

Reverse engineering metadata for the Materials Project

Shyam Dwaraknath

Research Scientist

Lawrence Berkeley Labs



BERKELEY LAB

What is Materials Project

Home About Apps Documentation API Login

The Materials Project

Harnessing the power of supercomputing and state of the art electronic structure methods, the Materials Project provides open web-based access to computed information on known and predicted materials as well as powerful analysis tools to inspire and design novel materials.

[Learn more](#) [Tutorials](#) [Sign In or Register](#) to start using

News

May 20, 2019: **Database V2019.05 Released**
We have discontinued nearly 15,000 materials and added over 3,000 new ones. [More info.](#)

March 11, 2019: **New App to Predict Interface Reactions Between Solids**
We are proud to officially launch an application that calculates possible interface reactions between solids using thermodynamic driving forces. [More info.](#)

Feb. 13, 2019: **Database V2019.02 Released**
We have added over 47,000 new compounds. We also improved feedback on structure submission via the Crystal Toolkit. [More info.](#)

Nov. 1, 2018: **Database V2018.11 Released**
In this major release, we have changed the grouping of magnetic materials and fixed many display/labeling issues. [More info.](#)

Aug. 23, 2018: **MP book chapter released**
Our contributed chapter to Springer's *Handbook of Materials Modeling* has been released online.

Electronic Structure

Click and drag to zoom

Indirect K - Γ bandgap = 7.7511 eV

Density of States

sign indicates spin \uparrow/\downarrow

Materials Data

Material Details

- First Magnetic Moment: 0.0000 μ_B
- Formation Enthalpy: -4.1500 eV
- Energy Above Hull: 0.0000 eV
- Density: 7.16 g/cm^3
- Space Group: $Fm\bar{3}m$
- Prototype: PbF_4
- IP: 20.240

EXPLORE MATERIALS

Search for materials information by chemistry, composition, or property

EXPLORE BATTERIES

Find candidate materials for lithium batteries. Get voltage profiles and oxygen evolution data.

VISUALIZE STABILITY

Generate phase and pourbaix diagrams to find stable phases and study reaction pathways

INVENT STRUCTURES

Design new compounds with our structure editor and substitution algorithms

CALCULATE

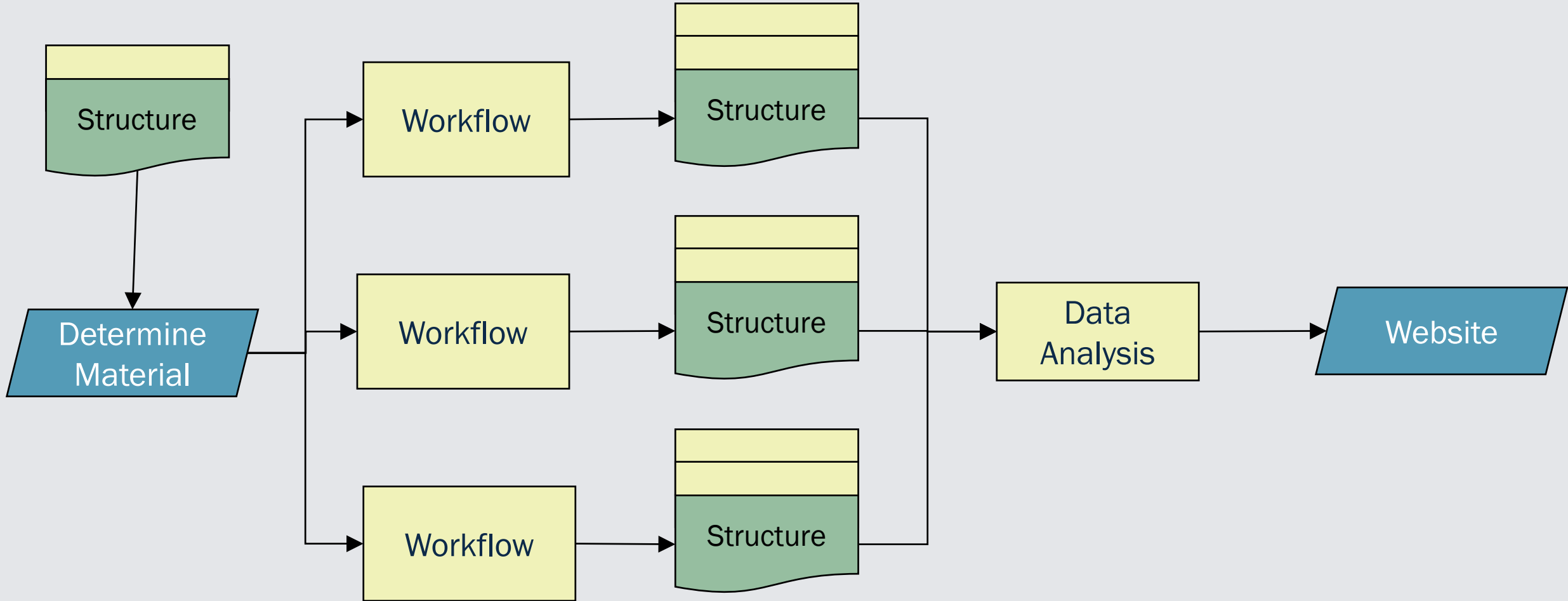
Calculate the enthalpy of 10,000+ reactions and compare with experimental values

Database Statistics

120,612	52,366	35,336	530,243
INORGANIC COMPOUNDS	BANDSTRUCTURES	MOLECULES	NANOPOROUS MATERIALS
13,621	3,003	4,401	16,128
ELASTIC TENSORS	PIEZOELECTRIC TENSORS	INTERCALATION ELECTRODES	CONVERSION ELECTRODES

- Free and open source database of computed materials properties
- Applications to explore data set in "material science" means
- Design computational workflows based on experimental ground truth
- Focus on property diversity over number of structures
- All the software bells and whistles that go with a large web data project

The good old days



What is metadata?

A set of data that describes and gives information about other data.

