

PDS4 Data Engineering Support

PDS System Design Review II

Mitch Gordon
PDS Rings Node

Overview

- Now for something different.
- Overview of "The Process" – Taking new data from initial mission design through PDS distribution & maintenance.
- Look at Data Engineering support, tools & mechanics, in the context of "The Process".

The Data Engineering Menu*

1. Direct Coordination
2. Documents
3. Design Tool
4. Validate Tool
5. Dictionary Tools
6. Ingest Tool
7. Registry Service
8. Subscription Service
9. Transport Service
10. Maintenance Tools
11. Node Tools

Green – Ready for Build 2

Plum – Prototypes available for Build 2

Blue – Not available for Build 2

* Not an exhaustive list (e.g., PDS3 to PDS4 migration tools).

The Process – Four Stages*

1. Planning and Design
 2. Development and Testing
 3. Data Production and Delivery
 4. Ingestion, Distribution, and Maintenance
- I'll go quickly through all four then come back to Phase 1 & 2 for a bit more detail.
 - * Not as linear as it looks

Stage 1: Planning and Design

- **Major Tasks**
 - Identify Archive Contents
 - Pipeline Design
 - Mission Dictionary Design
 - Observational Data Product Label Design
- **From the Data Engineering Menu**
 - Direct Coordination
 - Documents
 - XML Specific Editors

Stage 2: Development and Testing (1 of 2)

- **Major Tasks**

- Pipeline Development
- Mission Dictionary Development
- Observational Data Product Label Development
- Write Supplemental Products
- Validation & Testing
- Pipeline Peer Review & Lien Resolution

Stage 2: Development and Testing (2 of 2)

- **From the Data Engineering Menu**
 - Direct Coordination
 - Documents
 - XML Specific Editors
 - Dictionary Tools
 - Discipline Node Tools

Stage 3: Data Production and Delivery

- **Major Tasks**
 - Pipeline Execution & Validation
 - Delivery to PDS
- **From the Data Engineering Menu**
 - Documents
 - Validate Tool
 - Ingest Tool

Stage 4: Ingestion, Distribution, and Maintenance

- **Major Tasks**
 - Data Ingestion & Registration
 - Initial-Data Peer Review & Lien Resolution
 - Data Distribution
 - Data Maintenance
- **From the Data Engineering Menu**
 - Ingest Tool
 - Registry Service
 - Subscription Service
 - Transport Service
 - Archive Maintenance Tools

Phase 1: Planning and Design

- **Major Tasks**

- Identify Archive Contents
- Pipeline Design
- Mission Dictionary Design
- Observational Data Product Label Design

- **From the Data Engineering Menu**

- Direct Coordination
- Documents
- XML Specific Editors

Archive Contents & Organization

- Bundle
 - Logically an aggregation of collections
 - Physically an XML label with embedded inventory table
- Collections
 - Logically an aggregation of products of a common theme
 - Raw Data, Calibrated Data, Documents, etc.
 - Physically a file with an inventory table and an XML label file.
- Products
 - 'An XML label file plus the objects it describes.'

