GΦL
Global Phasing Limited

Synchrotron logistics with MXLIMS

starting with standardised experiment submission and result retrieval

Global Phasing

MXCuBE/ISPyB Meeting, Diamond, 18 November 2025

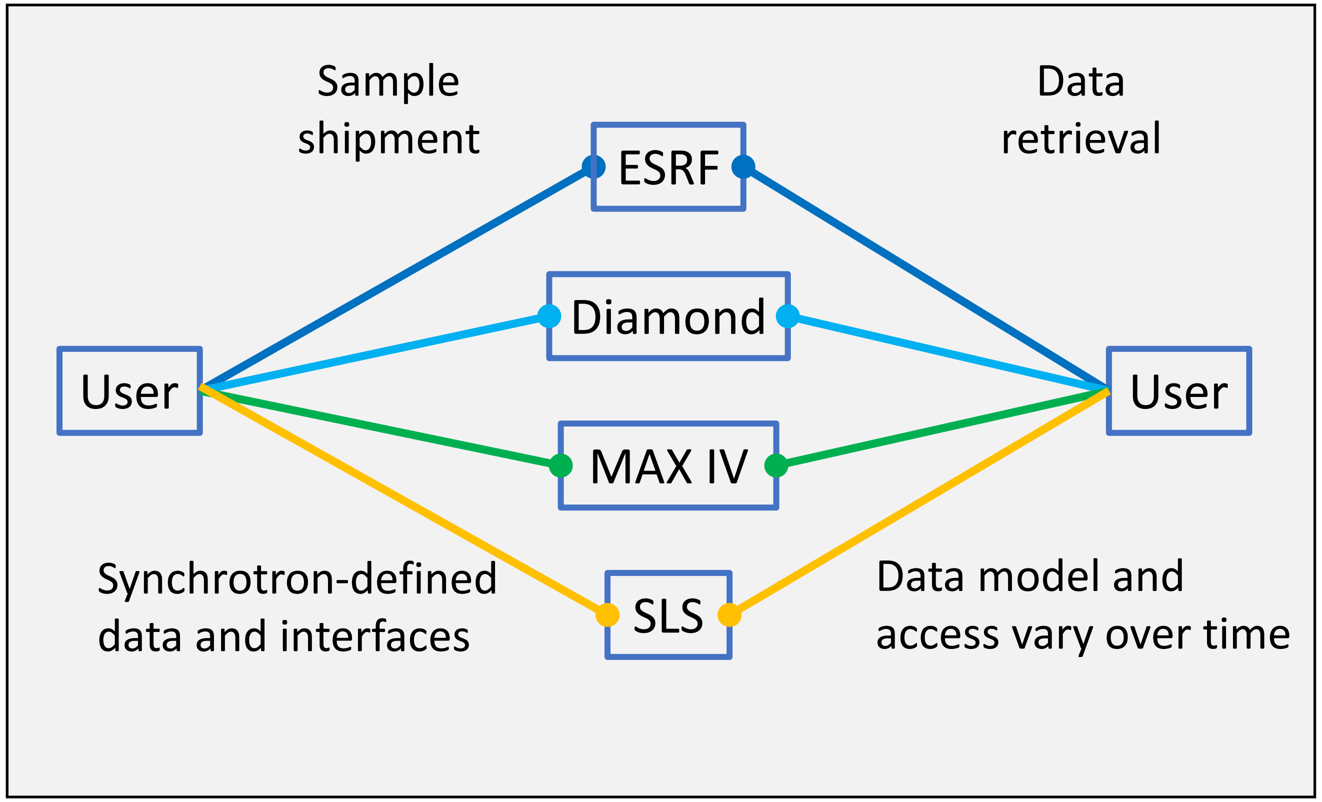
GΦL
Global Phasing Limited

- **Where we are now**
- Model, structure, and background
- Future development

GΦL
Global Phasing Limited

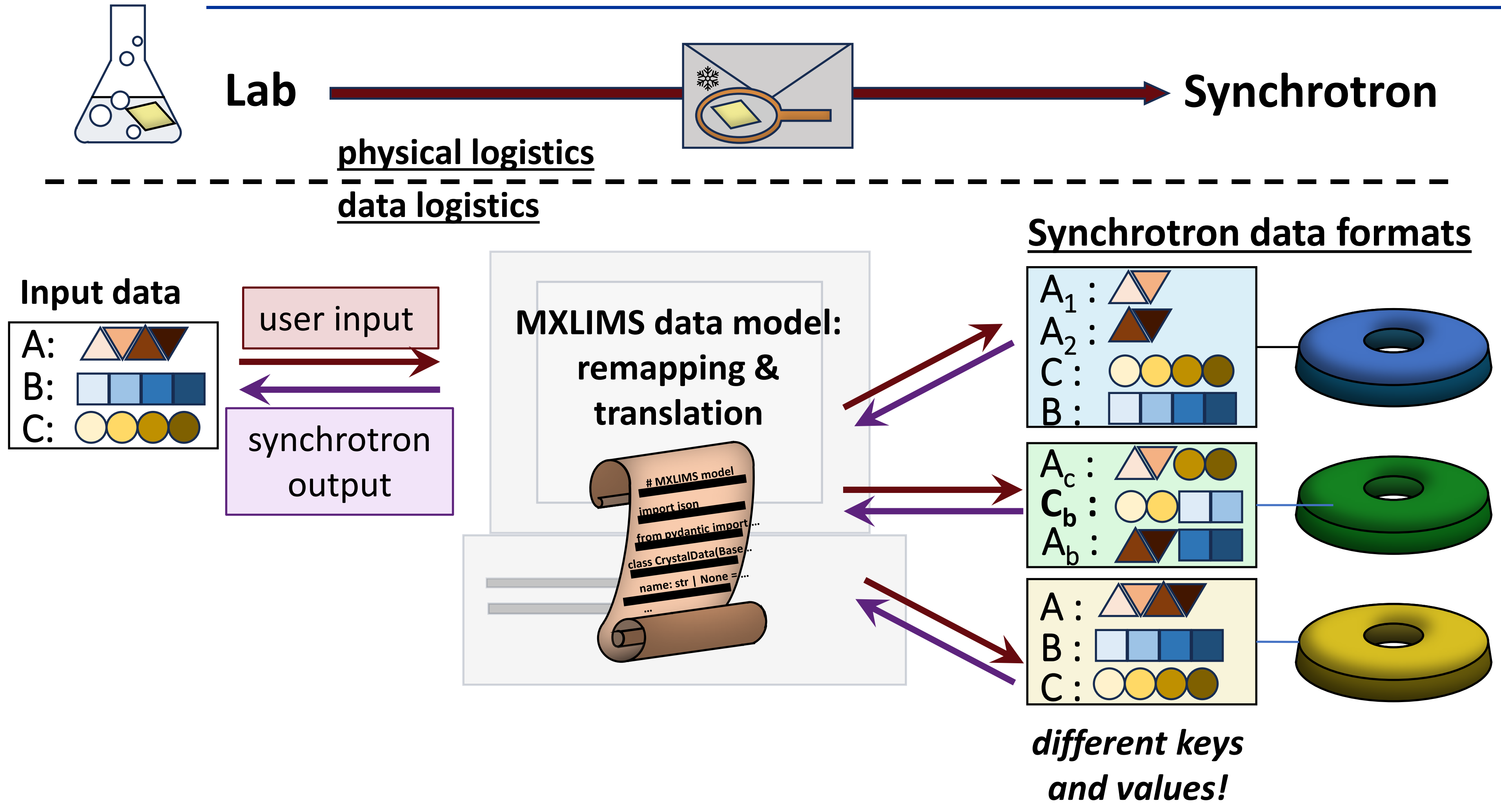
Shipping and data retrieval - now

- Separate proposals for each synchrotron
  - Different values, tag names, …
- Web site access, no software interface
  - Ad-hoc, fragile software connection
- Synchrotrons (re)define formats but users must update data I/O code
  - N users × M synchrotrons = N × M converters to maintain
- Diverse access control procedures and information requirements
  - Access controls will never be the same, but procedures might be harmonised