Descriptive

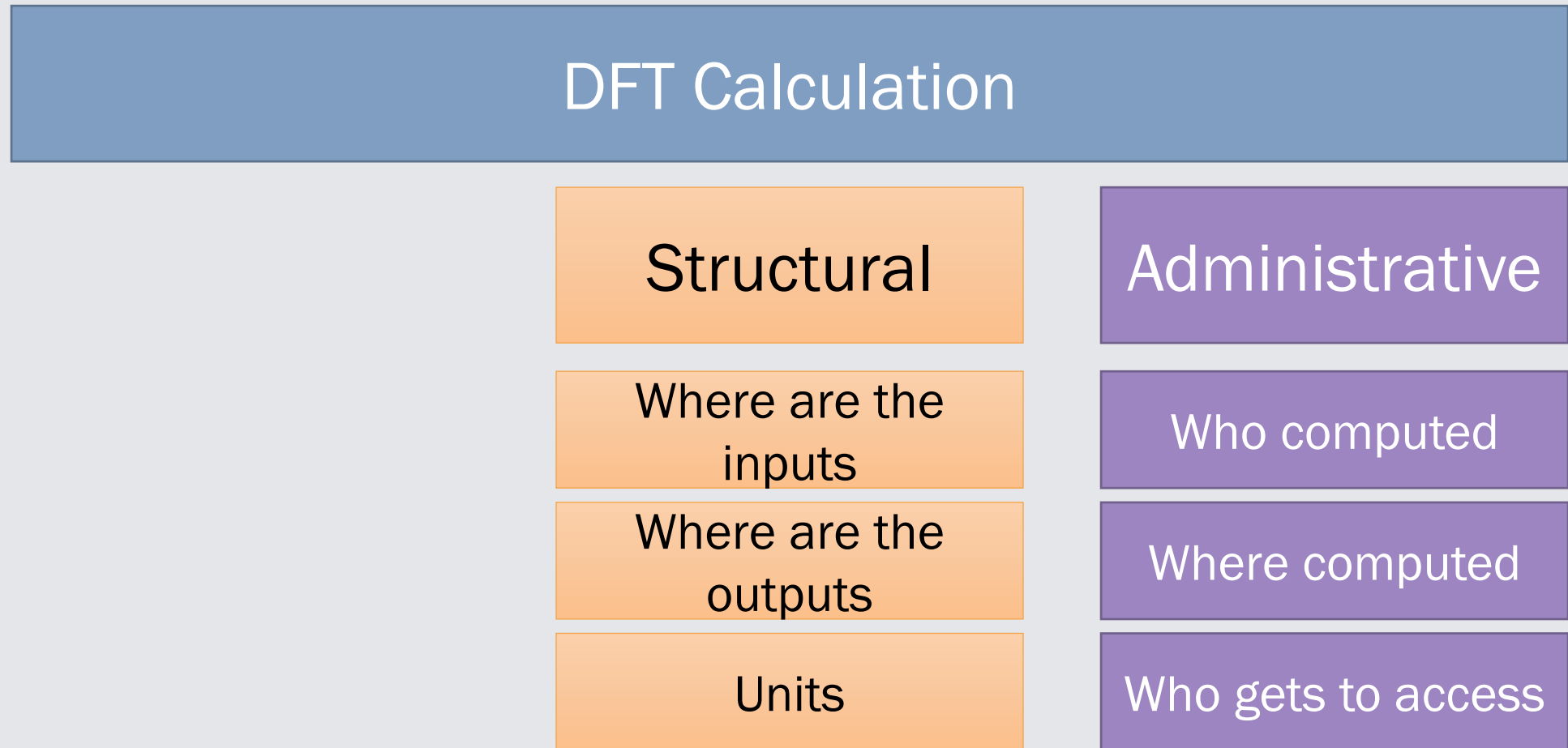
Structural

Administrative

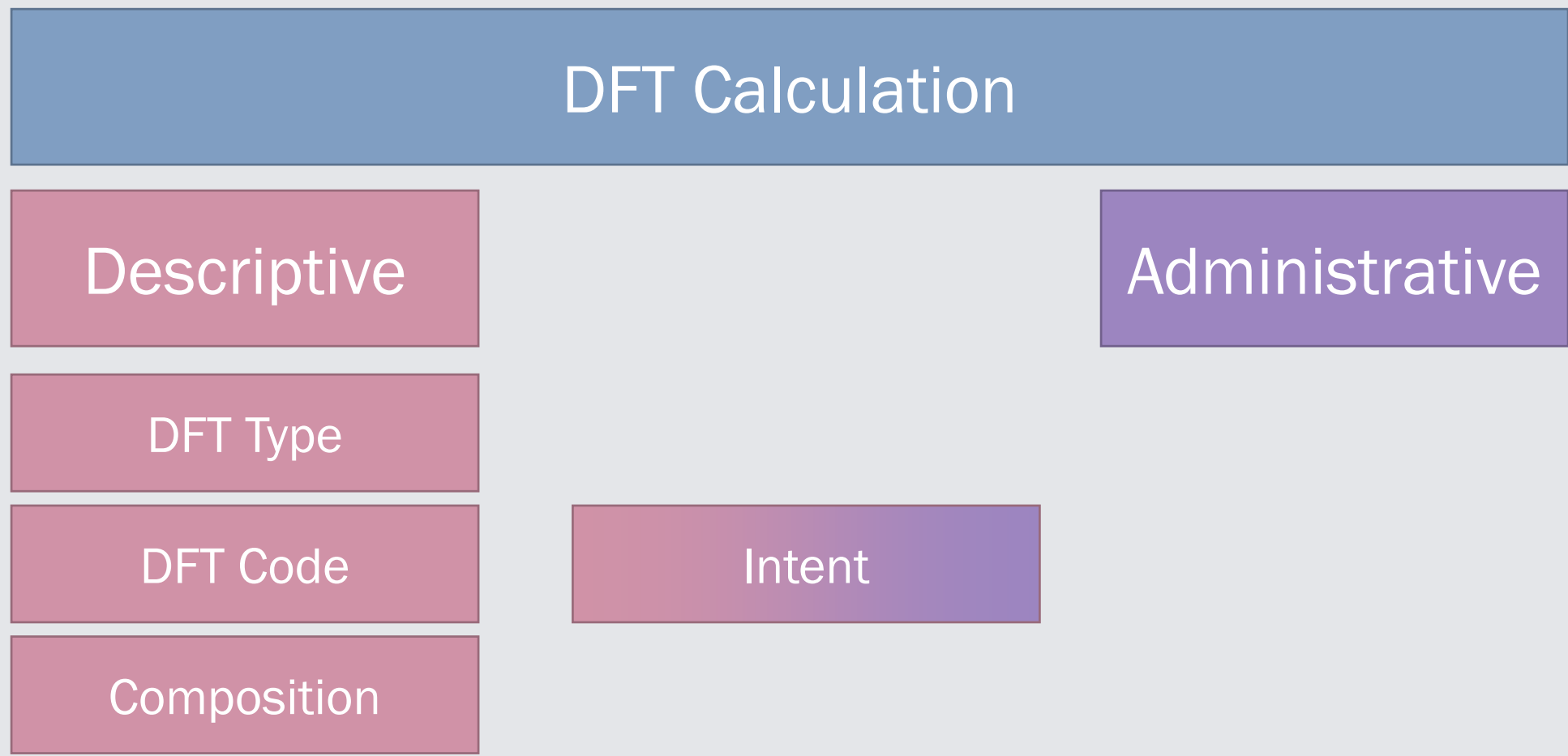
What is metadata?

