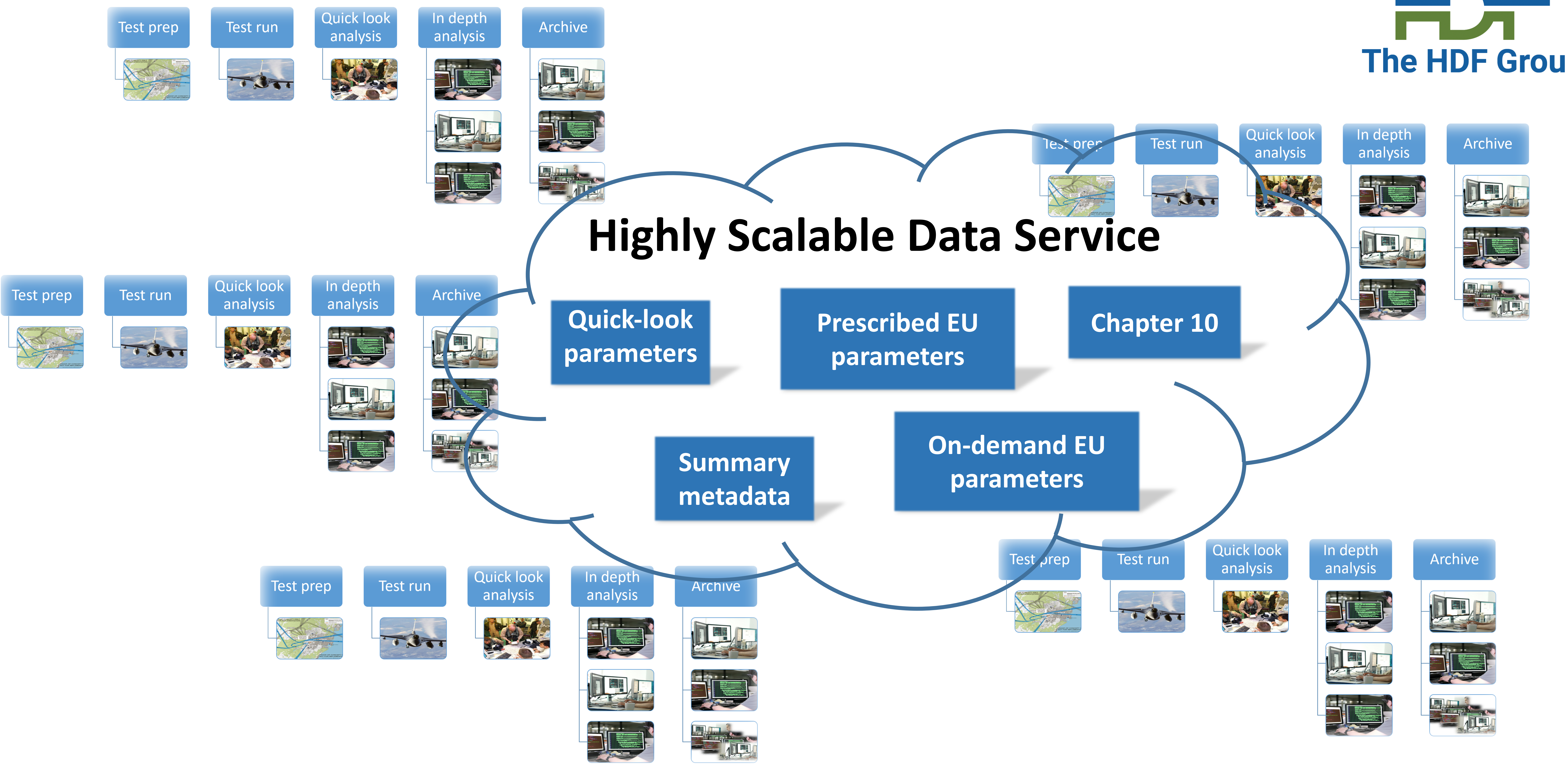
The Air Force asked for

A solution that would address these issues based on HSDS

- Why HSDS?
  - HDF5

The HDF Group

Highly Scalable Data Service



Test prep
Test run
Quick look analysis
In depth analysis
Archive

Quick-look parameters

Prescribed EU parameters

Chapter 10

Summary metadata

On-demand EU parameters

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Test prep
Test run
Quick look analysis
In depth analysis
Archive

