



DataPipe Intro

Project Management and Systems Engineering

Workshop Goals

Discuss requirements and design

Identify path from current prototypes to final product

- ▶ Identify and fix design issues
- ▶ Identify missing pieces
- ▶ Identify who does what

Introduce and indoctrinate new developers → Vision Sharing!



Overview

Project Management

Context

Use Cases

Releases

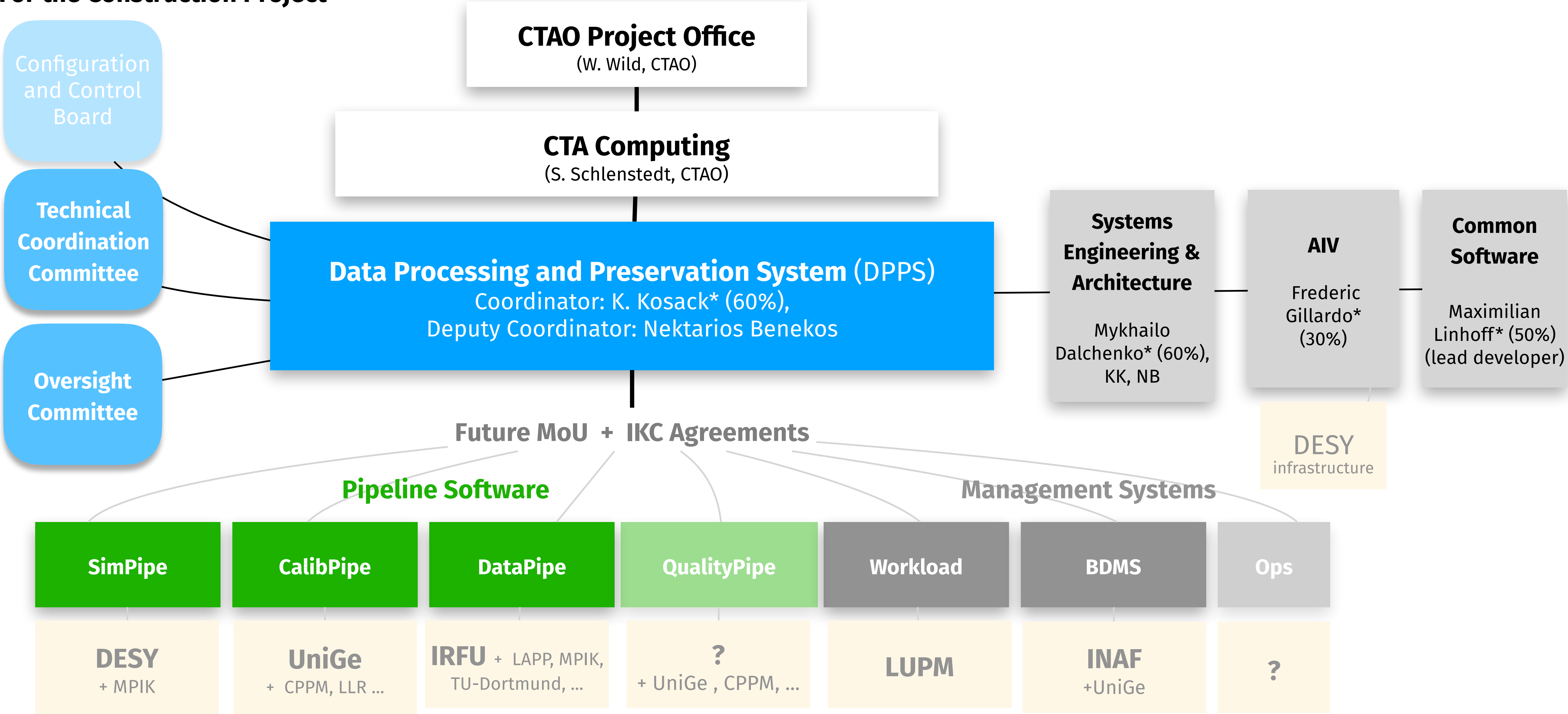


Project Management



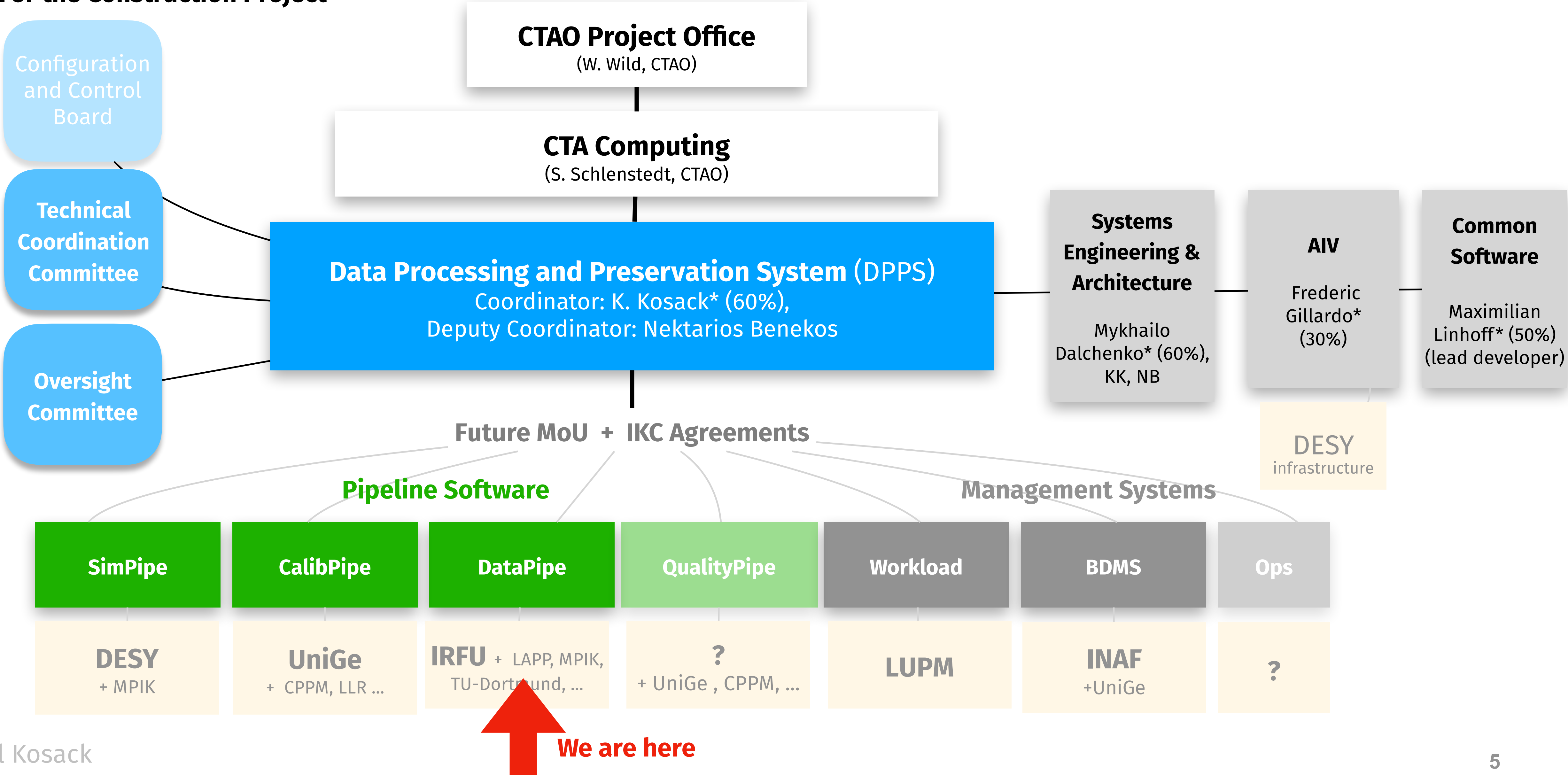
DPPS Management Structure

For the Construction Project



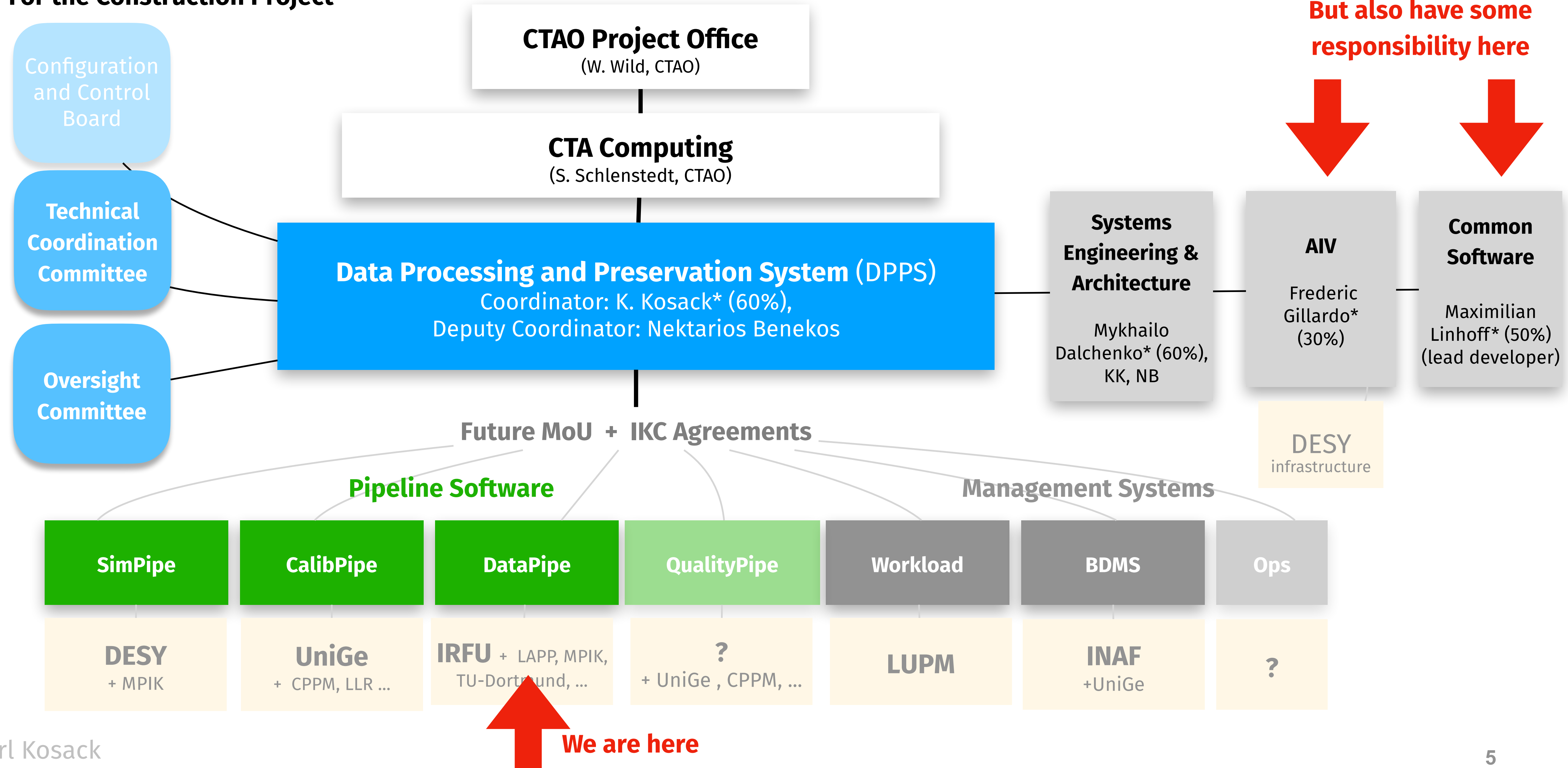
DPPS Management Structure

For the Construction Project



DPPS Management Structure