DFT Calculation		
Descriptive	Structural	Administrative
DFT Code	Where are the inputs	Who computed
Composition	Where are the outputs	Where computed
Volume Change	Units	Who gets to access

Some metadata are constructed implicitly

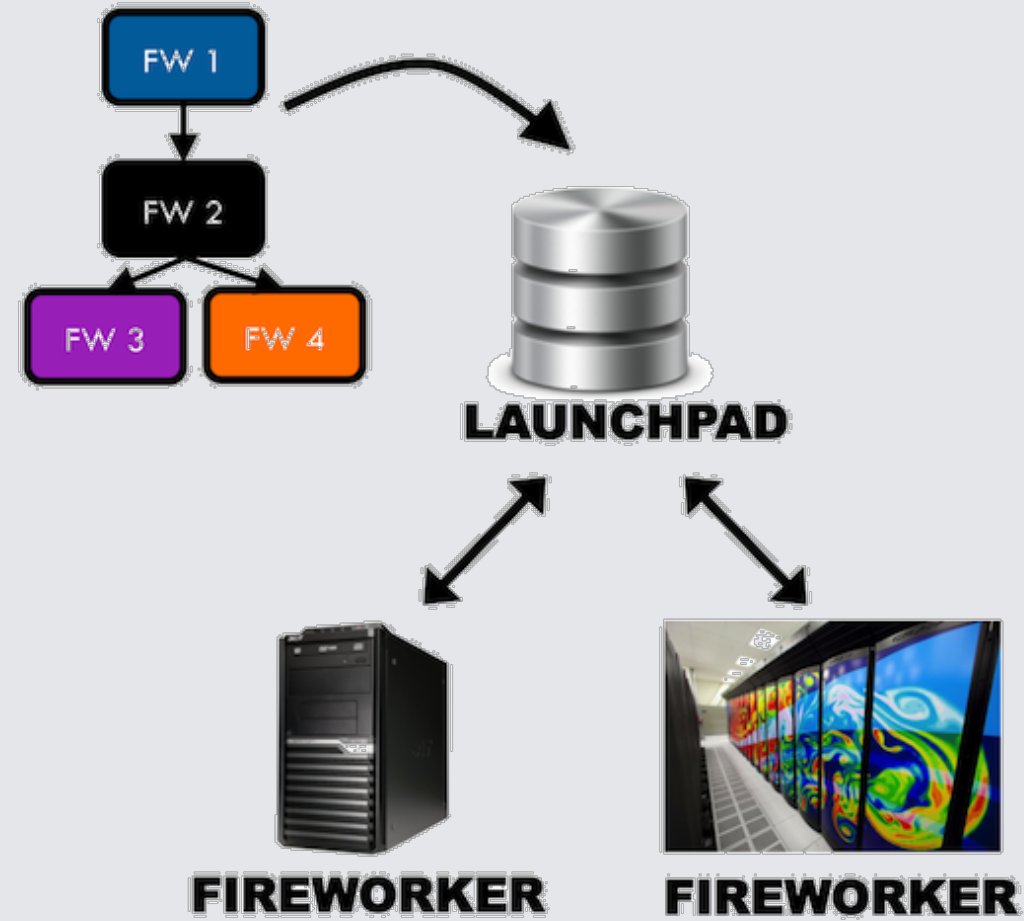


Some metadata are constructed explicitly



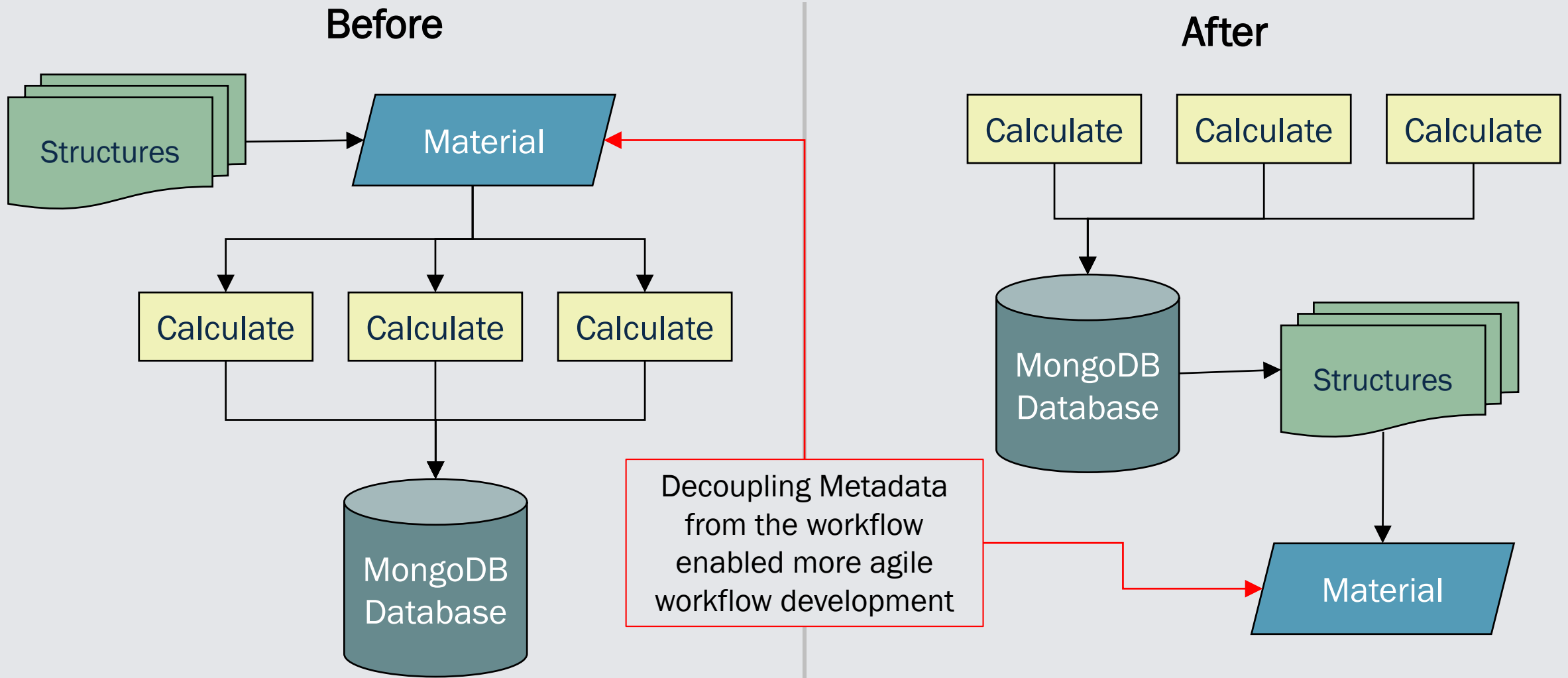
How do you determine the intent of a calculation?