Sample shipment
Data retrieval
ESRF
Diamond
User
User
MAX IV
SLS
Synchrotron-defined data and interfaces
Data model and access vary over time

GΦL
Global Phasing Limited

Alternative - using MXLIMS

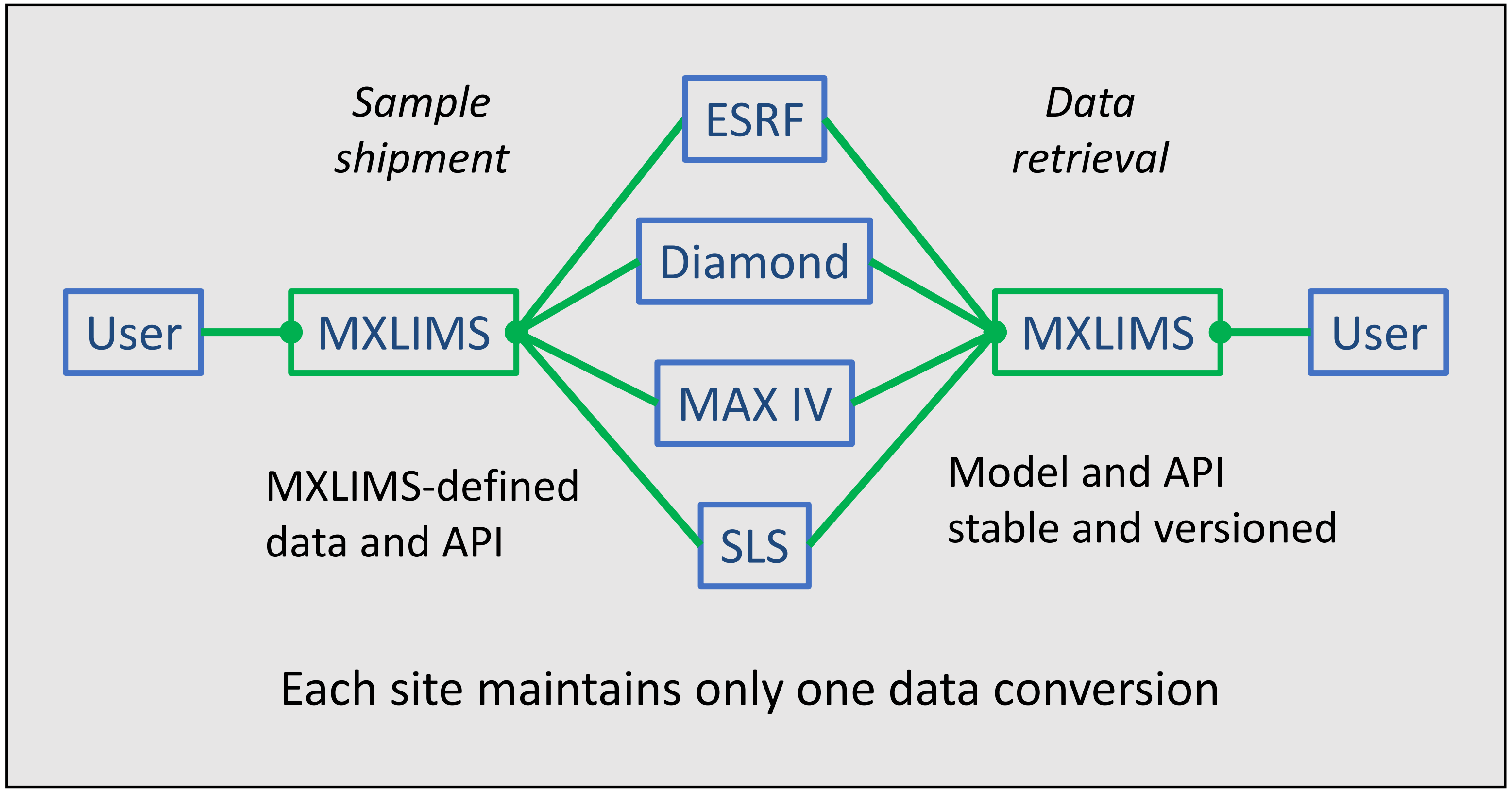


With thanks to Dennis Stegmann

GΦL
Global Phasing Limited

Harmonised data logistics

- MXLIMS exchange format and API used as stepping stone
- Users can prepare one MXLIMS document that will work anywhere
- Synchrotrons accept and export MXLIMS files, taking care of conversion to their internal formats
- Data retrieval has one metadata file (MXLIMS), even if directory structures may differ
- No requirement for using MXLIMS internally (though it remains possible)