For the Construction Project



Who are we? DataPipe IKC and Partners



Camera Teams:

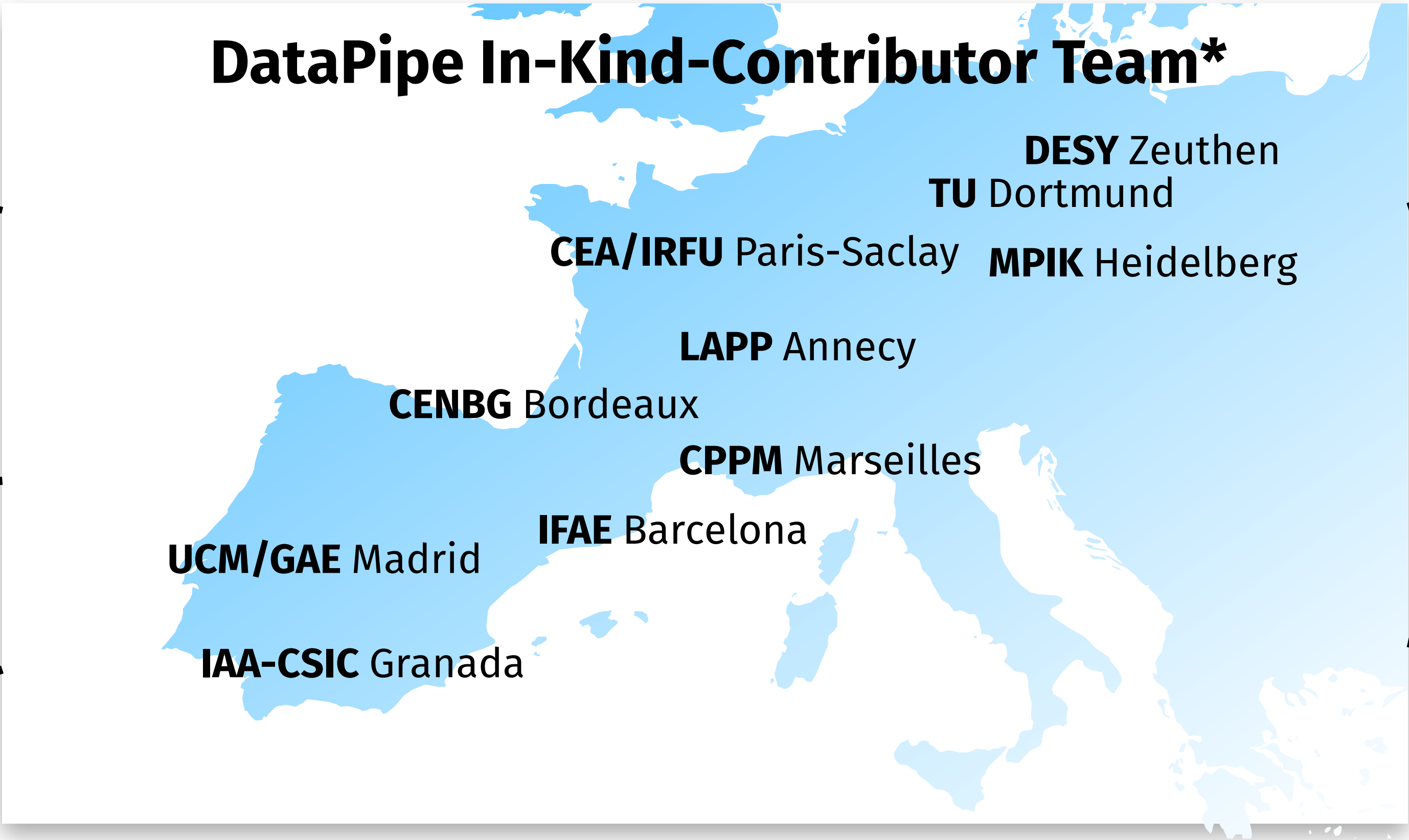
NectarCam
FlashCam
LSTCam
SSTCam
SCTCam

SimPipe

DESY Zeuthen (Lead)

CalibPipe

University of Geneva (Lead)



Workload

LUPM Montpellier

Common-Software

*Declared in CTAO Money Matrix

Communication: too much of a good thing?



Communication: too much of a good thing?



	Chat	Forum	Code	Docs	Issues	Milestones	Wiki	Calendar
--	------	-------	------	------	--------	------------	------	----------

Communication: too much of a good thing?



	Chat	Forum	Code	Docs	Issues	Milestones	Wiki	Calendar
GitHub		x	x		x	x	x	

Communication: too much of a good thing?