- The purpose of the workflow
- The purpose of the database it was in
- The purpose of the computer it ran on

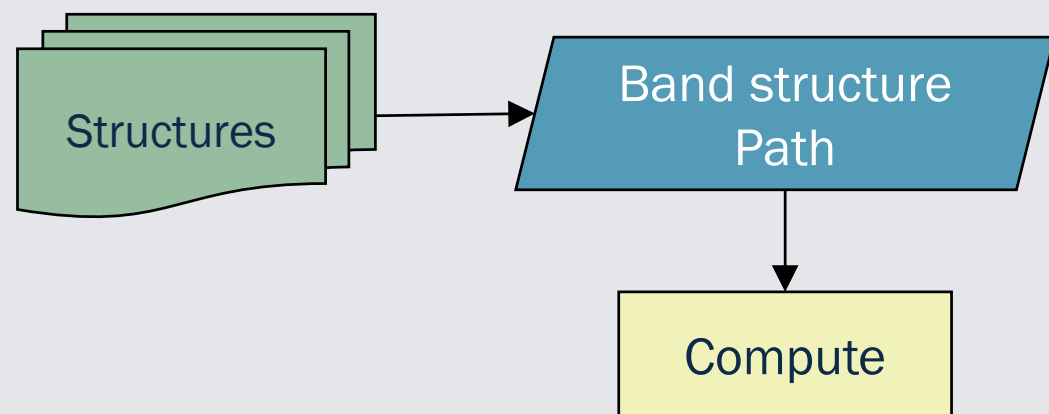


Why not determine the intent as the need arises?

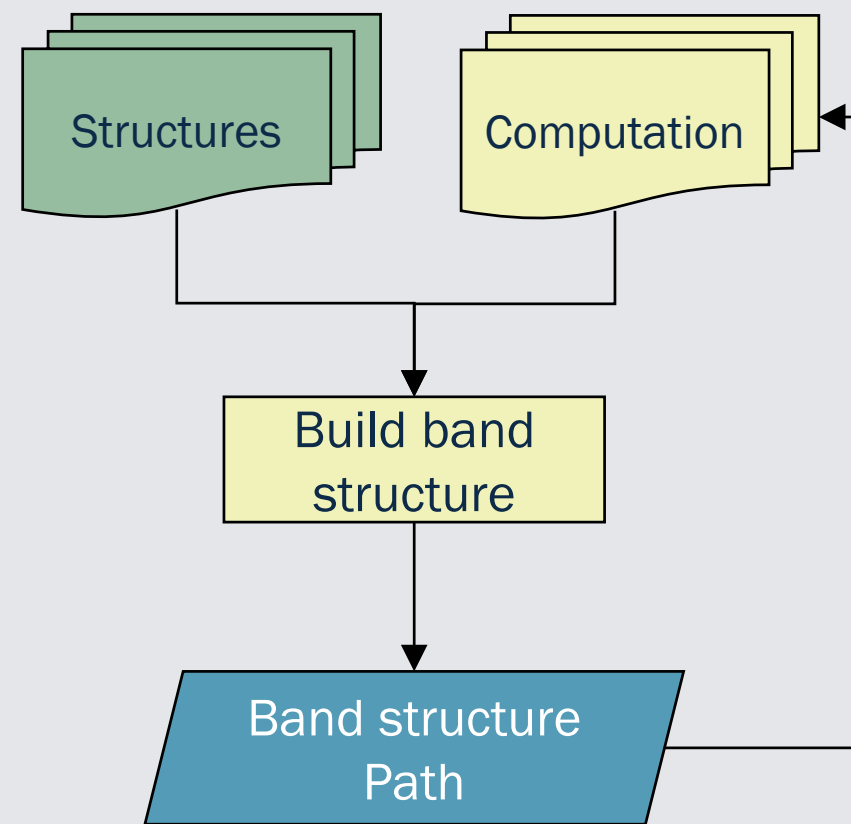
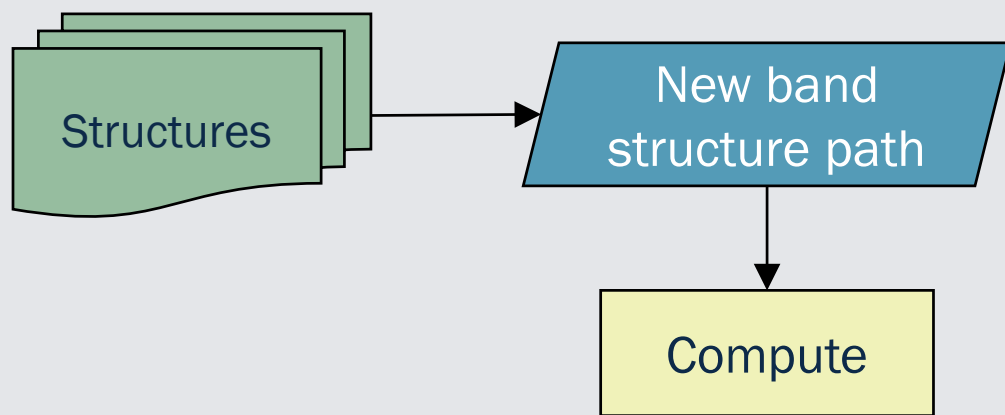
Metadata is technical debt



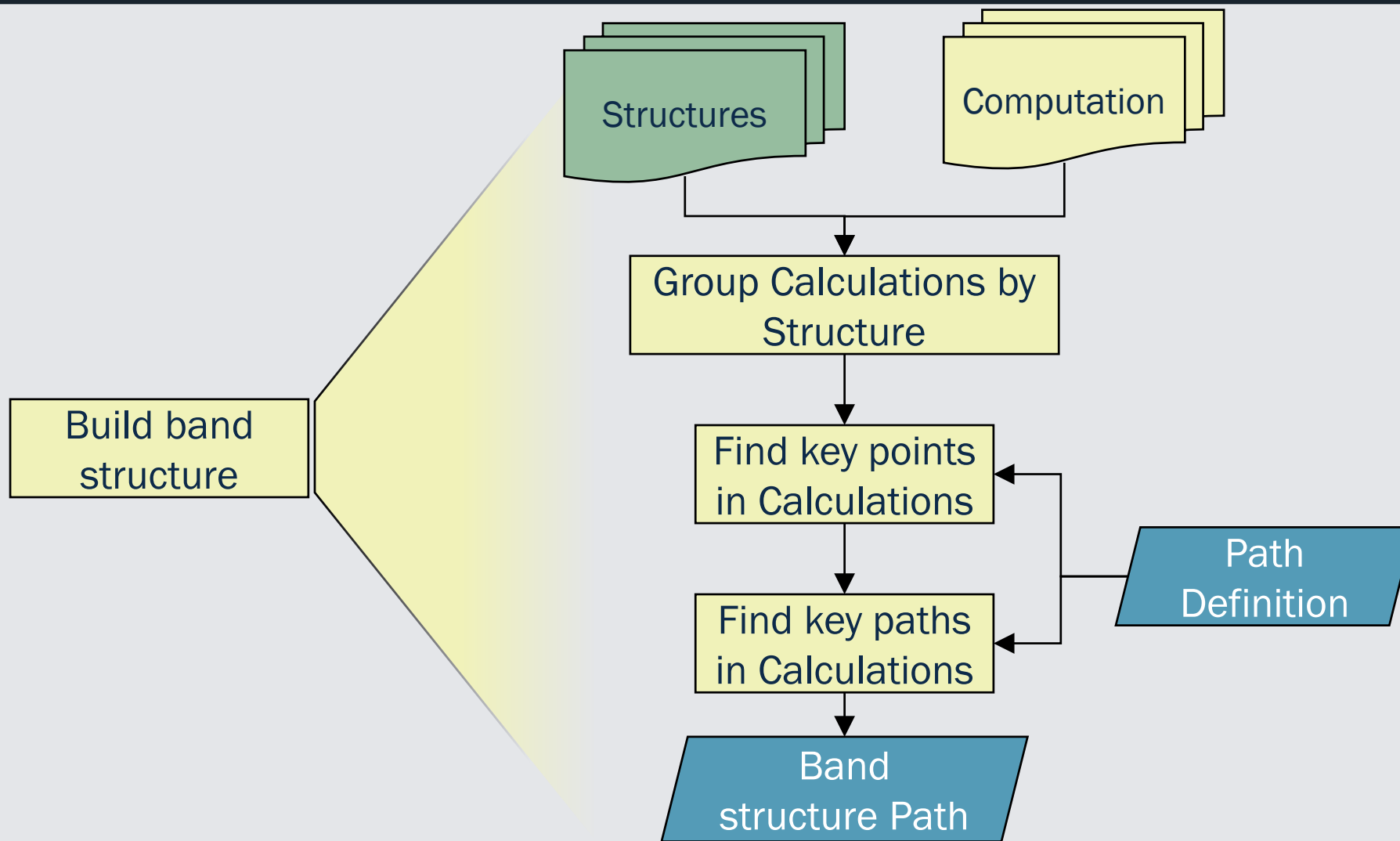
Decoupling metadata makes it declarative