Sample shipment
Data retrieval
User
MXLIMS
ESRF
Diamond
MAX IV
SLS
MXLIMS
User
MXLIMS-defined data and API
Model and API stable and versioned
Each site maintains only one data conversion

Global Phasing Limited

What do we need - initially ?

- A model that can store the data we need for shipping. **DONE**
- Implementation as file formats and in-memory objects (Pydantic) **DONE**
  - Includes validation through JSON schemas and Pydantic
- Mapping between the shipment-data .csv file of synchrotrons and MXLIMS **DONE**
  - Mappings set up for Diamond, ESRF, MAX IV, Petra III, SLS and SOLEIL
  - Import/export routines written by MXLIMS (mapping-driven, Python) and IceBear (PhP)
- Test data flow from user to synchrotron and back **WORK IN PROGRESS**
  - Send file, convert to synchrotron LIMS, and create MXLIMS file for result returns
  - Initial test can be for already-run data and can use conversion via .csv files
  - Test being set up at MAX IV, awaiting freeing of necessary ISPyB developer resources

Global Phasing Limited

For full integration

- Direct integration to LIMS systems via API (or file I/O)
  - Protype work done at
    - IceBear (Ed Daniel)
    - SLS (Dennis Stegmann, May Sharpe)
    - Diamond (Guilherme de Freitas)
  - Including support for model versioning
  - API access points for software-driven upload/download, separate from web interface
- Test and improvement of result upload

- Where we are now
- **Model, structure, and background**
- Future development

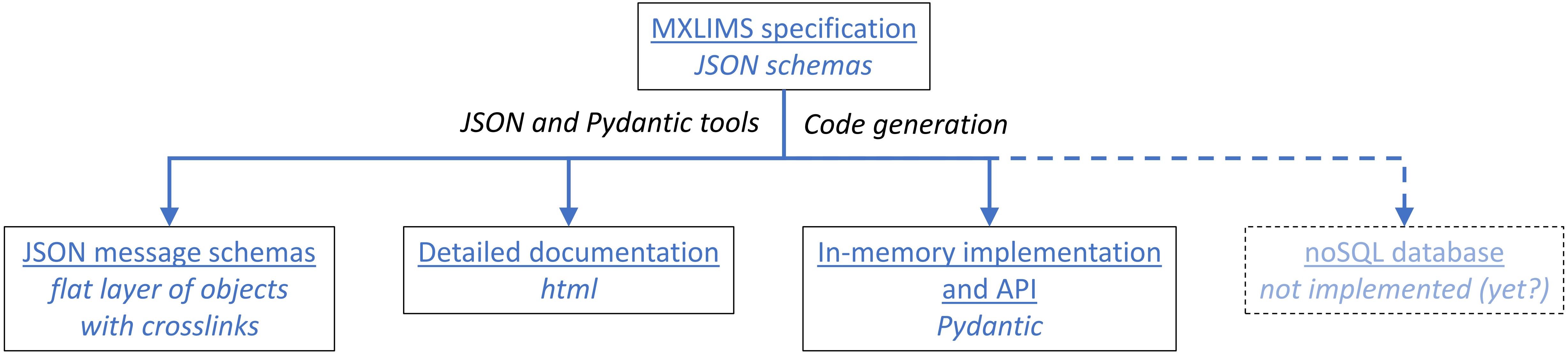
Global Phasing Limited

MXLIMS design principles

- A clearly defined standard that supports validation
  - JSON files, with JSON schema specification
- Flexibility and maintainability
  - NoSQL-database-compatible model with few main tables and varied metadata
  - Slots for site-specific data
  - versioning
- Support for provenance tracking
- Reuse metadata for multiple, similar objects (e.g. diffraction plan, results)
- Model applicable to LIMS building in addition to data exchange (if desired)