Our team

- Gerd Heber, John Readey, Aleksandar Jelenak, Mike Folk (The HDF Group)
- Bob Baggerman (Independent flight test consultant)
- Chris Colbert (SCC Software)
- Jeff Corn (Chief, Test Techniques Development Flight, U.S. Air Force)
- Other Air Force folks

Proposed proof of concept:
Data server with four core functions

INGEST - ingest scientific datasets, especially in Ch10 and HDF5 formats

FILTER - perform SQL-like data queries to find datasets of interest

REDUCE - view datasets of interest using a Web-based user interface, and upload Python algorithms for server-side data analysis

EXTRACT - return subsets of datasets to the client in formats convenient for client-side analysis

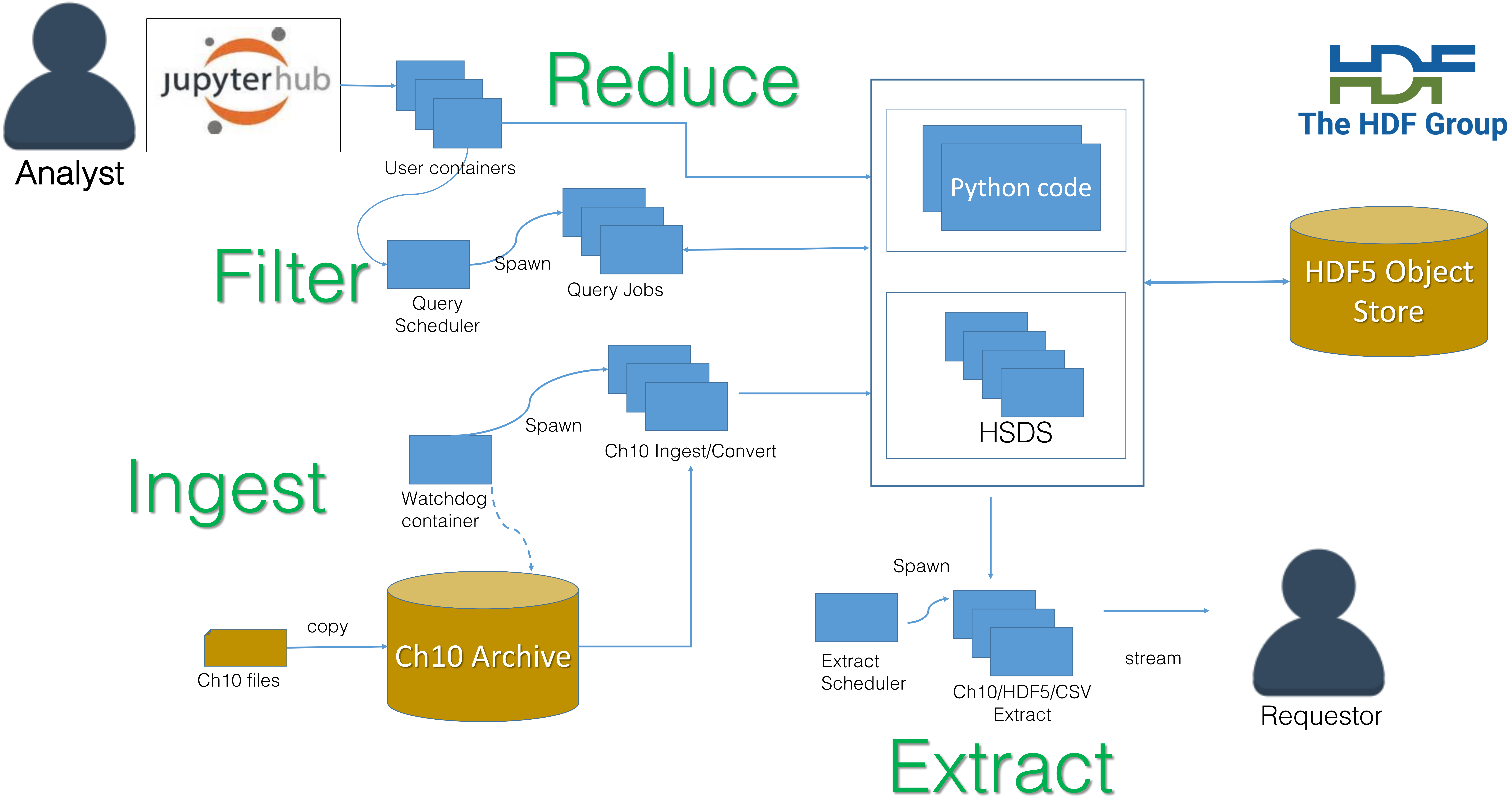
Proposed proof of concept:
Data server with four core functions