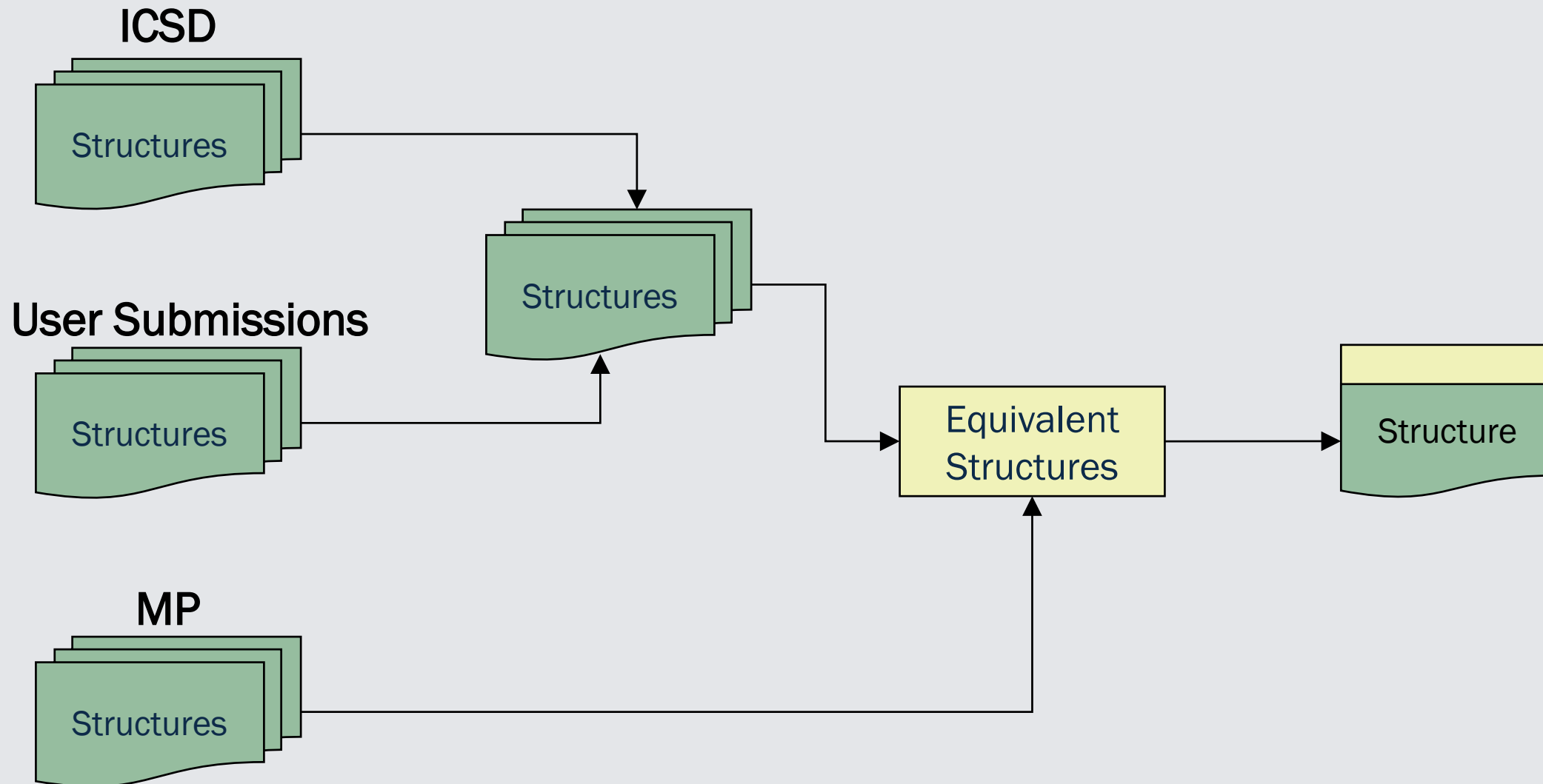
I changed my path!?