Stage 2: Development and Testing

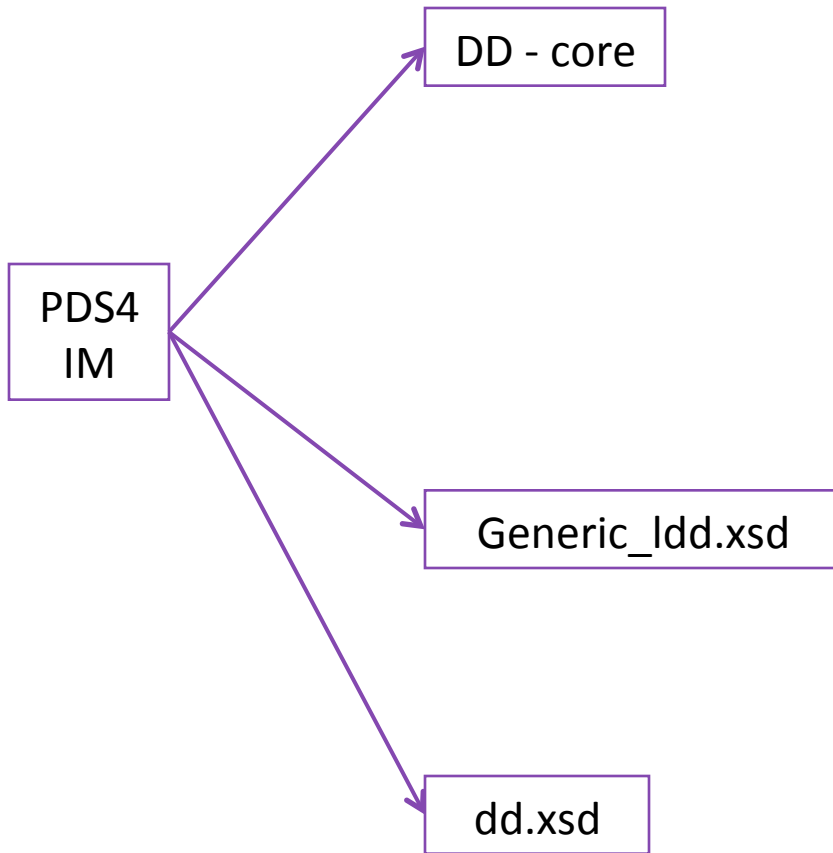
- **Major Tasks**

- Mission Dictionary Development
- Observational Data Product Label Development

- **From the Data Engineering Menu**

- Direct Coordination
- Documents
- XML Specific Editors
- Dictionary Tools

PDS4
IM



→ Protégé or some other entity indistinguishable from Magic