Global Phasing Limited

MXLIMS implementation

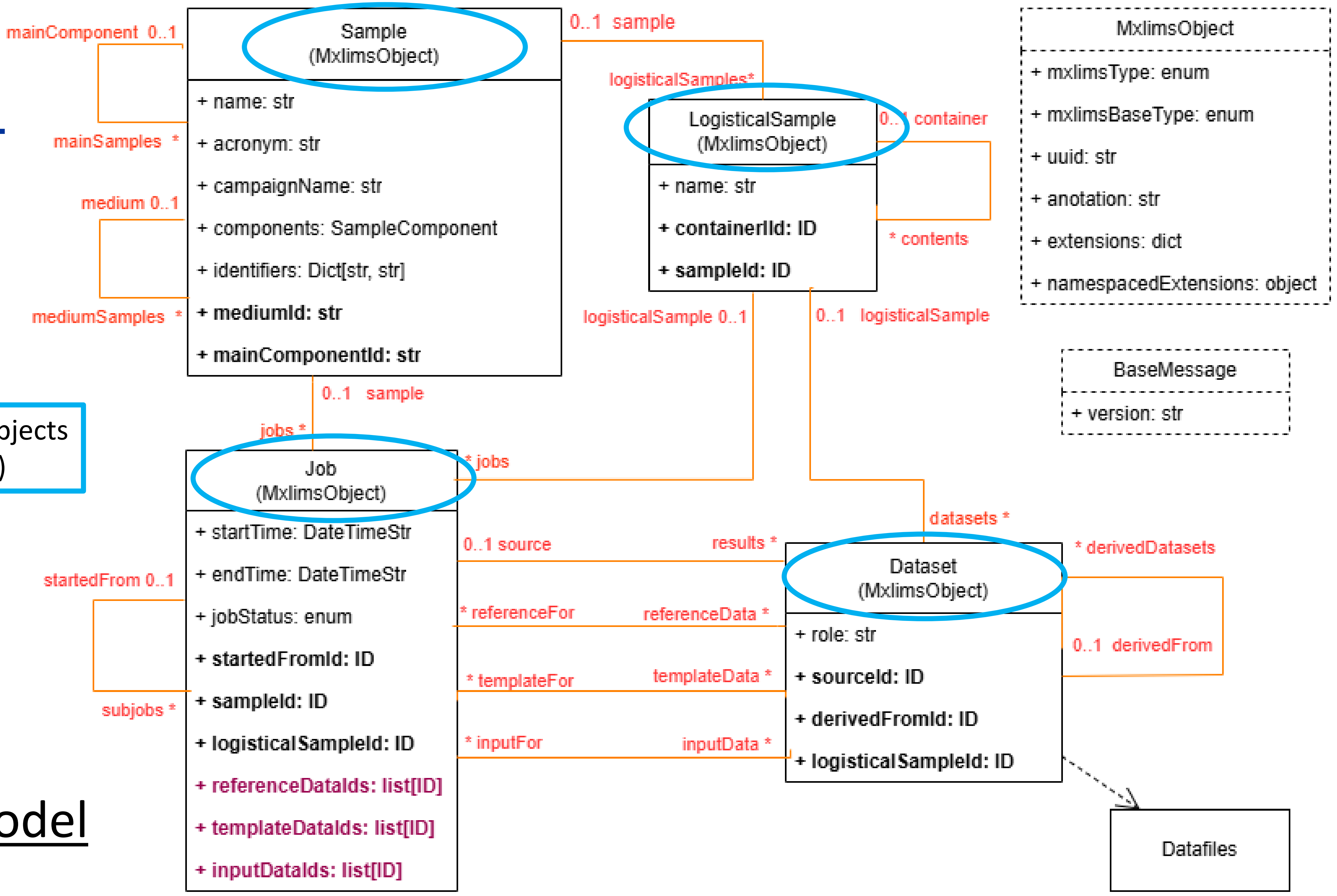


JSON supports tree structures, but objects form a web, not a tree.
Links between objects are stored as object ID pointers (foreign keys, UUIDs) also within messages.