FILTER

INGEST

REDUCE

EXTRACT



Analyst

jupyterhub

User containers

Reduce

Python code

HDF5 Object Store

HSDS

Filter

Query Scheduler

Spawn

Query Jobs

Ingest

Watchdog container

Spawn

Ch10 Ingest/Convert

Ch10 files

copy

Ch10 Archive

Spawn

Extract Scheduler

Ch10/HDF5/CSV Extract

stream

Requestor

Extract

The HDF Group

What existing open source components?

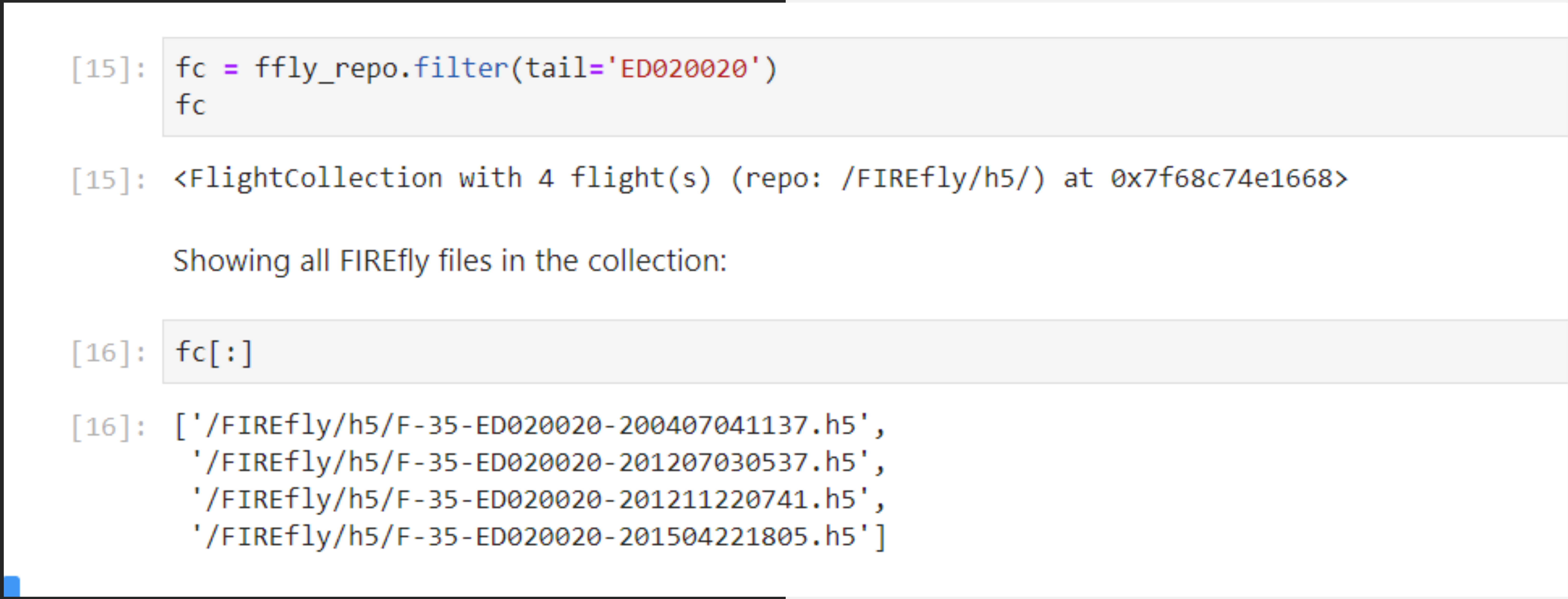
- **HSDS** - share data between widely distributed users
- **JupyterLab -** interactive computing via remote applications
- **Docker containers** - deploy without traditional VM overhead
- **Kubernetes -** deploy/monitor containers across the cluster
- **Py106 -** python package for irig 106 library