DD - core

DD - Mission

PDS4
IM

Generic_Idd.xsd

dd.xsd

mission_Idd.xsd

DD - core

DD - Mission

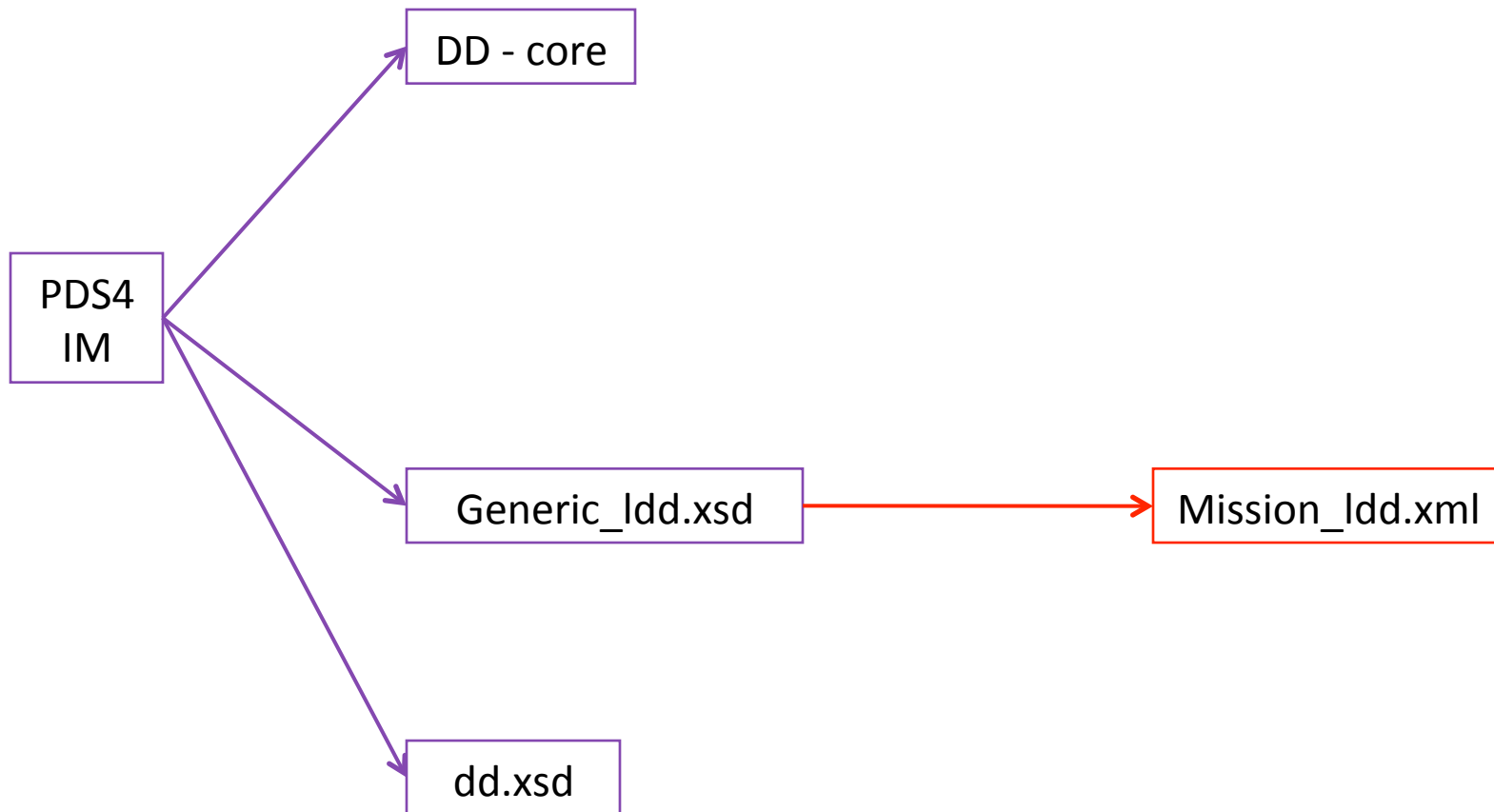
PDS4
IM

Mission
IM

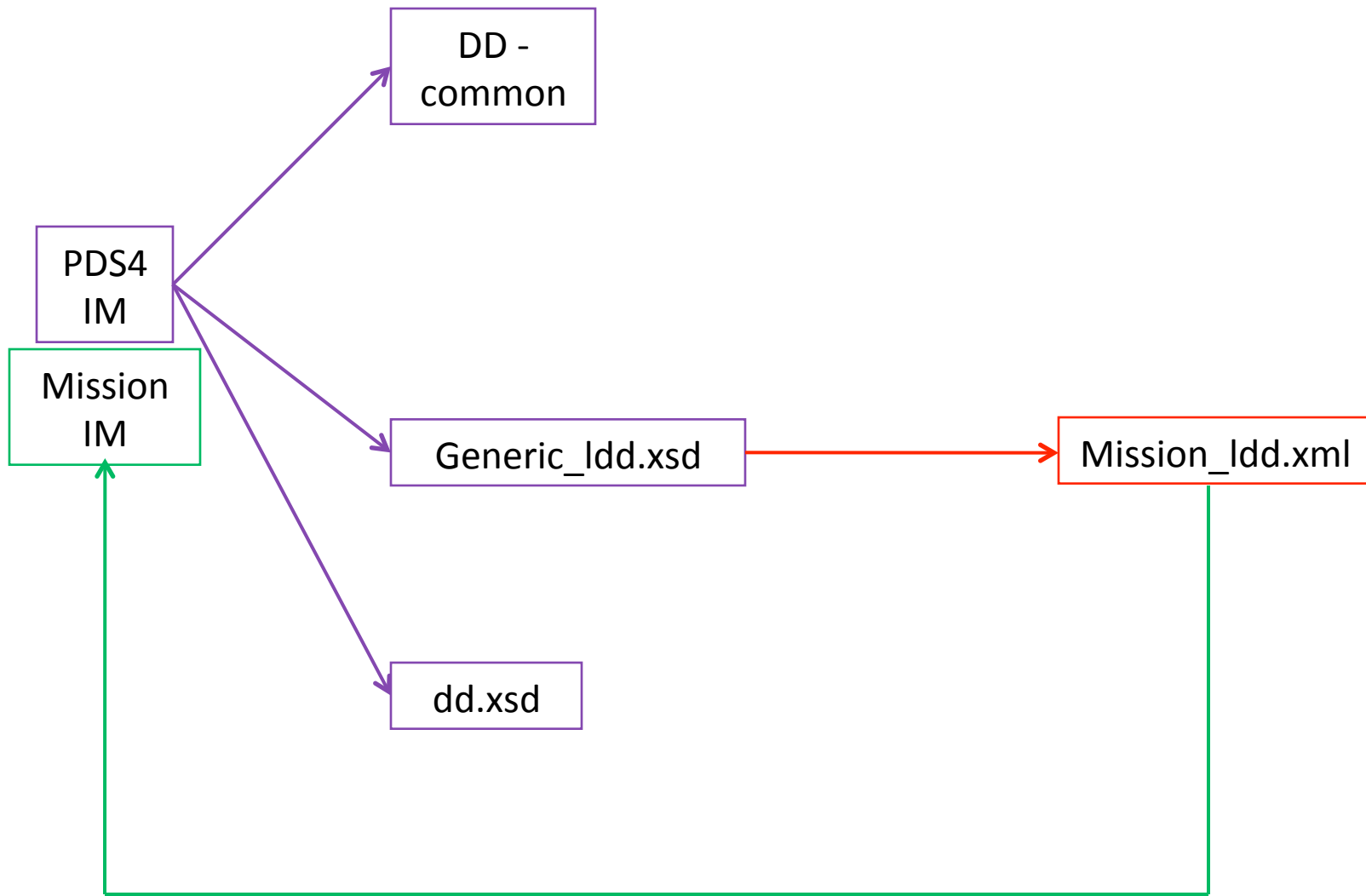
Generic_Idd.xsd

dd.xsd

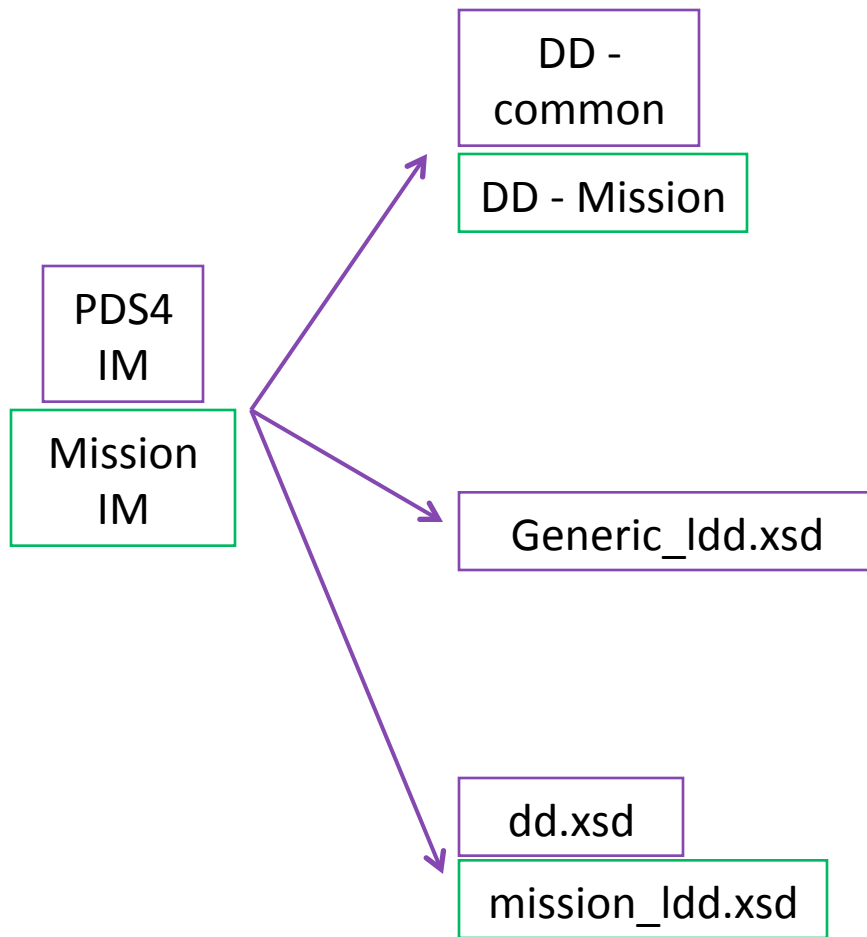
mission_Idd.xsd



→ Protégé
→ Tool 1