	Chat	Forum	Code	Docs	Issues	Milestones	Wiki	Calendar
GitHub		x	x		x	x	x	
GitLab (CTAO)			x		x	x	x	

Communication: too much of a good thing?



	Chat	Forum	Code	Docs	Issues	Milestones	Wiki	Calendar
GitHub		x	x		x	x	x	
GitLab (CTAO)			x		x	x	x	
Slack: computing	x							

Communication: too much of a good thing?



	Chat	Forum	Code	Docs	Issues	Milestones	Wiki	Calendar
GitHub		x	x		x	x	x	
GitLab (CTAO)			x		x	x	x	
Slack: computing	x							
Slack: ctapipe	x							

Communication: too much of a good thing?



	Chat	Forum	Code	Docs	Issues	Milestones	Wiki	Calendar
GitHub		x	x		x	x	x	
GitLab (CTAO)			x		x	x	x	
Slack: computing	x							
Slack: ctapipe	x							
Slack: ASWG	x							

Communication: too much of a good thing?



	Chat	Forum	Code	Docs	Issues	Milestones	Wiki	Calendar
GitHub		x	x		x	x	x	
GitLab (CTAO)			x		x	x	x	
Slack: computing	x							
Slack: ctapipe	x							
Slack: ASWG	x							
Redmine		x		x	x	x	x	x

Communication: too much of a good thing?



	Chat	Forum	Code	Docs	Issues	Milestones	Wiki	Calendar
GitHub		x	x		x	x	x	
GitLab (CTAO)			x		x	x	x	
Slack: computing	x							
Slack: ctapipe	x							
Slack: ASWG	x							
Redmine		x		x	x	x	x	x
mailing list		x						x

Communication: too much of a good thing?



	Chat	Forum	Code	Docs	Issues	Milestones	Wiki	Calendar
GitHub		x	x		x	x	x	
GitLab (CTAO)			x		x	x	x	
Slack: computing	x							
Slack: ctapipe	x							
Slack: ASWG	x							
Redmine		x		x	x	x	x	x
mailing list		x						x
CTAO Office365		x		x			x	x

Communication: too much of a good thing?



	Chat	Forum	Code	Docs	Issues	Milestones	Wiki	Calendar
GitHub		X	X		X	X	X	
GitLab (CTAO)			X		X	X	X	
Slack: computing	X							
Slack: ctapipe	X							
Slack: ASWG	X							
Redmine		X		X	X	X	X	X
mailing list		X						X
CTAO Office365		X		X			X	X
Indico								X

Communication: too much of a good thing?



	Chat	Forum	Code	Docs	Issues	Milestones	Wiki	Calendar
GitHub		x	x		x	x	x	
GitLab (CTAO)			x		x	x	x	
Slack: computing	x							
Slack: ctapipe	x							
Slack: ASWG	x							
Redmine		x		x	x	x	x	x
mailing list		x						x
CTAO Office365		x		x			x	x
Indico								x
Teams	x							x

Communication: too much of a good thing?



	Chat	Forum	Code	Docs	Issues	Milestones	Wiki	Calendar
GitHub		X	👍		👍	👍	X	
GitLab (CTAO)			👍		👍	👍	👍	
Slack: computing	👍							
Slack: ctapipe	X							
Slack: ASWG	X							
Redmine		X	X	👍	X	👍	X	X
mailing list		X						X
CTAO Office365		X		X			X	X
Indico								👍
Teams	X							👍

Could use a Project Librarian!

Do you have access?



GitHub: member of cta-dev group?

GitLab: member of computing/dpps/datapipe group?

CTAO Computing Slack: invited? subscribed to #dpps-datapipe?

***computing-dpps-datapipe* Mailinglist**: subscribed?

Redmine: can see the *dpps-coordination* and *datapipe* groups?

DataPipe Deliverables (for each major release)



So far, we have not made a DataPipe release, only individual packages

Software

- ▶ selected packages in a conda env (+ optionally in a Docker container)

Code Documentation:

- ▶ Users guides (installation + high-level usage)
- ▶ Developer Docs (API, etc)

Project Documentation:

- ▶ Use Cases / Functionality
- ▶ Level-C Requirements Specification derived from use cases
- ▶ Requirements Traceability matrices (from UC → LevelC, from Level-C → Level-B)
- ▶ Requirements Verification Plan (how to check each requirement, includes tests/benchmarks)

Release Documentation:

- ▶ Requirements Verification Checklist (ensure we did all the checks)

Could use a Release Manager!

≈week or so of work for every release...

Question for later:

How do we make a DataPipe release?

DPPS will have a Preliminary Design Review (PDR) within a year

- ▶ External + Internal reviewers
- ▶ Will look at all documentation (particularly Use Cases and Requirements) and any existing design and implementation
- ▶ Will give official "recommendations" on what to improve or clarify (RIXes)