Packages developed for this project

- **firefly-py python** package for data query, analysis and visualization
- **Python 106 extensions** for video format 0
- **HSDS query** enhancements
- **FIREfly GUI query** capability
- **Chapter 10 ingest** watchdog
- **Chapter 10-to-HDF5** converter
- **Synthetic Chapter 10** file generator

Some screenshots

Search by tail number returns four flights

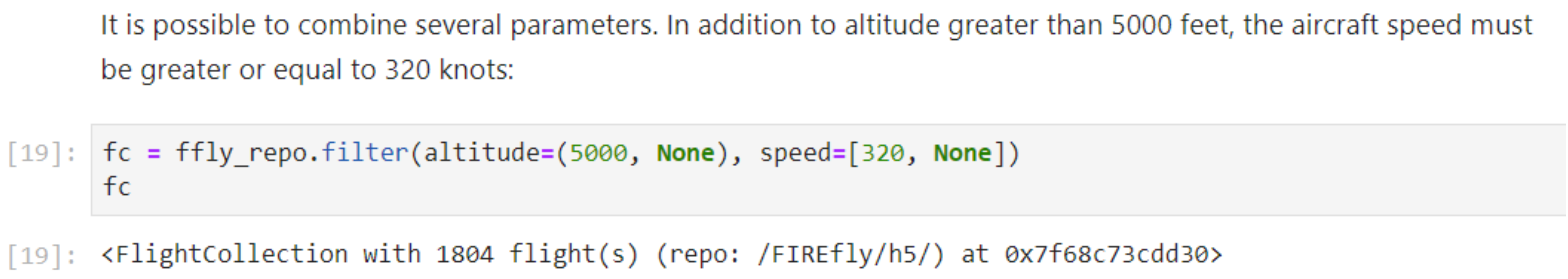


```
[15]: fc = ffly_repo.filter(tail='ED020020')
      fc

[15]: <FlightCollection with 4 flight(s) (repo: /FIREfly/h5/) at 0x7f68c74e1668>
```

Showing all FIREfly files in the collection:

```
[16]: fc[:]

[16]: ['/FIREfly/h5/F-35-ED020020-200407041137.h5',
       '/FIREfly/h5/F-35-ED020020-201207030537.h5',
       '/FIREfly/h5/F-35-ED020020-201211220741.h5',
       '/FIREfly/h5/F-35-ED020020-201504221805.h5']
```

Find flights with altitude > 5000 feet and speed > or equal to 320 knots.