- Protégé
- Tool 1
- Tool 2



→ Protégé

Stage 2: Development and Testing

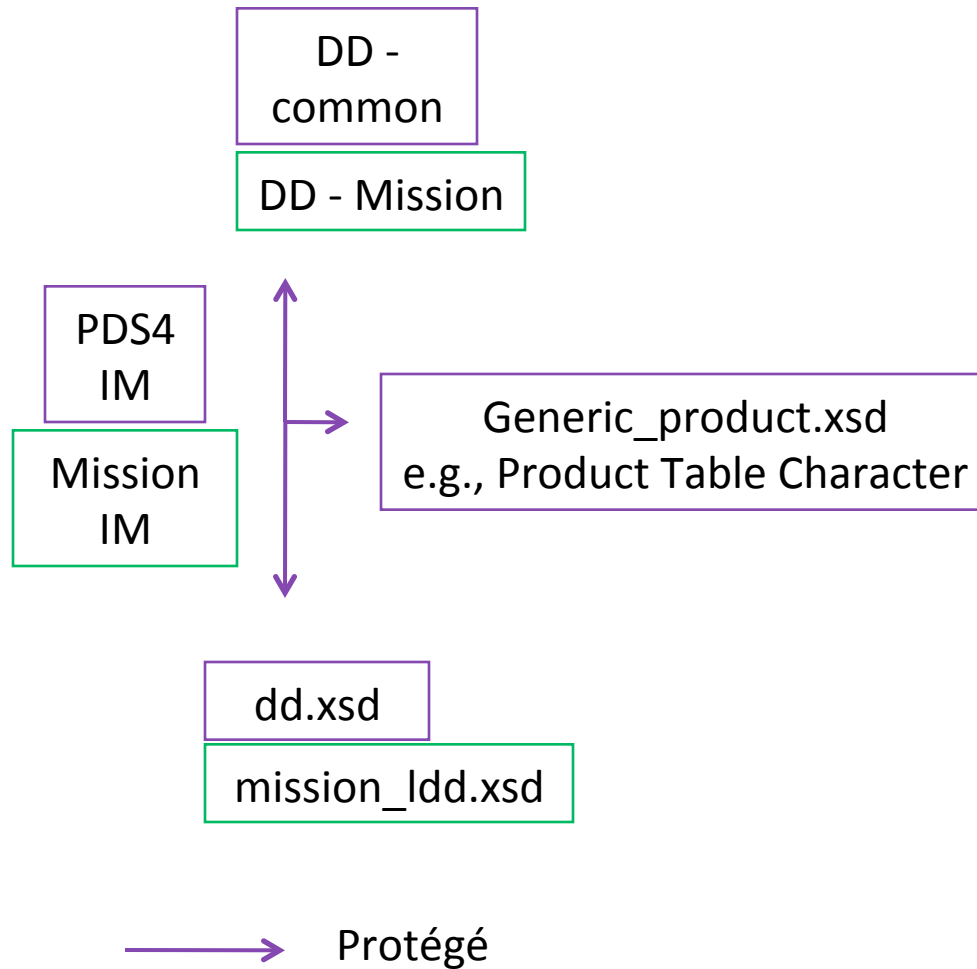
- **Major Tasks**

- Mission Dictionary Development
- Observational Data Product Label Development

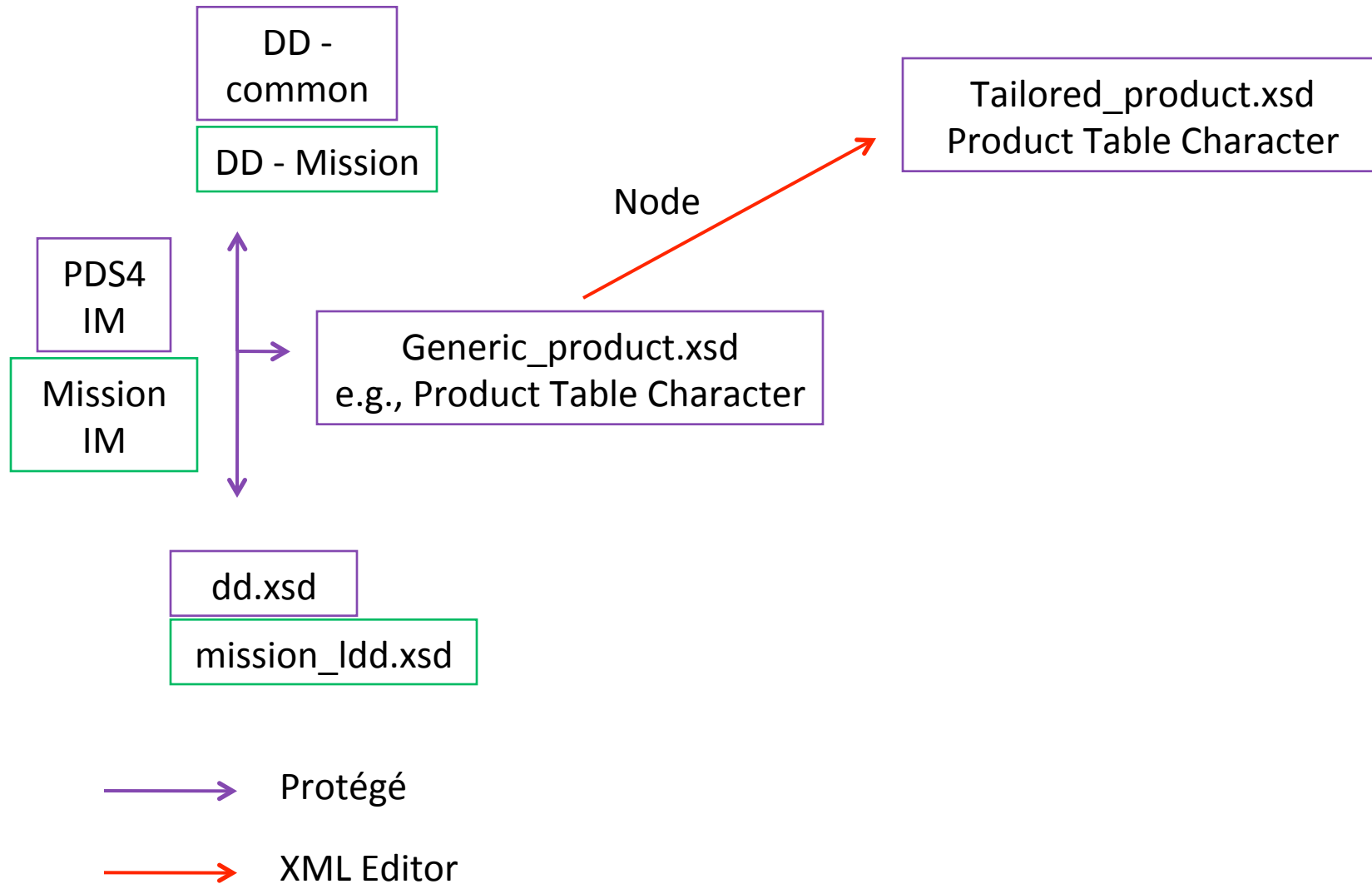
- **From the Data Engineering Menu**

- Direct Coordination
- Documents