Declarative metadata gives us connectivity

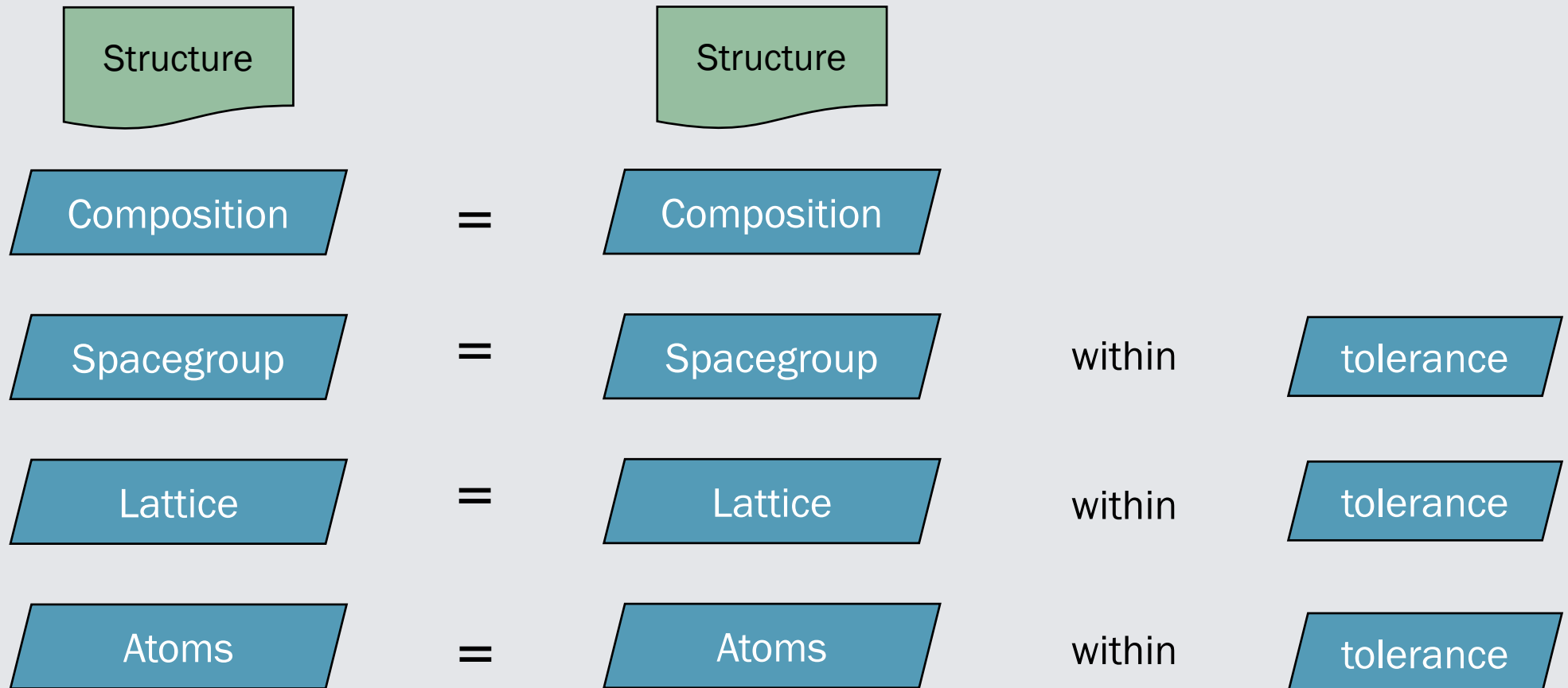


Reverse Engineering Provenance



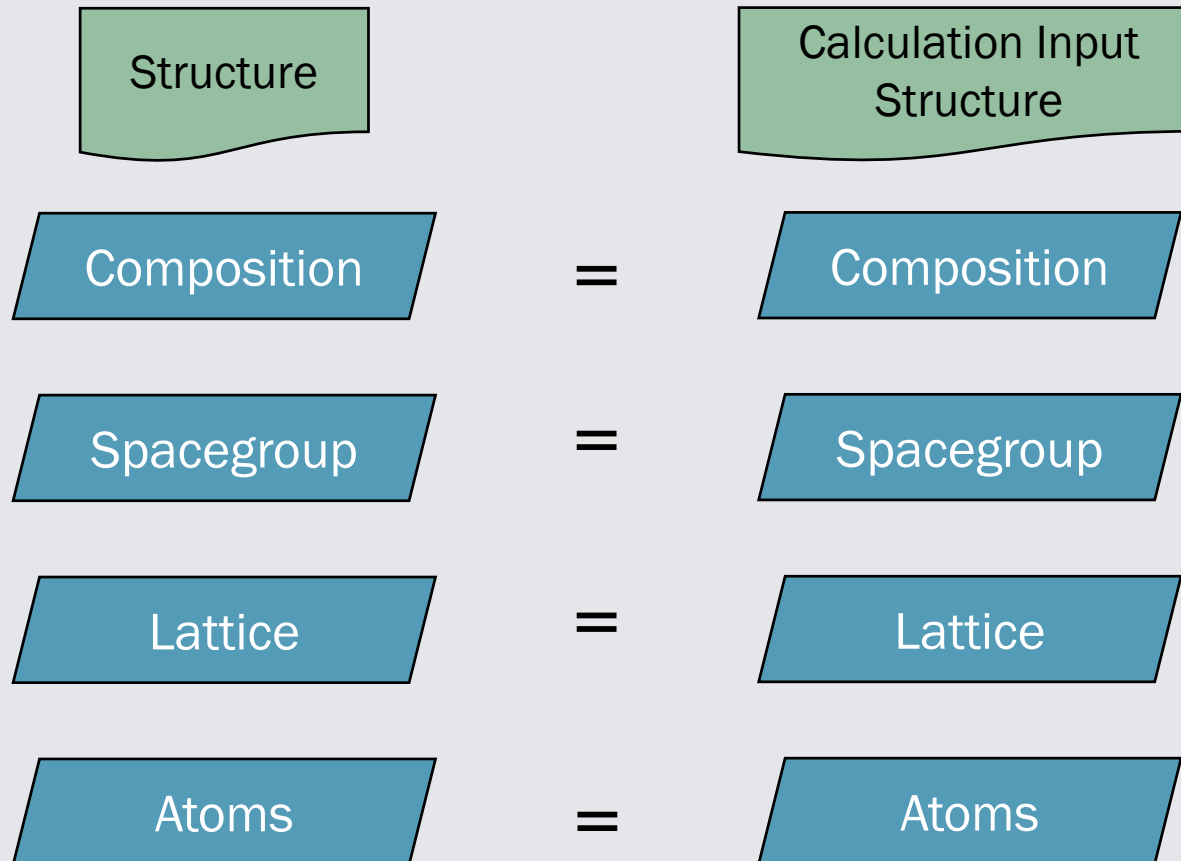
Define process metadata in one spot

Equivalent Structures