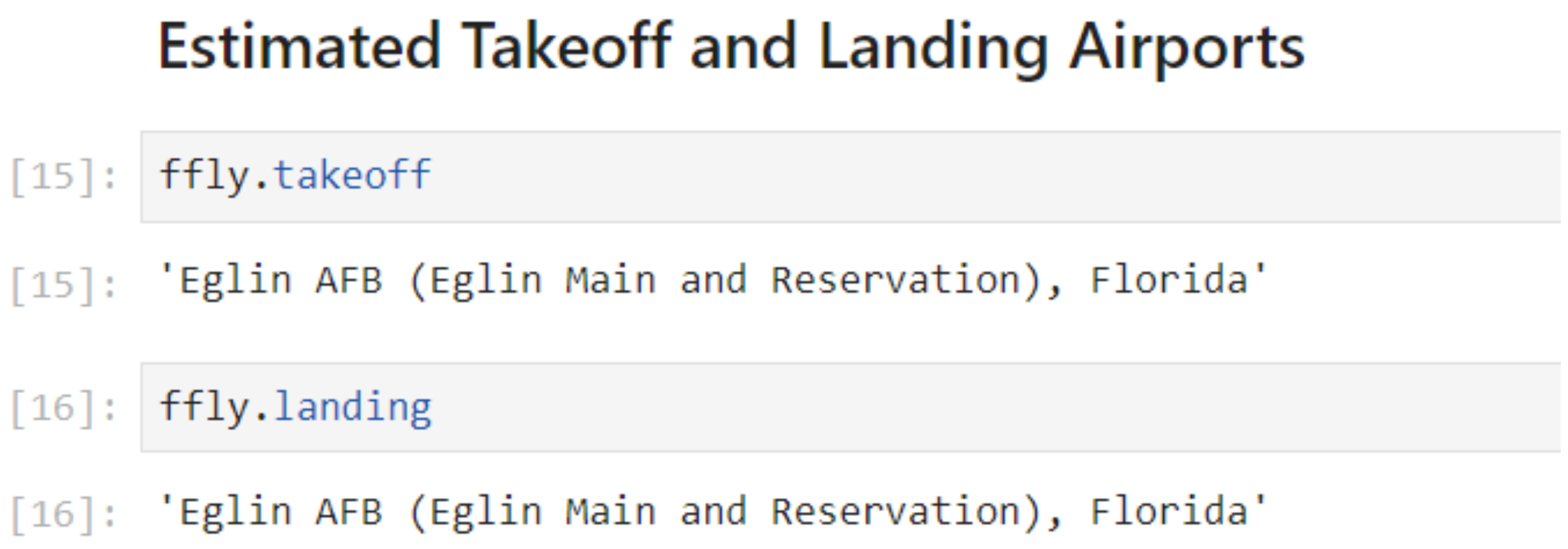
It is possible to combine several parameters. In addition to altitude greater than 5000 feet, the aircraft speed must be greater or equal to 320 knots:

```
[19]: fc = ffly_repo.filter(altitude=(5000, None), speed=[320, None])
      fc
```

```
[19]: <FlightCollection with 1804 flight(s) (repo: /FIREfly/h5/) at 0x7f68c73cdd30>
```