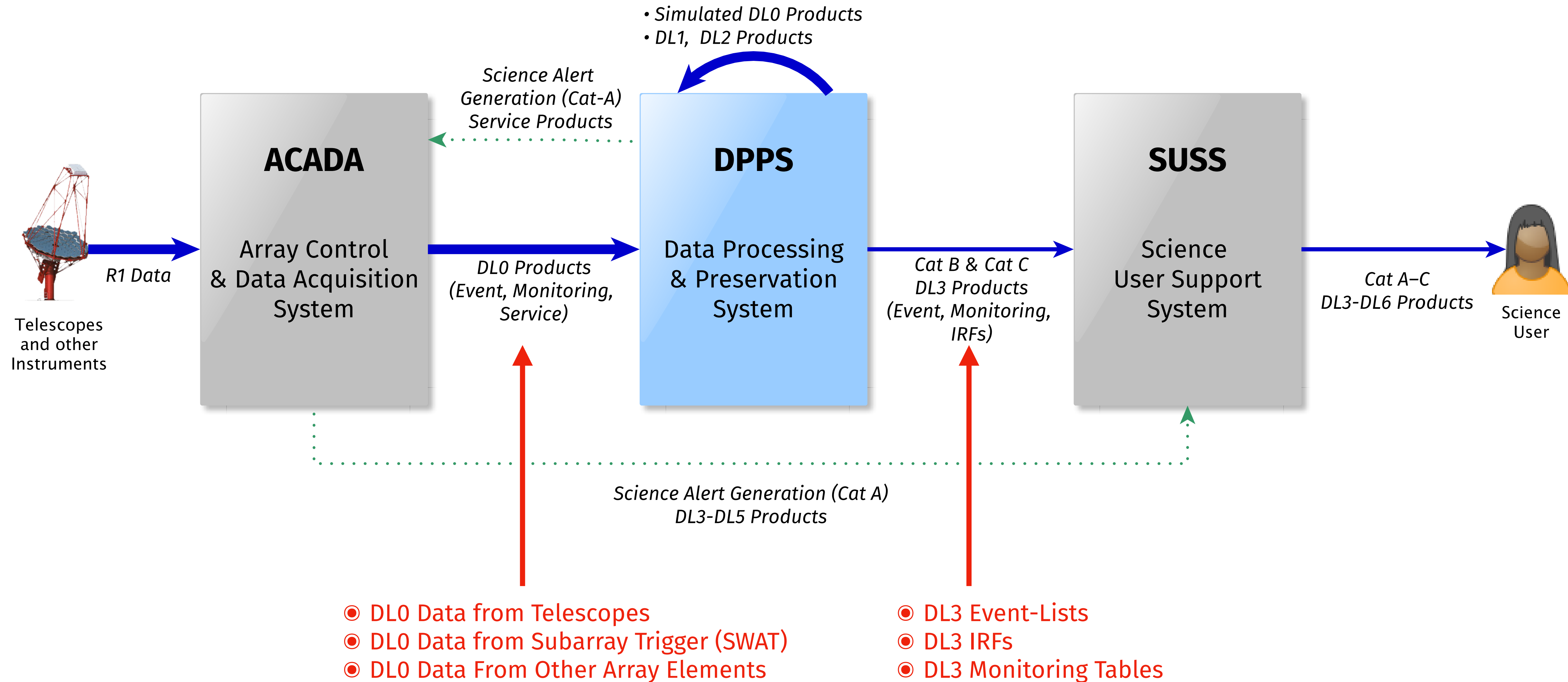
We need to be prepared

Context



DPPS in the CTAO Context

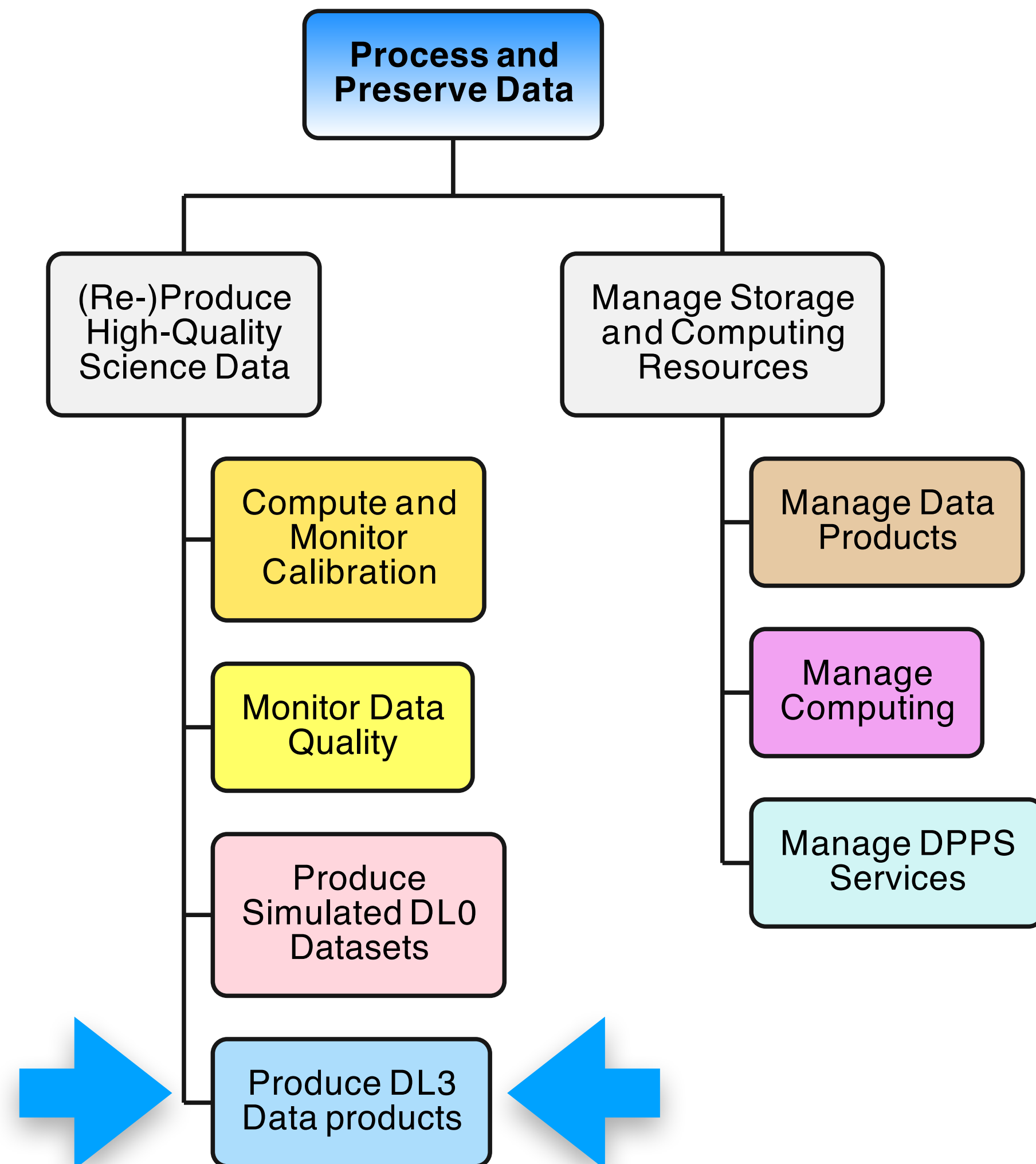
Preserve Data at 2+ off-site
data centers
Process it in-place to DL3



DataPipe in the DPPS Context



DPPS Requirements Specification: <https://redmine.cta-observatory.org/dmsf/files/17752/view>