- XML Specific Editors
- Dictionary Tools

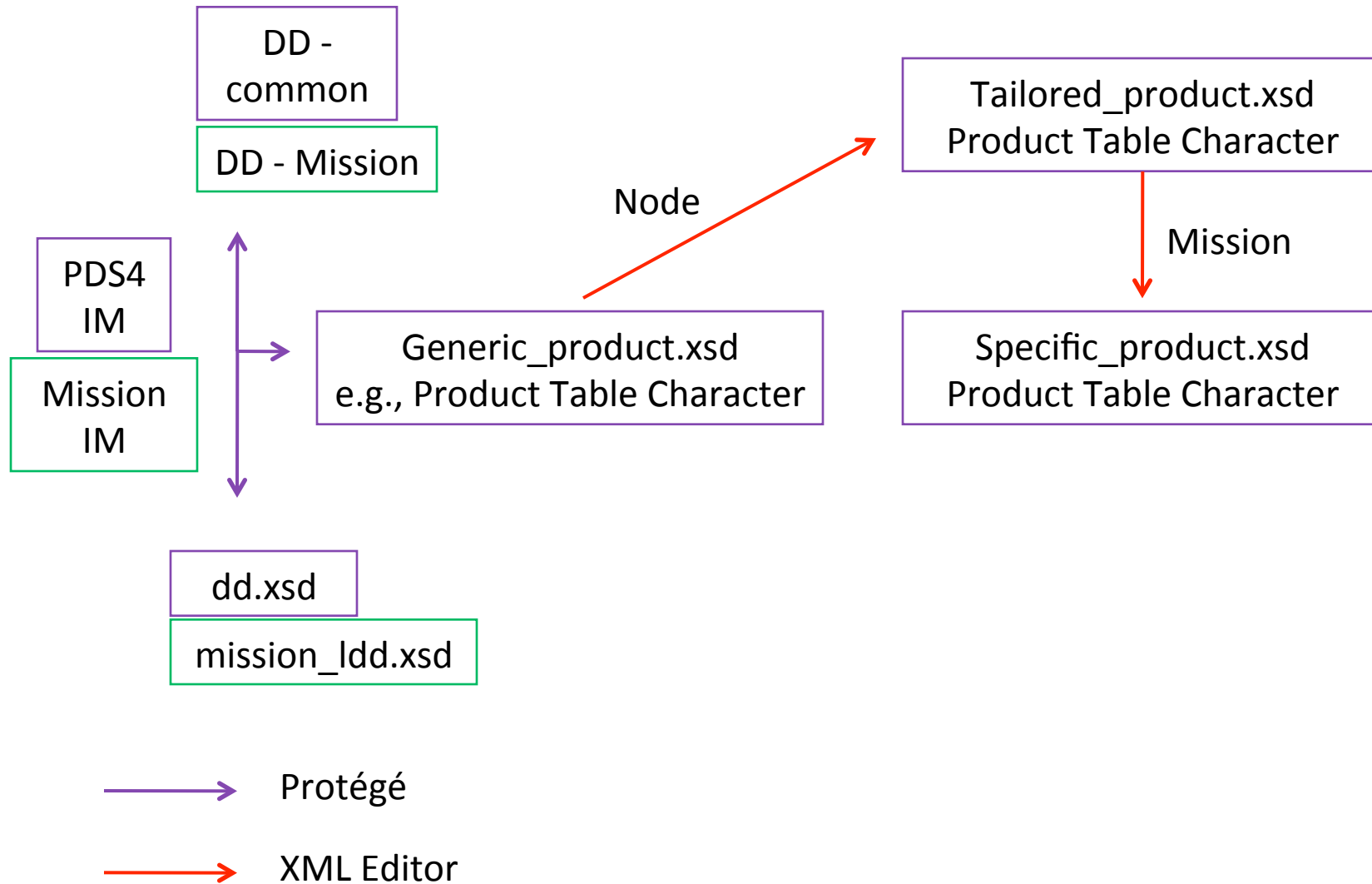
Stage 2: Development and Testing



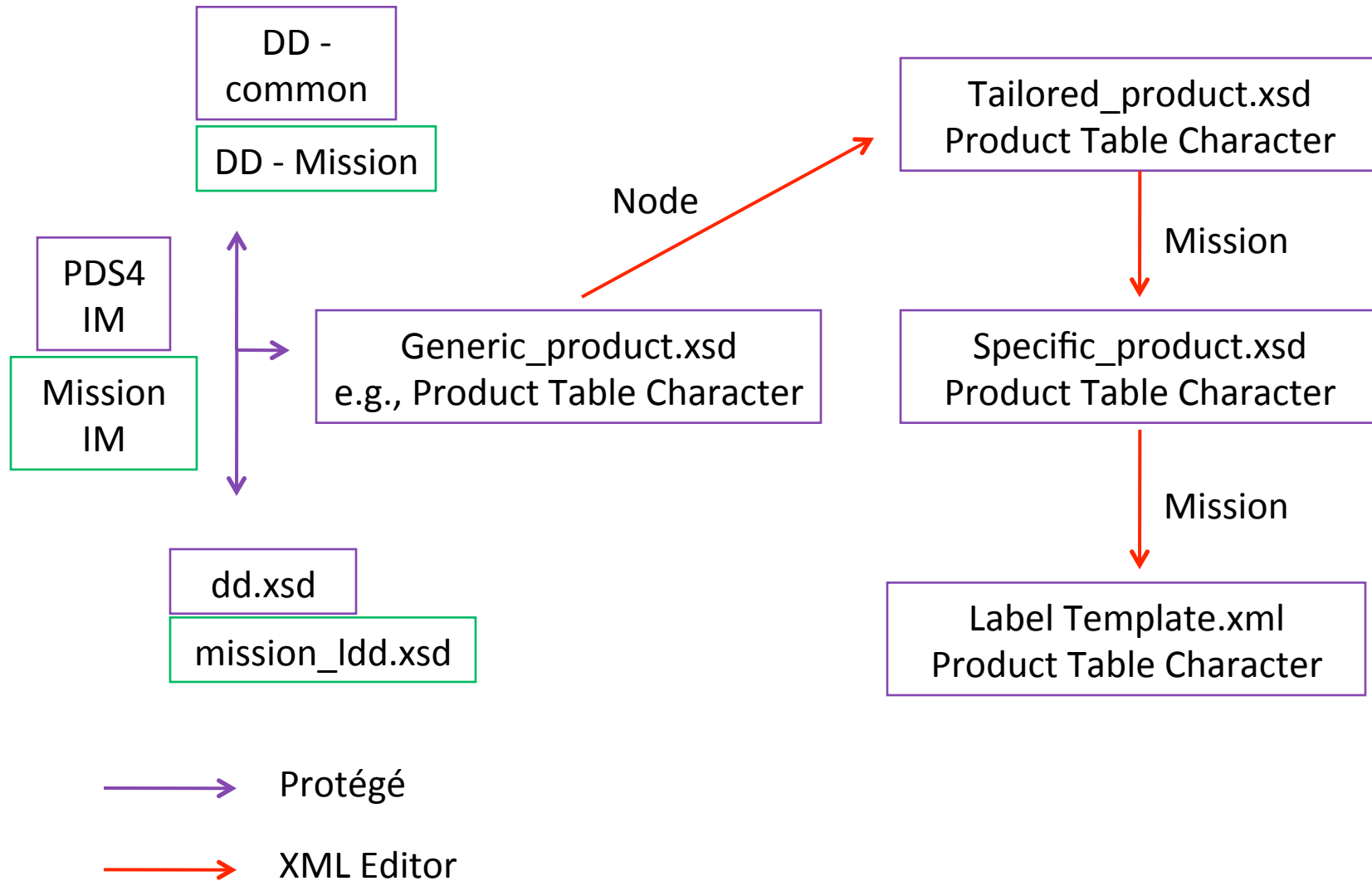
Stage 2: Development and Testing



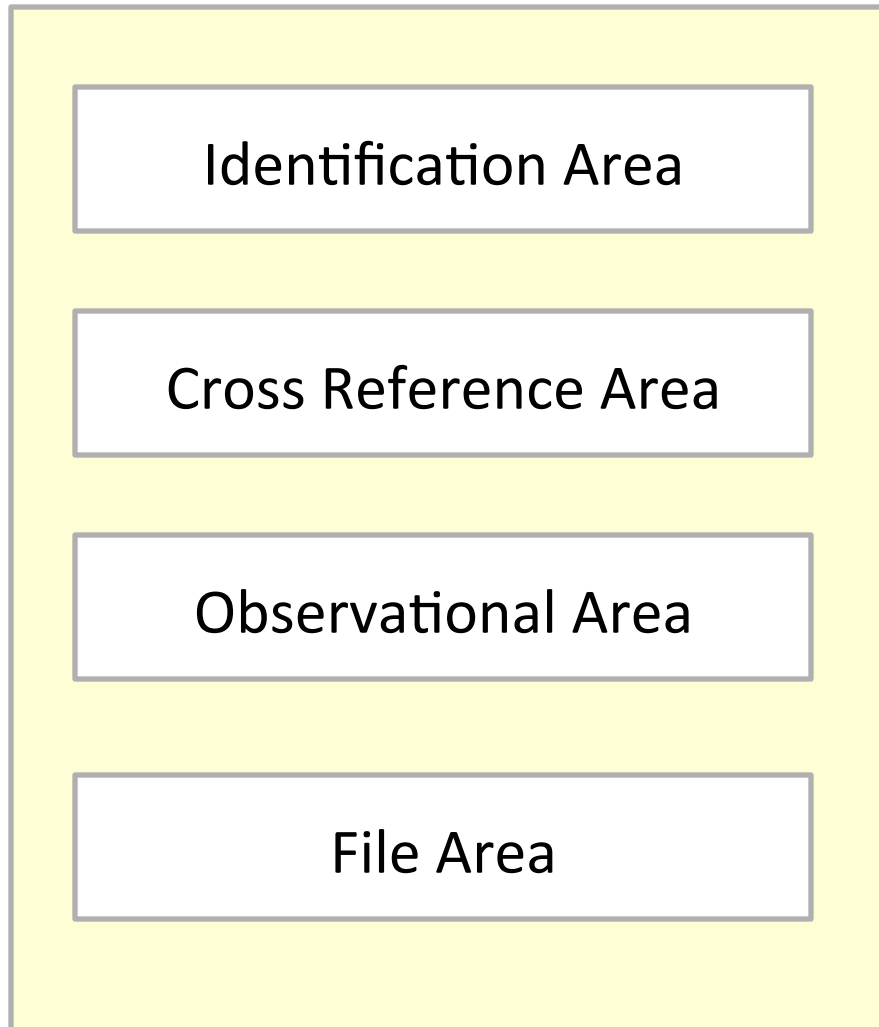
Stage 2: Development and Testing



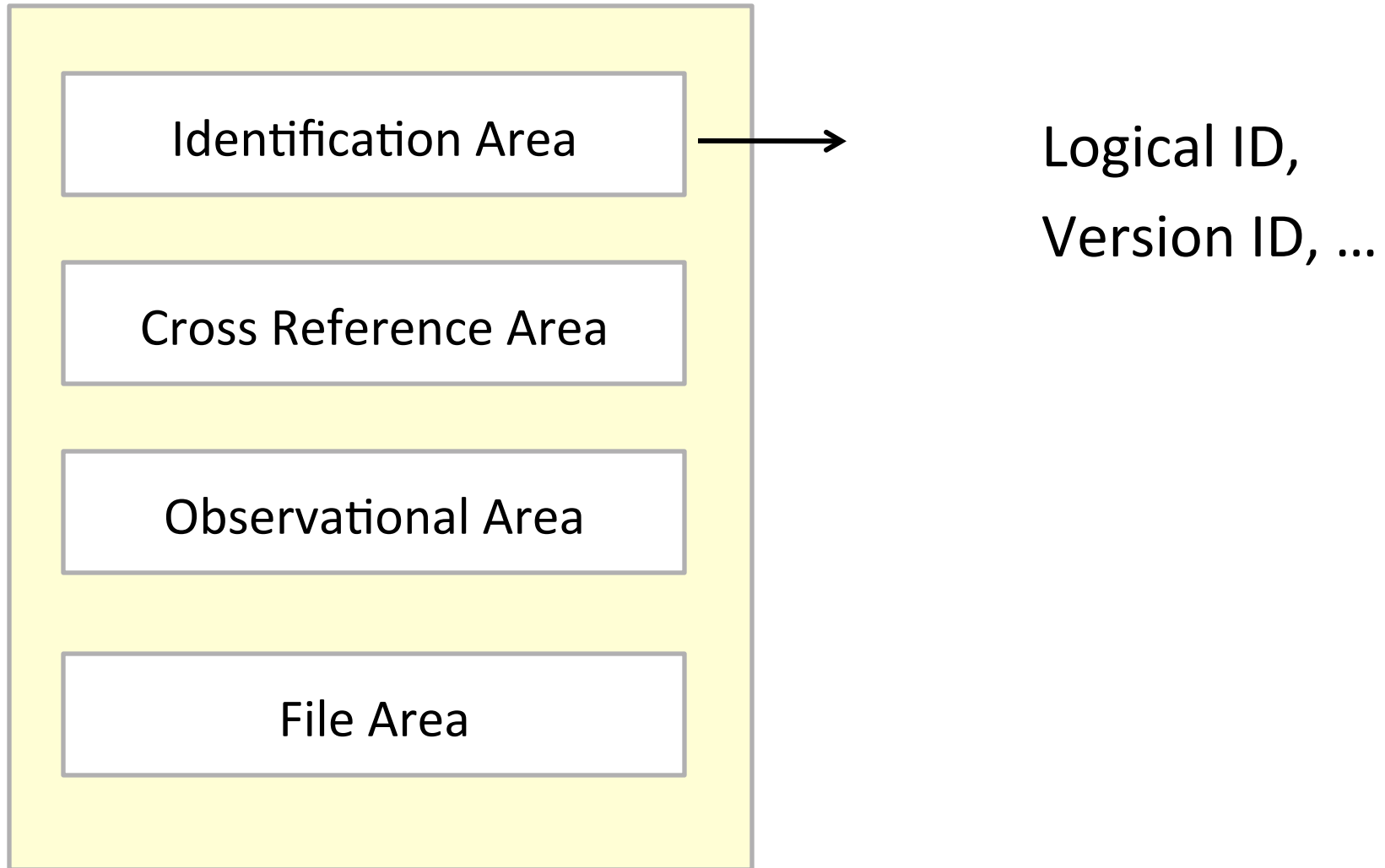
Stage 2: Development and Testing