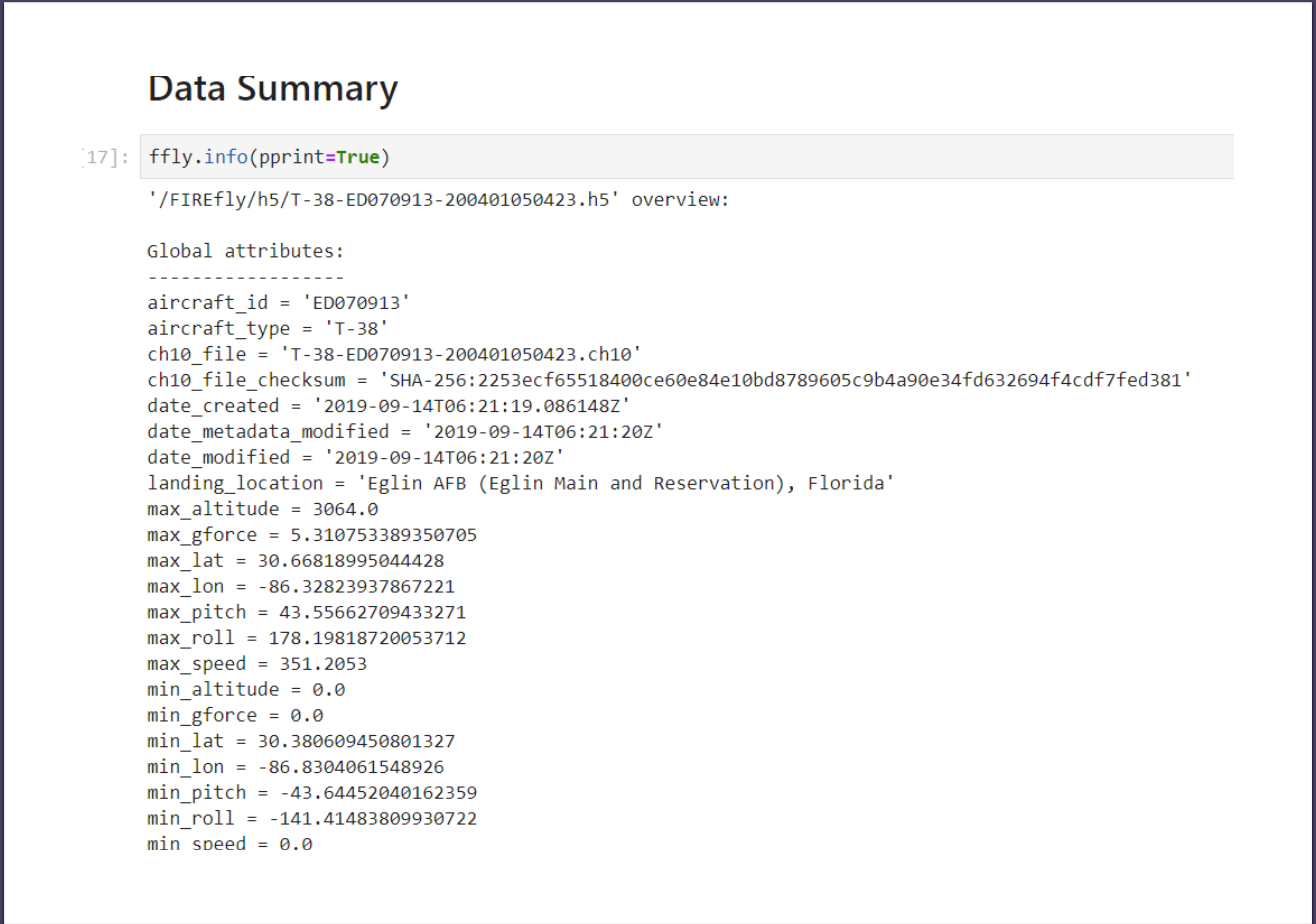
For a given flight find takeoff and landing airports computed from metadata

The HDF Group



Estimated Takeoff and Landing Airports

```
[15]: ffly.takeoff

[15]: 'Eglin AFB (Eglin Main and Reservation), Florida'

[16]: ffly.landing

[16]: 'Eglin AFB (Eglin Main and Reservation), Florida'
```