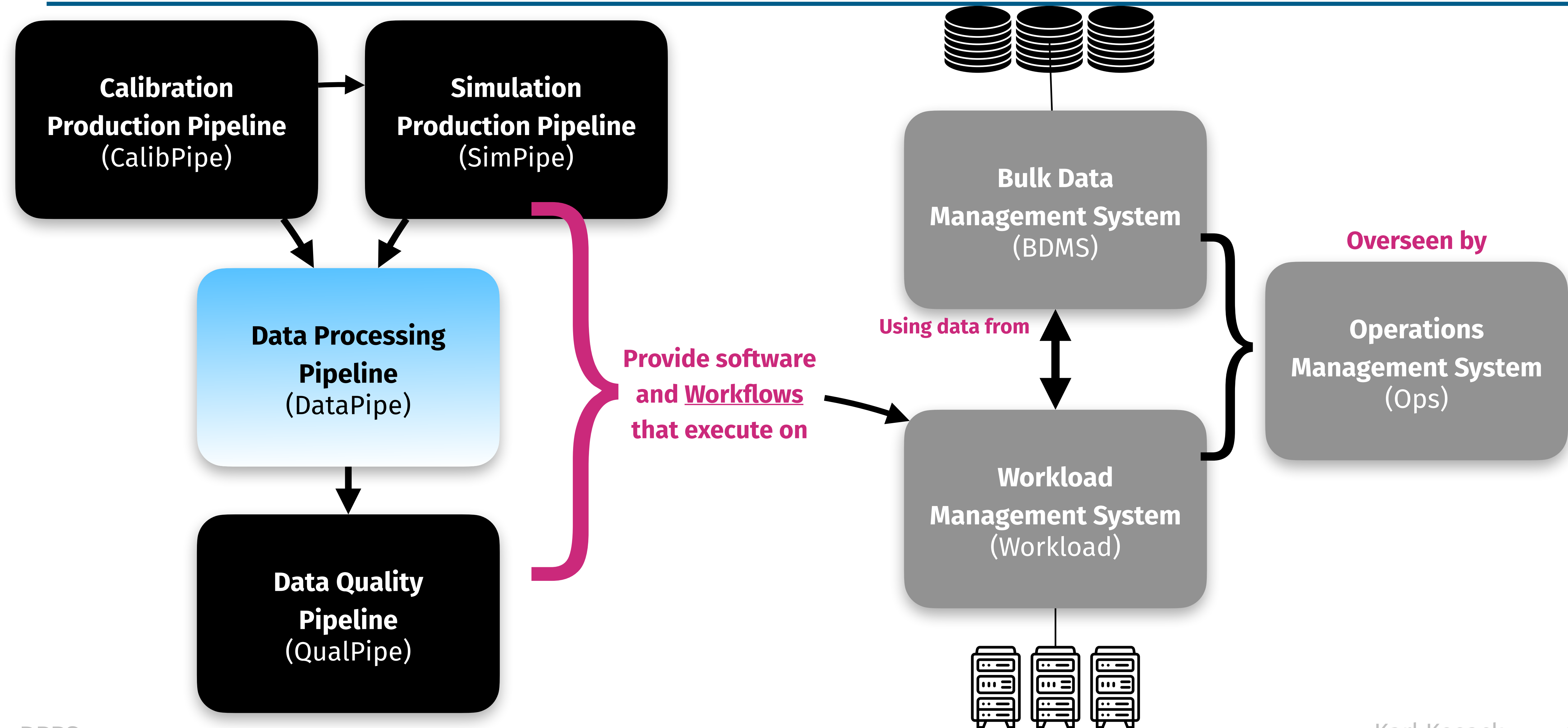


Generated on Mon Oct 16 10:51:38 CEST 2023

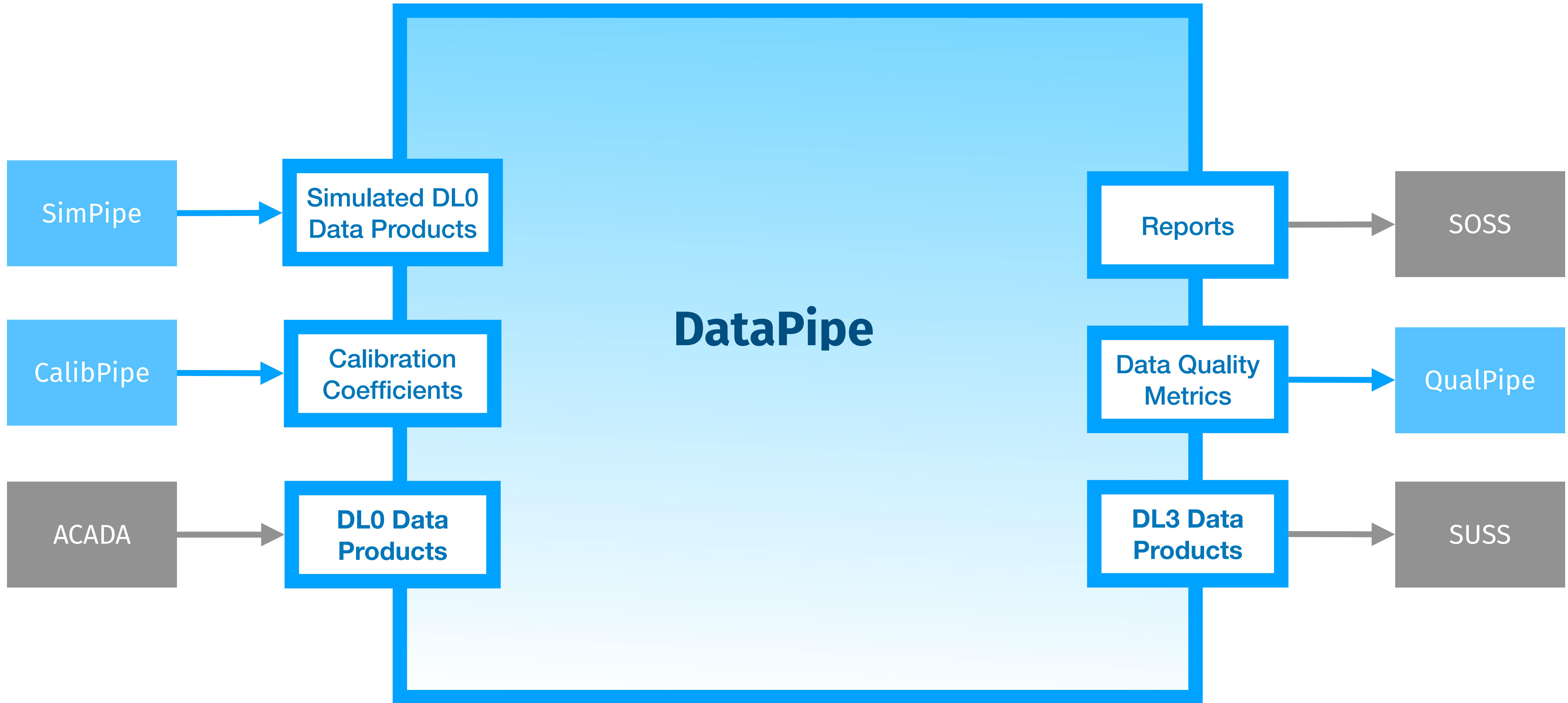
Provide Science-Ready (DL3) data products for each observation block to SUSS for eventual dissemination to users.

- ▶ ...
- ▶ Generate DL3 event, monitoring, and IRF data products from a combination of observed and simulated DL0 data
- ▶ Reproduce after upgrades to calibration or reconstruction algorithms
- ▶ ...

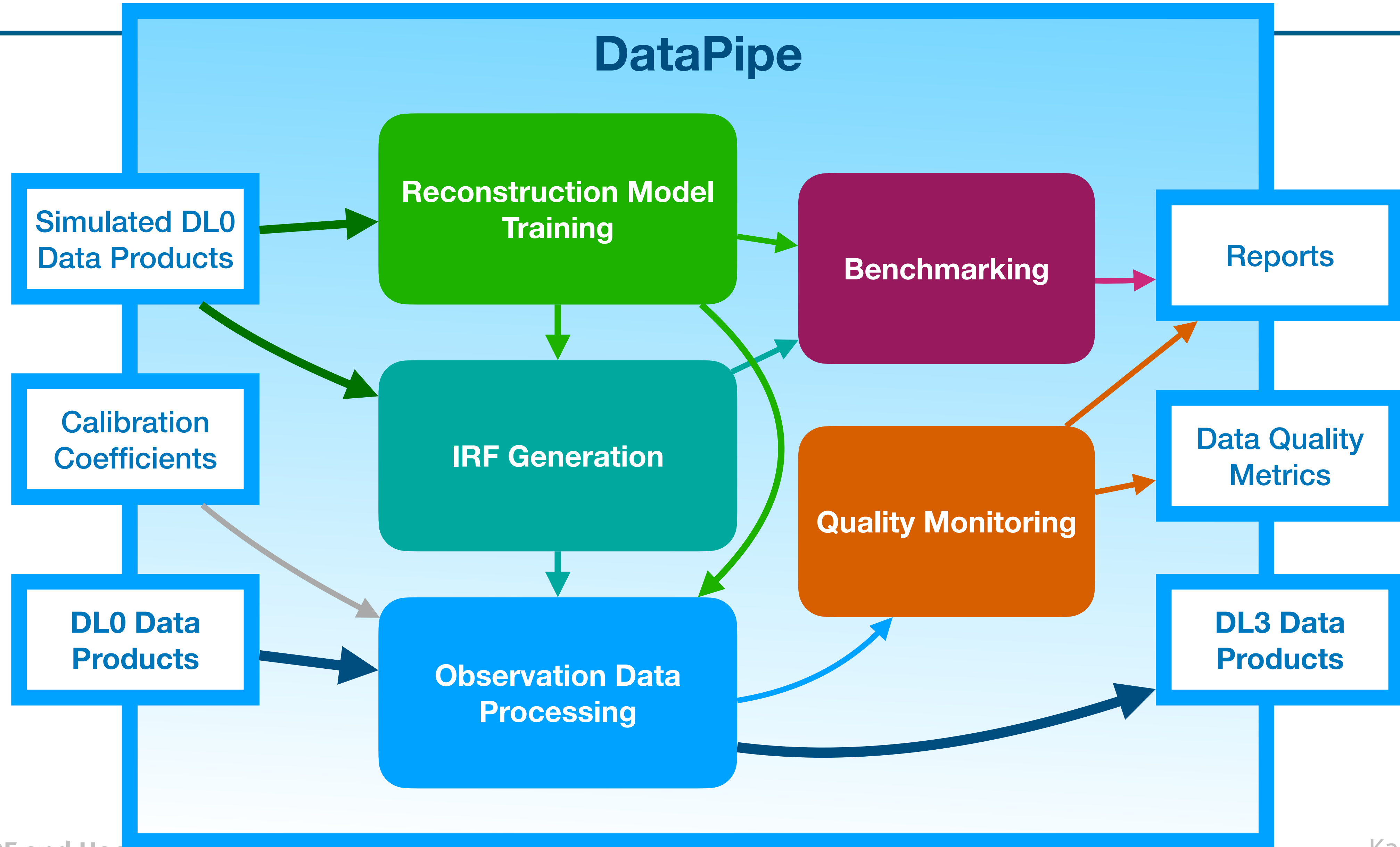
DPPS Subsystems: Pipelines and Management



Workflow View (Black Box)



Workflow View: (White Box)