Link handling and link type checking differs between messages, API, and different implementations.

Code generation is used to adapt the basic specification to each case.

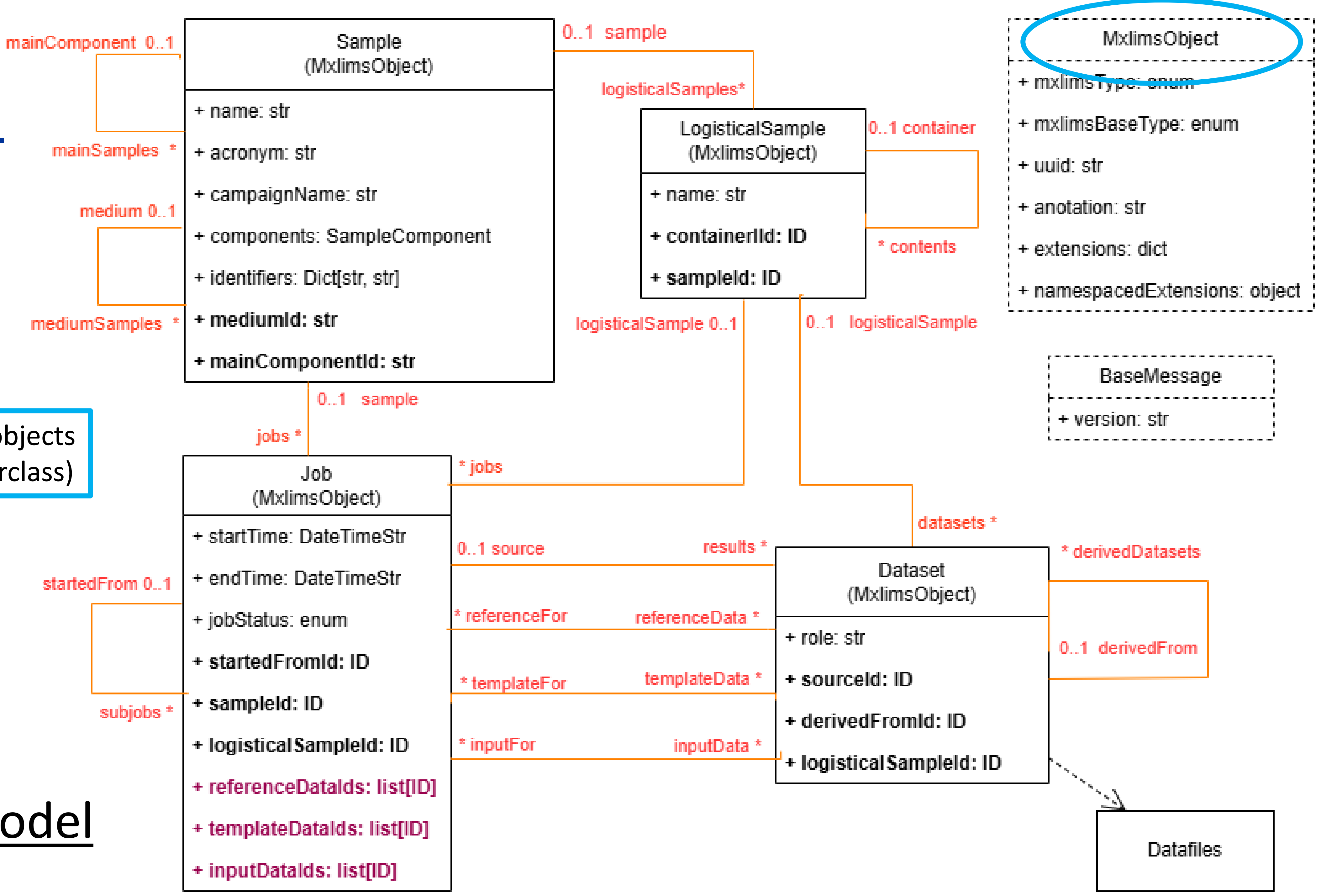
GΦL
Global Phasing Limited



Four kinds of objects
('noSQL tables')