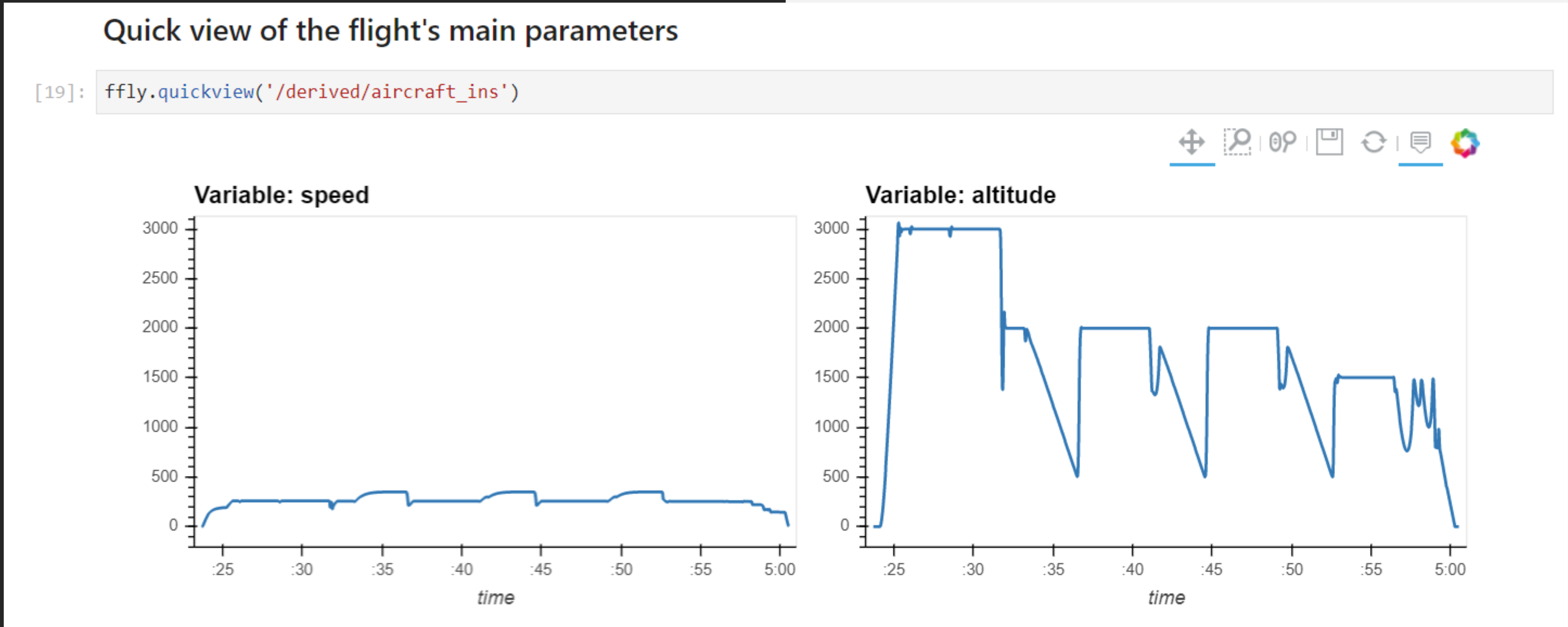
Data Summary

```
[17]: ffly.info(pprint=True)
```

```
'/FIREfly/h5/T-38-ED070913-200401050423.h5' overview:

Global attributes:
------------------
aircraft_id = 'ED070913'
aircraft_type = 'T-38'
ch10_file = 'T-38-ED070913-200401050423.ch10'
ch10_file_checksum = 'SHA-256:2253ecf65518400ce60e84e10bd8789605c9b4a90e34fd632694f4cdf7fed381'
date_created = '2019-09-14T06:21:19.086148Z'
date_metadata_modified = '2019-09-14T06:21:20Z'
date_modified = '2019-09-14T06:21:20Z'
landing_location = 'Eglin AFB (Eglin Main and Reservation), Florida'
max_altitude = 3064.0
max_gforce = 5.310753389350705
max_lat = 30.66818995044428
max_lon = -86.32823937867221
max_pitch = 43.55662709433271
max_roll = 178.19818720053712
max_speed = 351.2053
min_altitude = 0.0
min_gforce = 0.0
min_lat = 30.380609450801327
min_lon = -86.8304061548926
min_pitch = -43.64452040162359
min_roll = -141.41483809930722
min_speed = 0.0
```