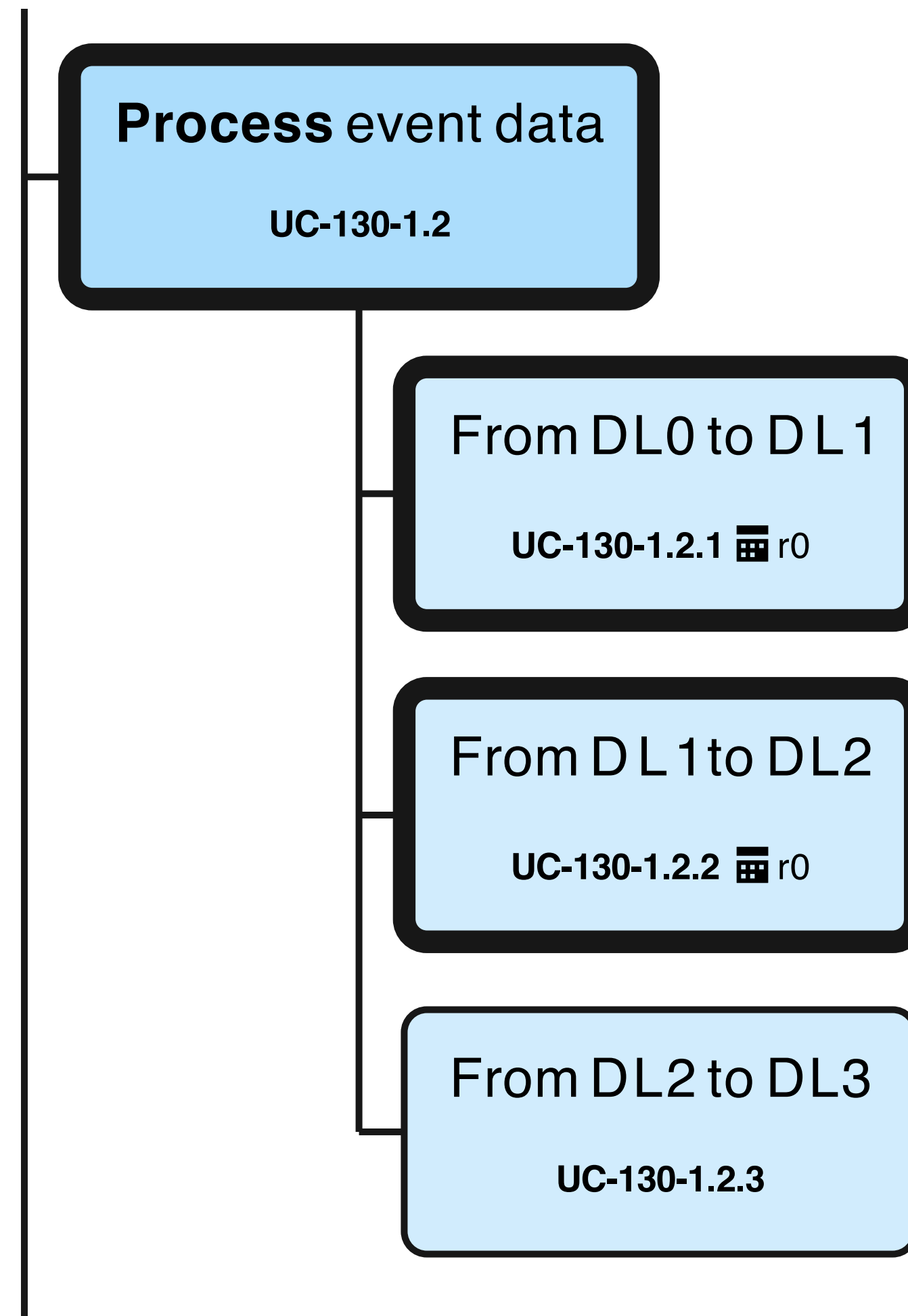
DataPipe Use Cases



Use Cases / Functionality



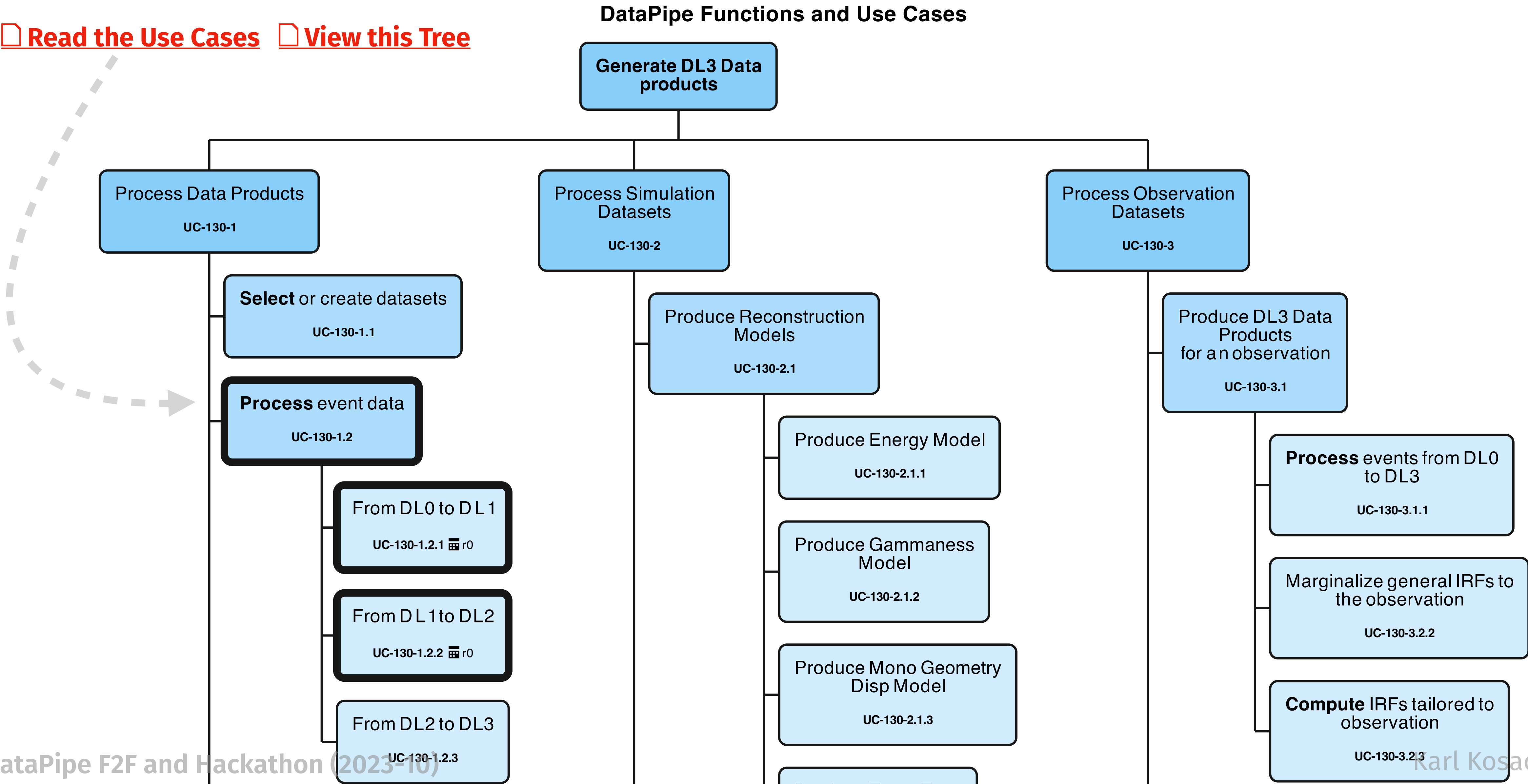
[Read the Use Cases](#) [View this Tree](#)



Use Cases / Functionality



[Read the Use Cases](#) [View this Tree](#)