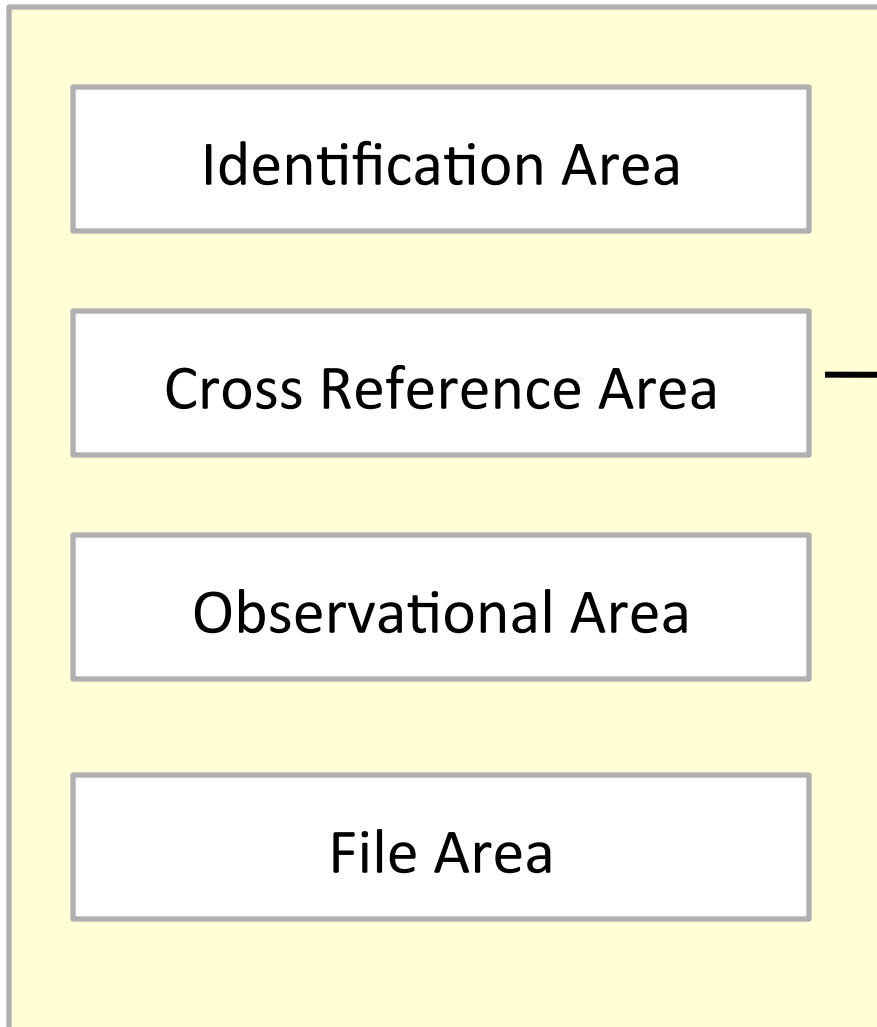
Observational Data Product Labels



Observational Data Product Labels



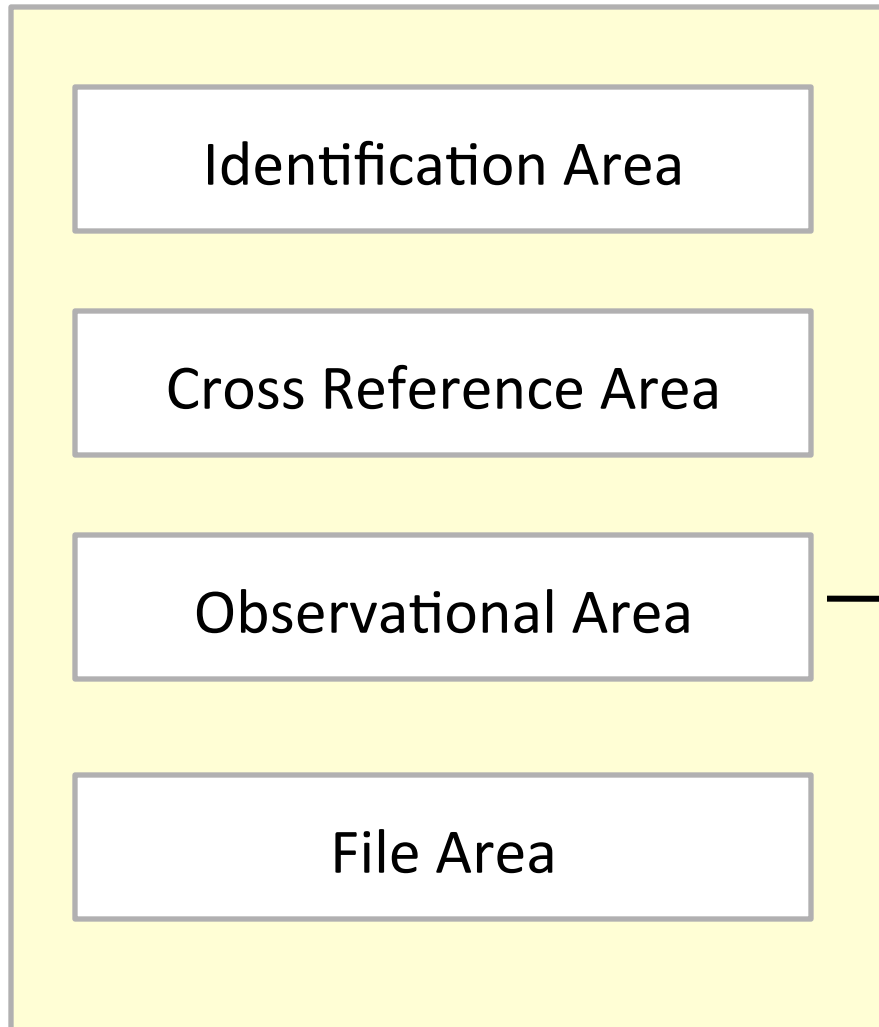
Observational Data Product Labels



For example, in label for a calibrated image:

browse image,
raw image,
calibration files,
calibration docs,
instrument,
etc.