For a given flight write out global metadata from HDF5 file

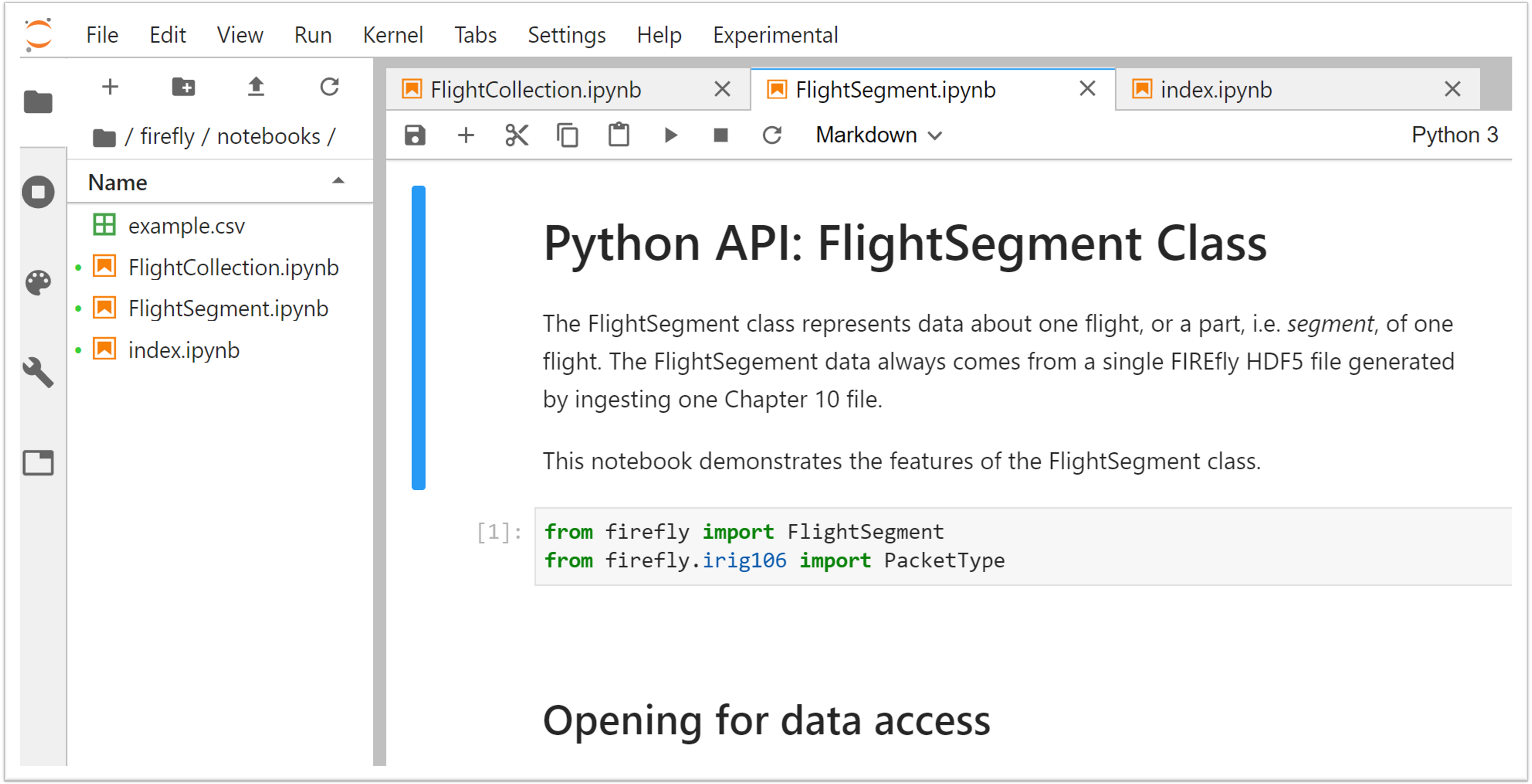
Display “quickview” graphs of flight's main parameters



Quick view of the flight's main parameters

```
[19]: ffly.quickview('/derived/aircraft_ins')
```

Test drive at KitaLab: https://www.hdfgroup.org/hdfkitalab/
Login to KitaLab and find the directory /firefly



Python API: FlightSegment Class

The FlightSegment class represents data about one flight, or a part, i.e. *segment*, of one flight. The FlightSegement data always comes from a single FIREfly HDF5 file generated by ingesting one Chapter 10 file.

This notebook demonstrates the features of the FlightSegment class.

```
[1]: from firefly import FlightSegment
     from firefly.irig106 import PacketType
```

Opening for data access

Next steps

Next steps

Us

- Take FIREfly to the next stage
- Apply to other domains

You

- Talk to us if interested
- Have a look
  - KitaLab: https://www.hdfgroup.org/hdfkitalab/
  - Code: https://github.com/AkadioInc/firefly

Thank you

Questions and comments

Acknowledgement/disclaimer

The HDF Group

Acknowledgment: This material is based upon work supported by the United States Air Force under Contract No. FA9302-19-P-1021.

Disclaimer: Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Air Force.