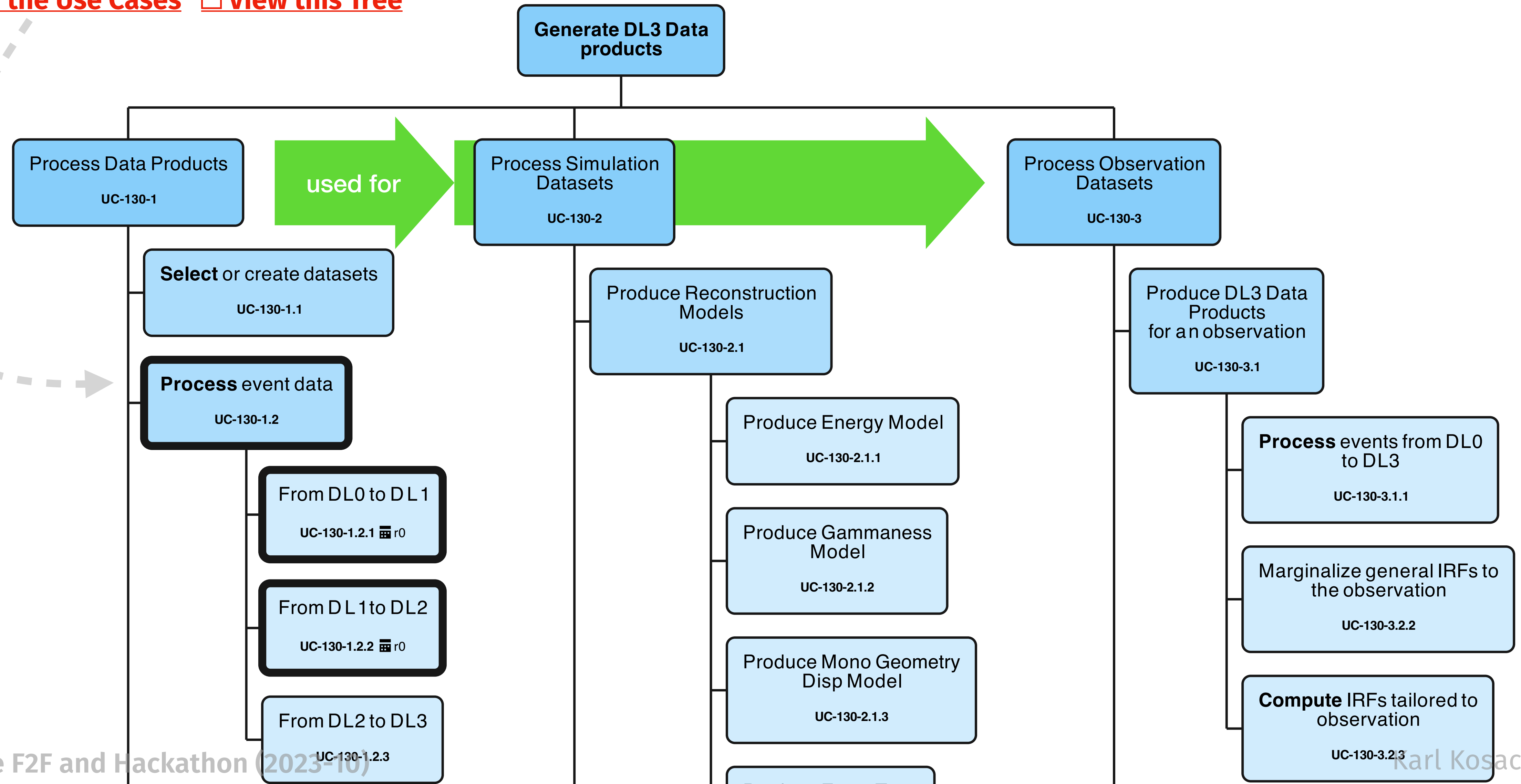


Use Cases / Functionality

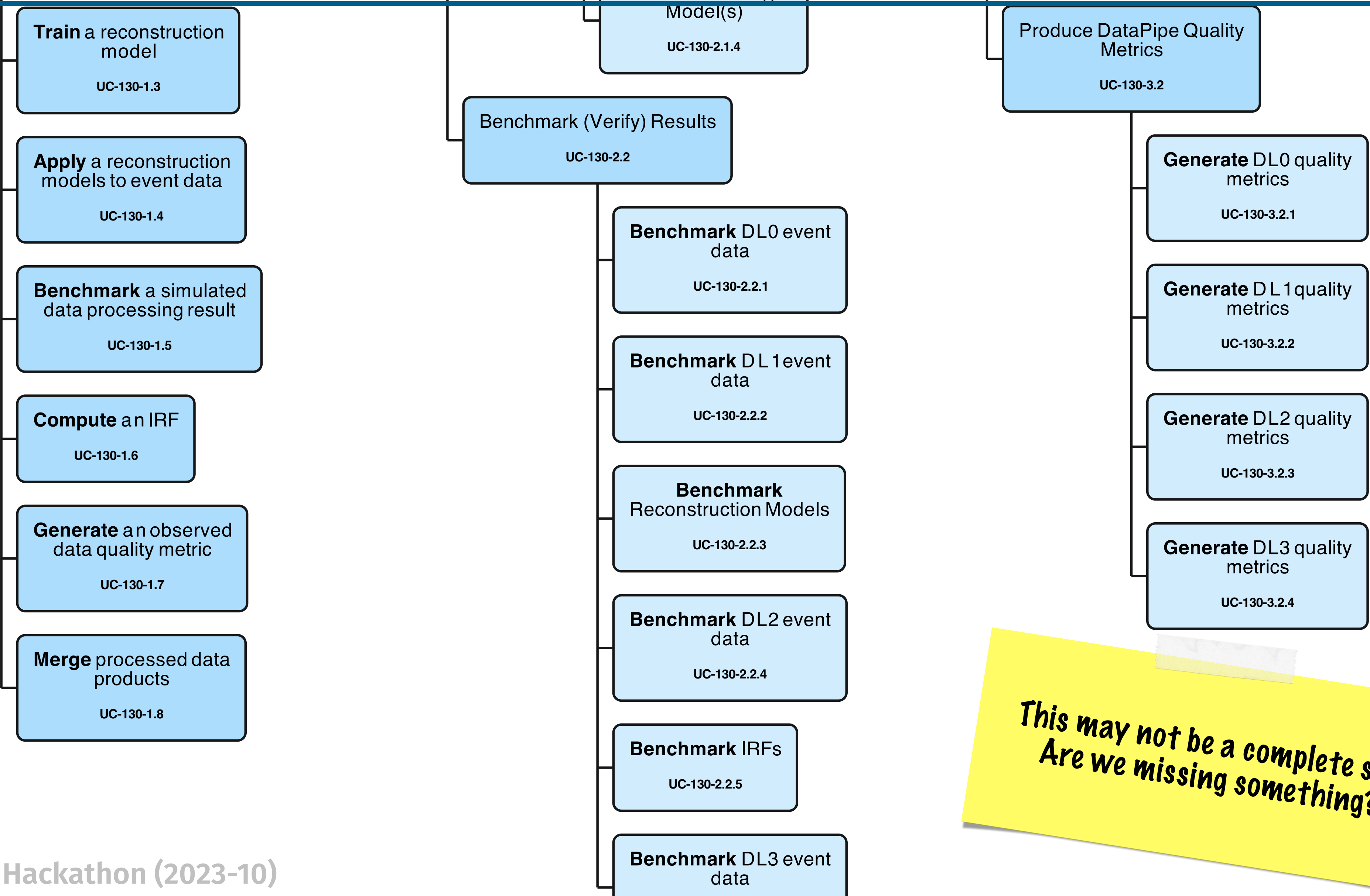


[Read the Use Cases](#) [View this Tree](#)