Observational Data Product Labels



basic observation
parameters:

start time, ...

node classes, e.g.,

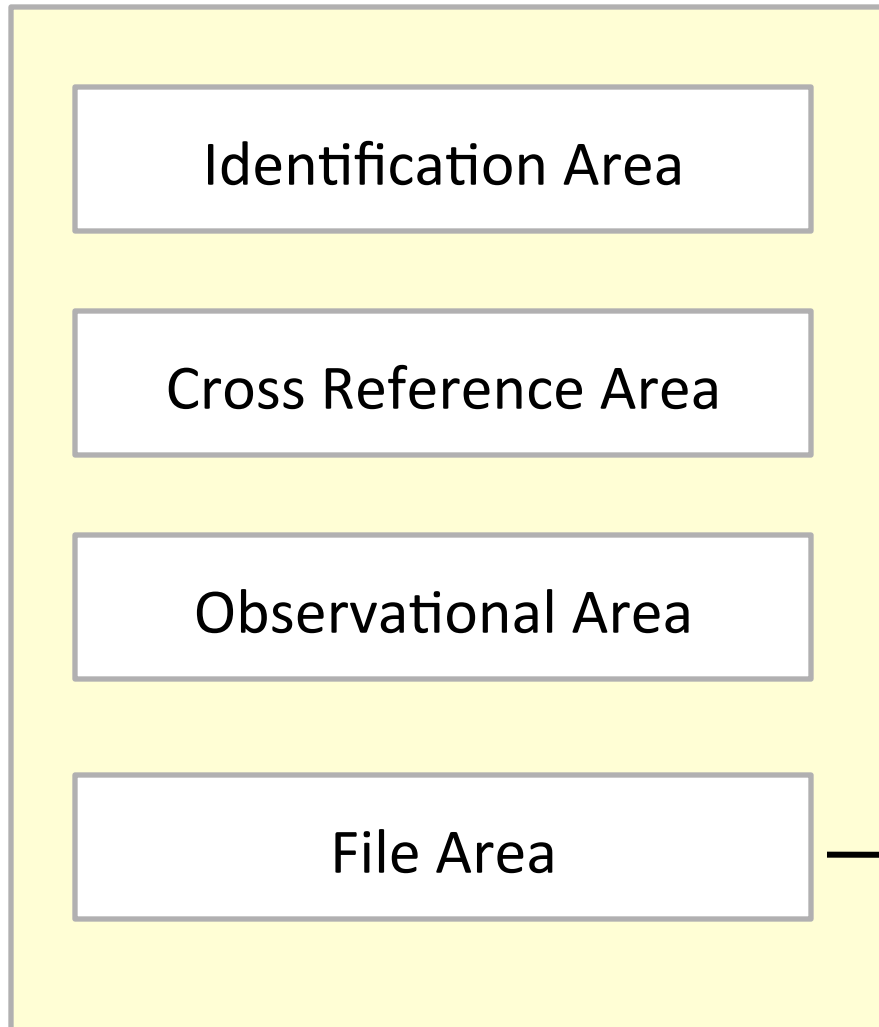
ring geometry,

mission classes, e.g.,

instrument

settings, ...

Observational Data Product Labels



Description and name of the file,

Descriptions of the objects in the file:

Header

Image

Character Table

etc.

Product Labels in oXygen

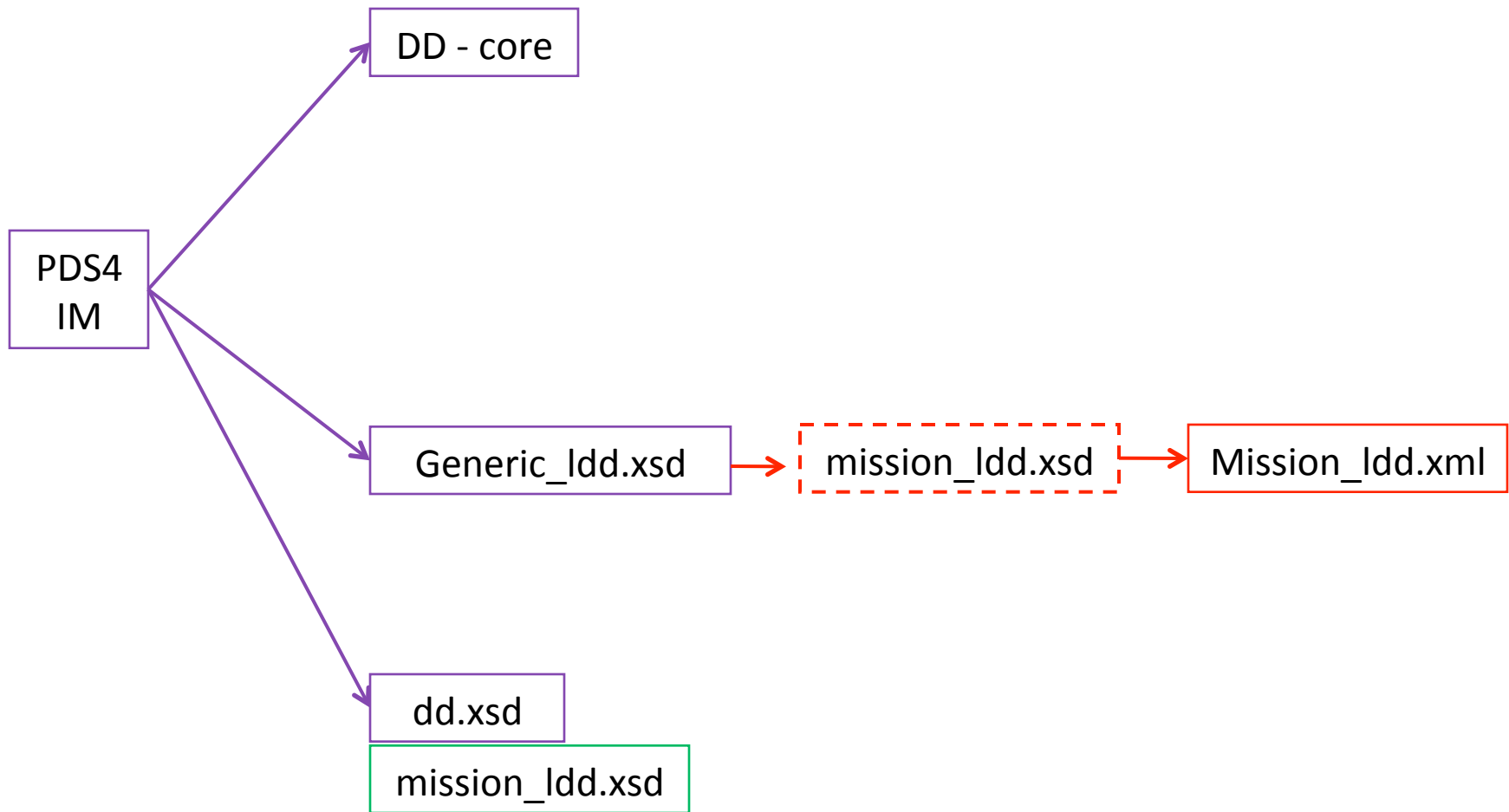
- Generic schema
 - Design View
 - Text view
- XML Label
 - Values
 - Validation

PDS4 Data Engineering Support

A black and white photograph of a desert landscape. The foreground is a dark, shadowed sand dune. In the middle ground, there are several bright, glowing horizontal lines that appear to be reflections on a wet surface or perhaps light trails. The background shows a dark horizon line under a dark sky.