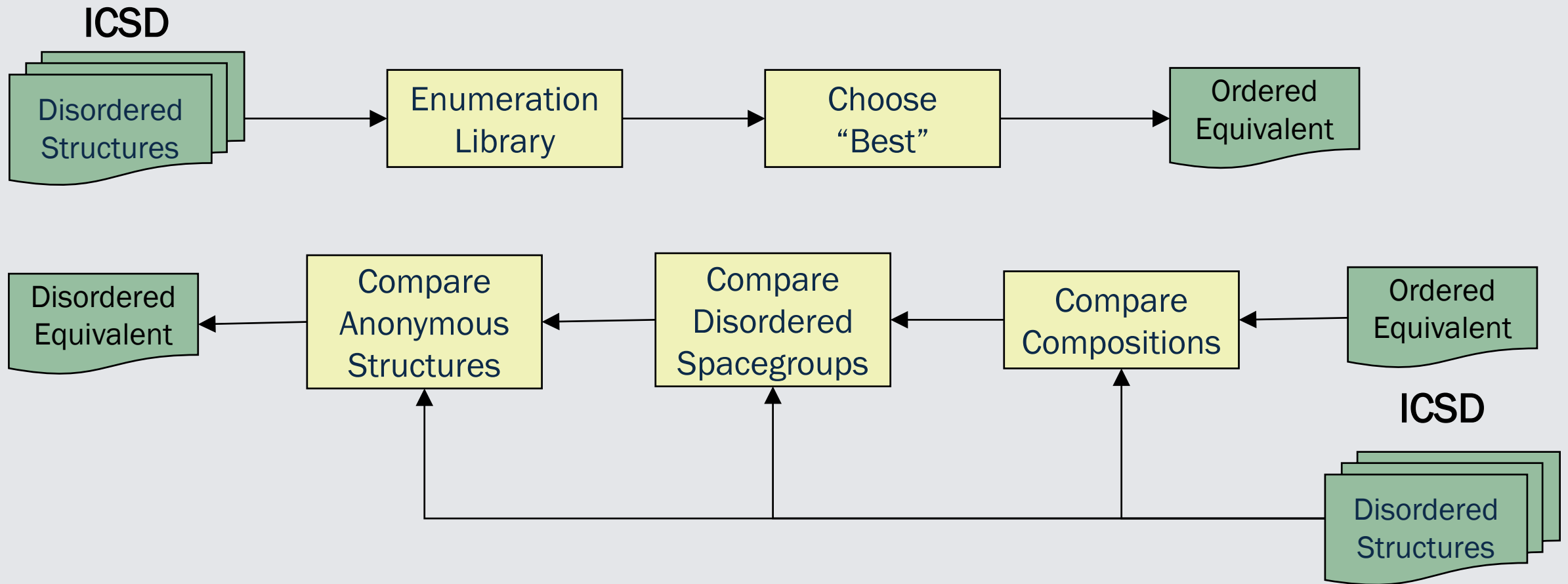


Define process metadata in one spot

Equivalent Structures

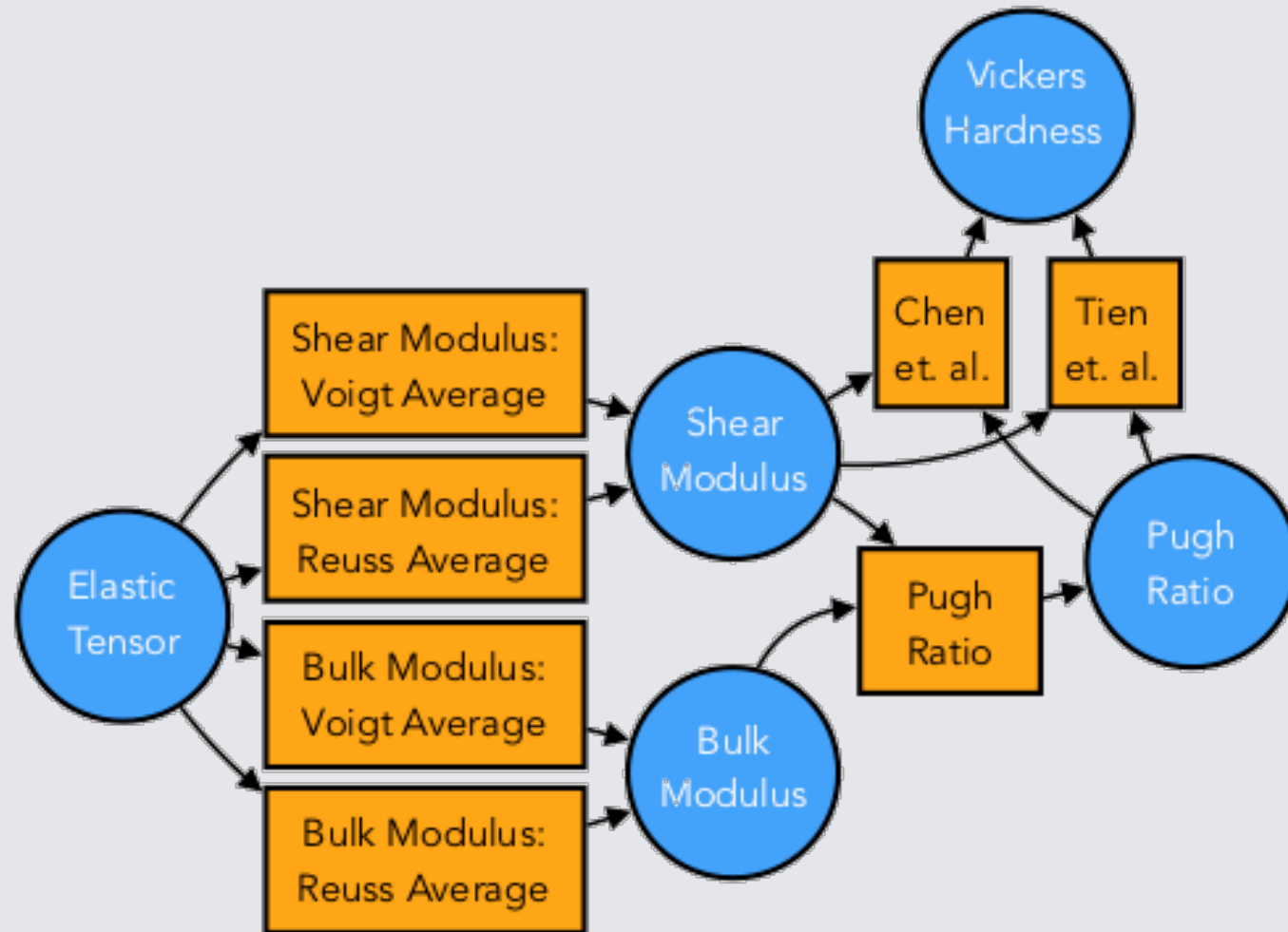


Codify human processes



Can we generalize this to more than computation?