DataPipe Functions and Use Cases



Use Cases / Functionality



*This may not be a complete set!
Are we missing something?*

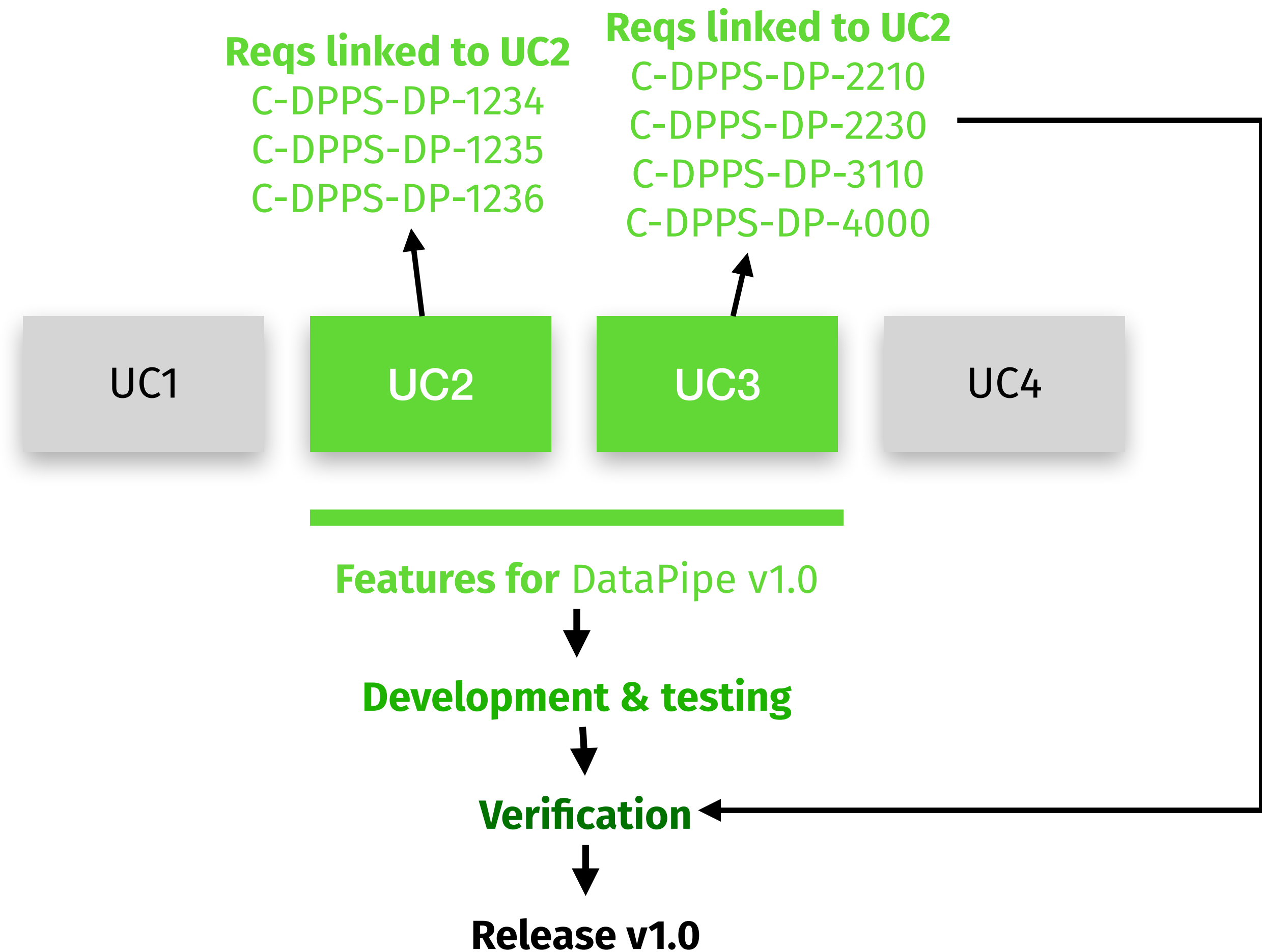
DataPipe Releases

(connected to DPPS releases)



For each release:

- ▶ Identify which use cases to implement and test
- ▶ Update Use Case descriptions if unclear
- ▶ Implement code
- ▶ Verification



Requirements Specification Consists of:

- ▶ Product Definition → see [current draft](#)
- ▶ Requirements list
- ▶ Coverage matrix
- ▶ Verification Statements

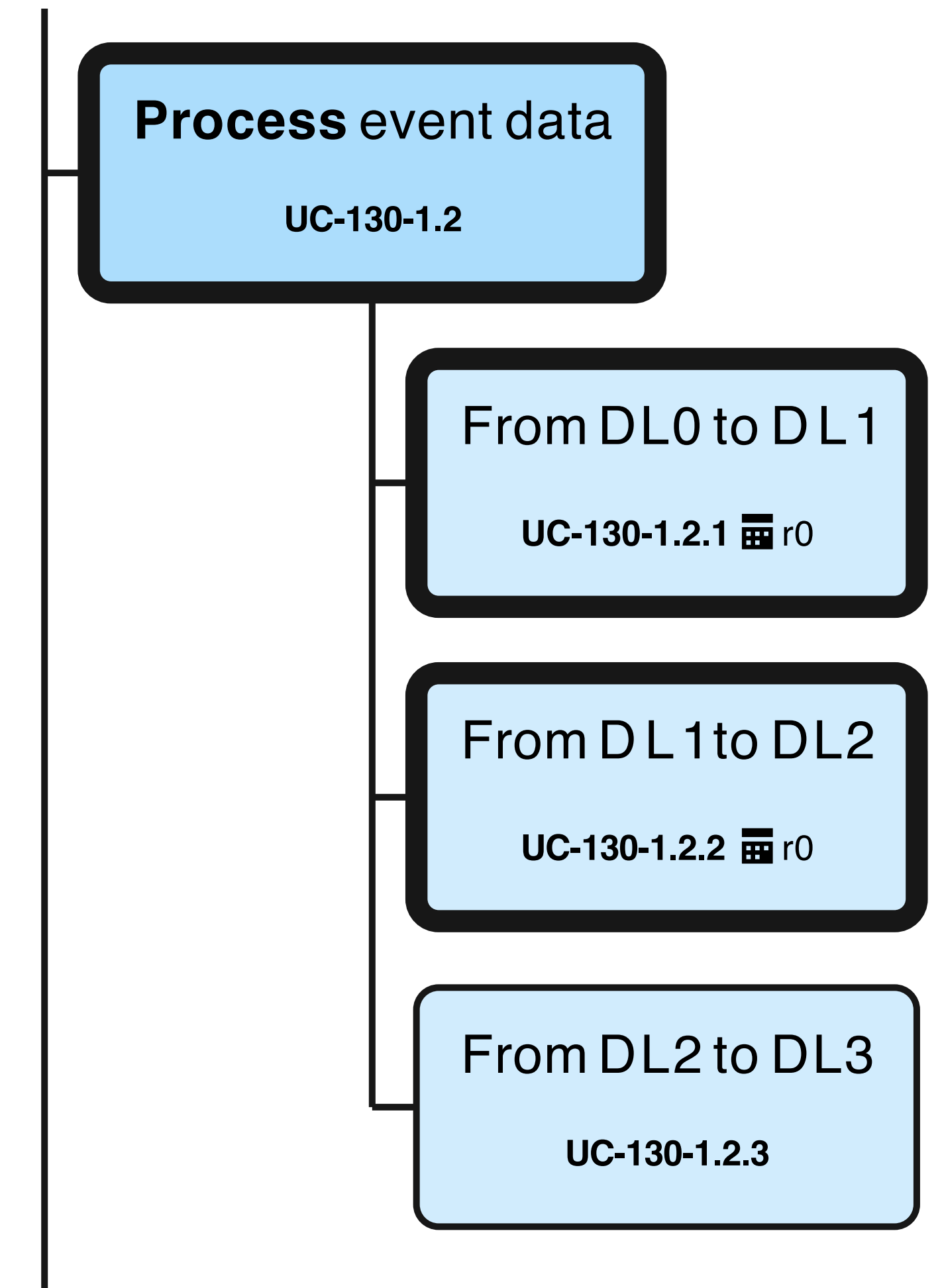
```
requirement_list:  
- number: 0000  
  name: DataPipe software licensing  
  statement: >=  
    All DataPipe software shall follow the CTA0 open-source  
    policy defined in [AD-4].  
  category: Software License  
  explanatory_notes: >=  
    nothing  
  verification: Inspection  
  
- number: 0100  
  name: data product model and format  
  statement: >=  
    DataPipe shall produce data products that follow the "C  
    mo    igned by the CTA0 data model group and in a  
    F    shall be HDF5. For DL3 the f  
    F  
  ca  
  ve
```

*Could use some expert help
with this!*

Goal: Integration of existing software on DPPS test cluster

Needs from DataPipe:

- ▶ **Rel 0.0 (Dec 2023):** Nothing but project docs (need to have Workload + BDMS integrated first)
- ▶ **Rel 0.x (May 2023):** UCs shown to the right →
 - Test will be to run SimPipe to generate simulations, process them to DL2-geometry
 - Could go further if we want



Release 1.0: Generate IRFs for Phys and Public Performance Plots

▶ DataPipe → Most simulation-processing UCs!

Rel 1.x: generate IRFs for Data Challenge

▶ DataPipe → Improved processing, event types, etc. Maybe missing UCs

Release 2.x: Process first Observations from LST-1 (Cat-B)

▶ DataPipe → Observation processing UCs, CalibPipe interface

Release 3.0: Process data from multi-telescope array (Cat-B)

▶ DataPipe → Refined observation processing UCs + CalibPipe interface

...

Release N: Re-process all data with full CTAO array with excellent science performance (Cat-C)

▶ Everything

**We don't know the
detailed timescale
yet**

Software Programming Standards:

GitLab platform + SonarQube