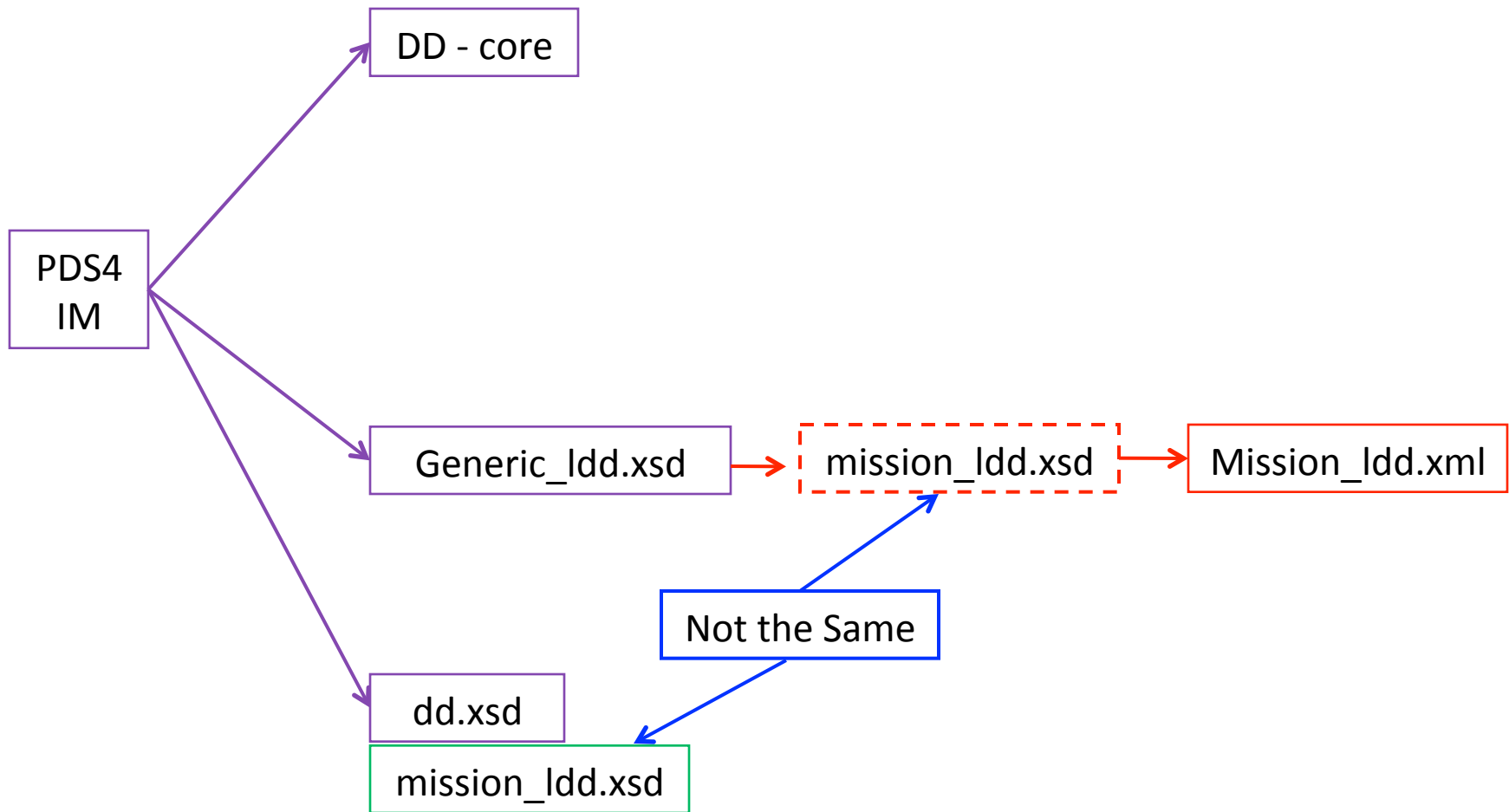
Questions?

Backups



—→ Protégé

—→ Tool 1



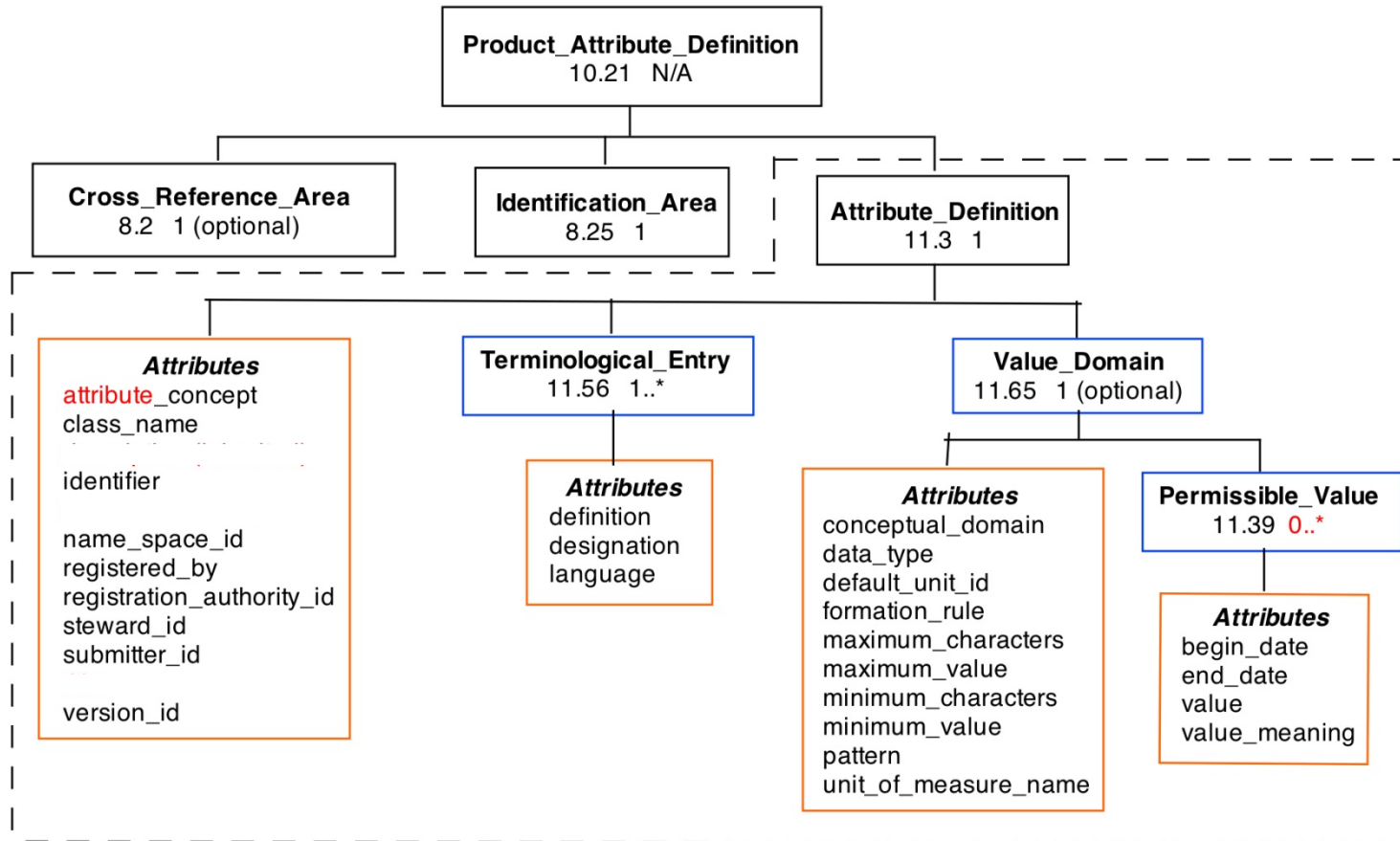
—→ Protégé

—→ Tool 1

Discipline Node Tools

- Growing out of our own prototyping exercises.
- May see an enhanced 'Red Tool'.
- We are seeing at least a couple of tools that could form the basis for part of a mission pipeline – the next talk describes one of these in more detail.

IM & DD Overhead



- For each PDS4 Data Dictionary attribute, you need the classes in blue boxes and meta-attributes in orange boxes.
- The Green tool deals with much of this.