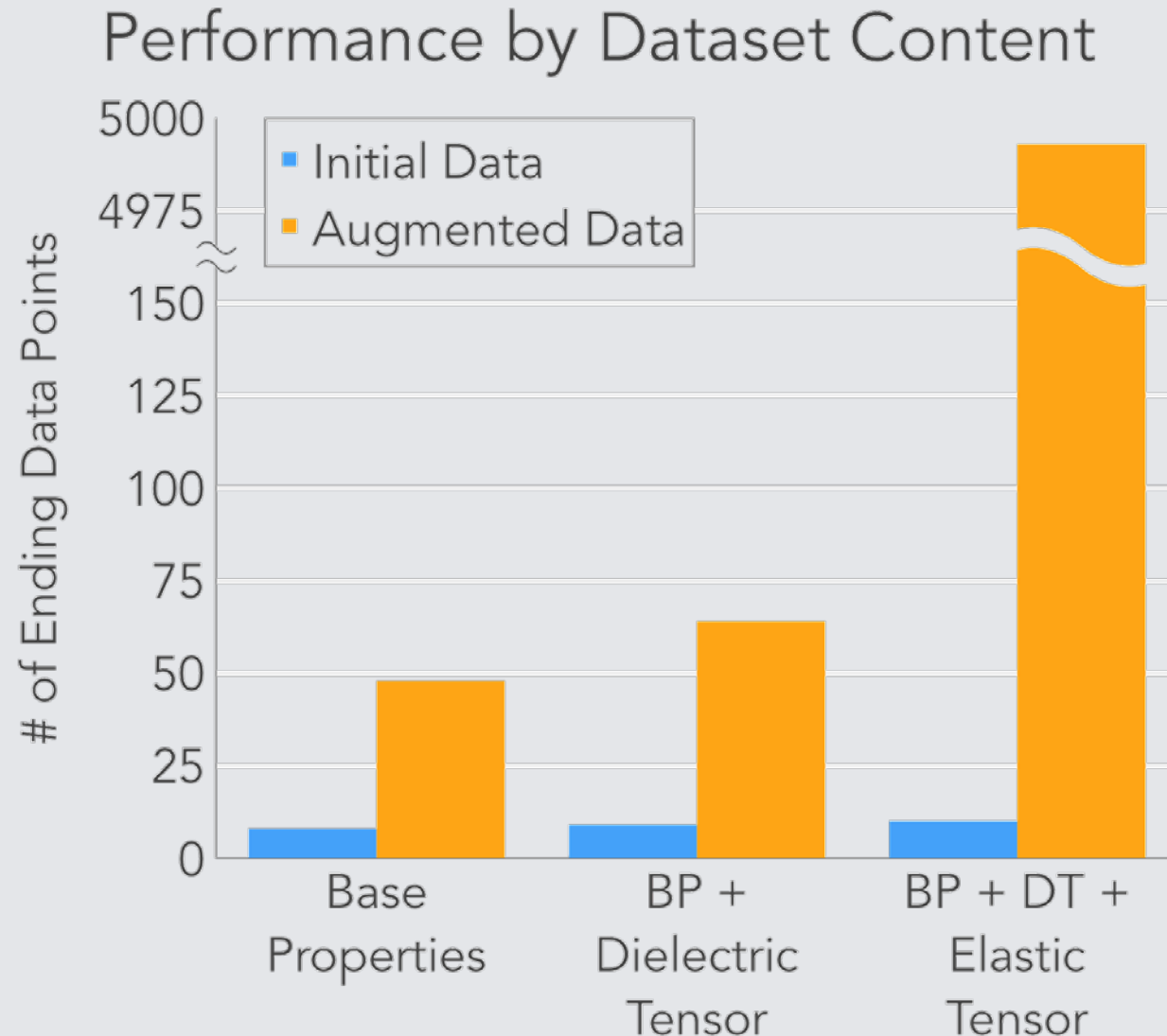
Propnet – Connecting Materials Models



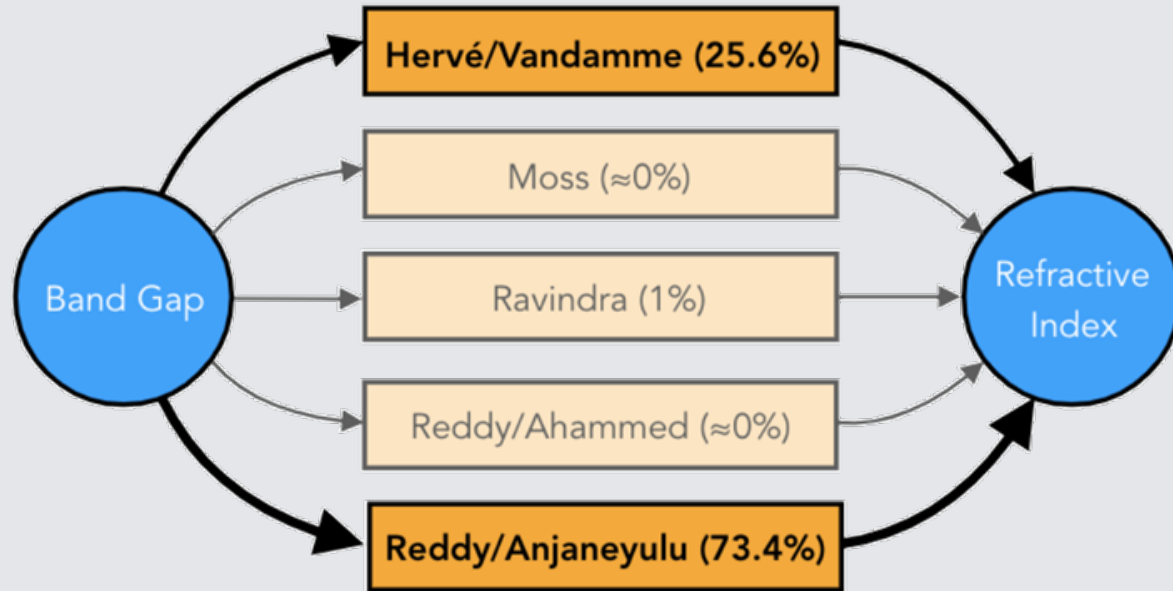
Too many models, now what?



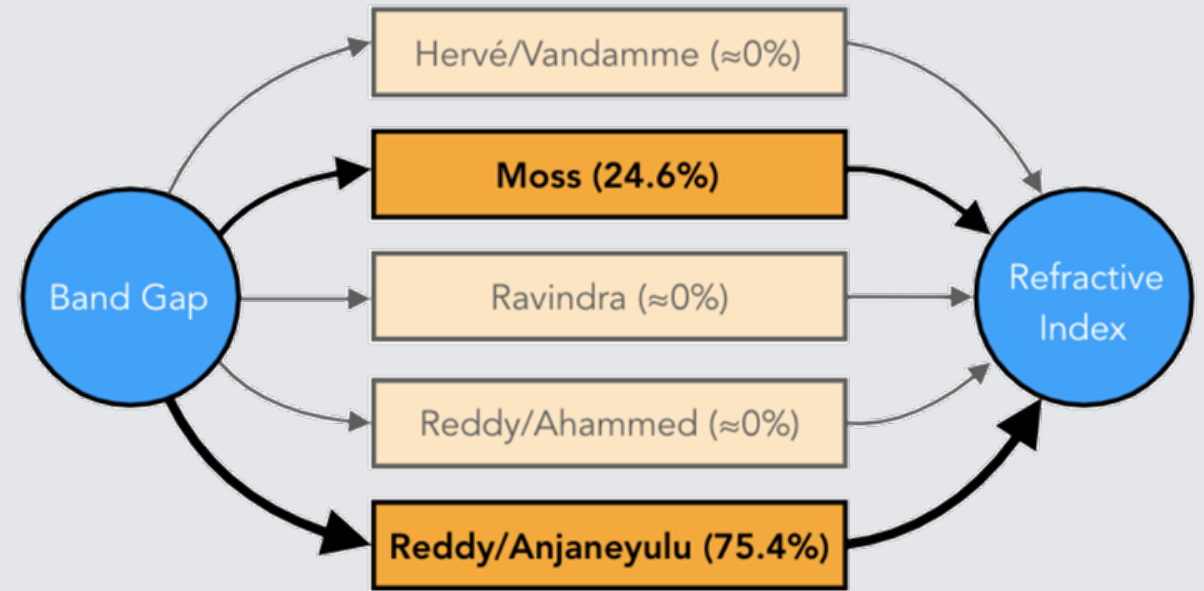
Metadata lets us augment properties



Reverse engineering metadata for experiments



(a)
Visible Light



(c)
Infrared Light

Questions?