Core Model

GΦL
Global Phasing Limited



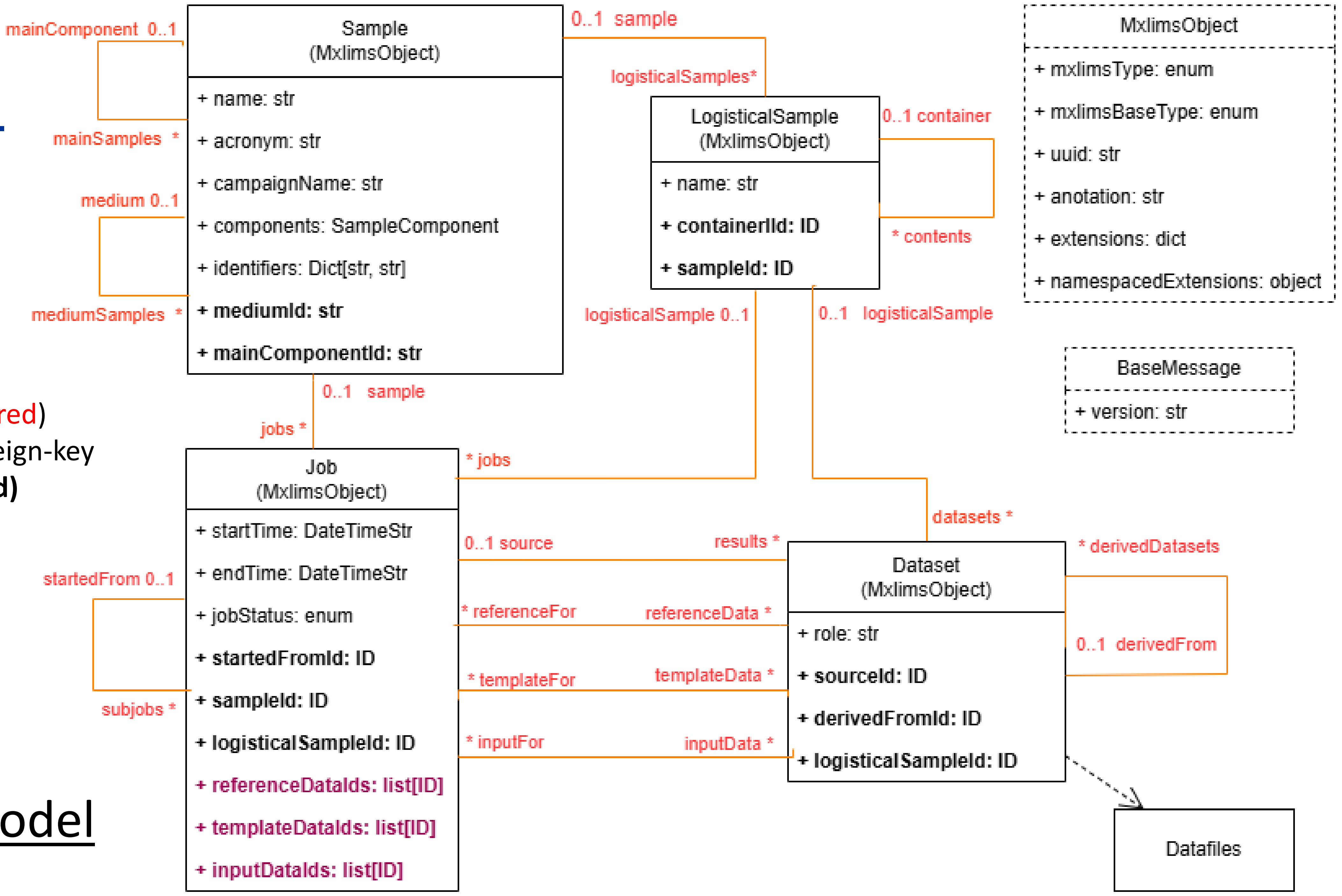
Shared by all objects
(abstract superclass)

Core Model

Global Phasing Limited

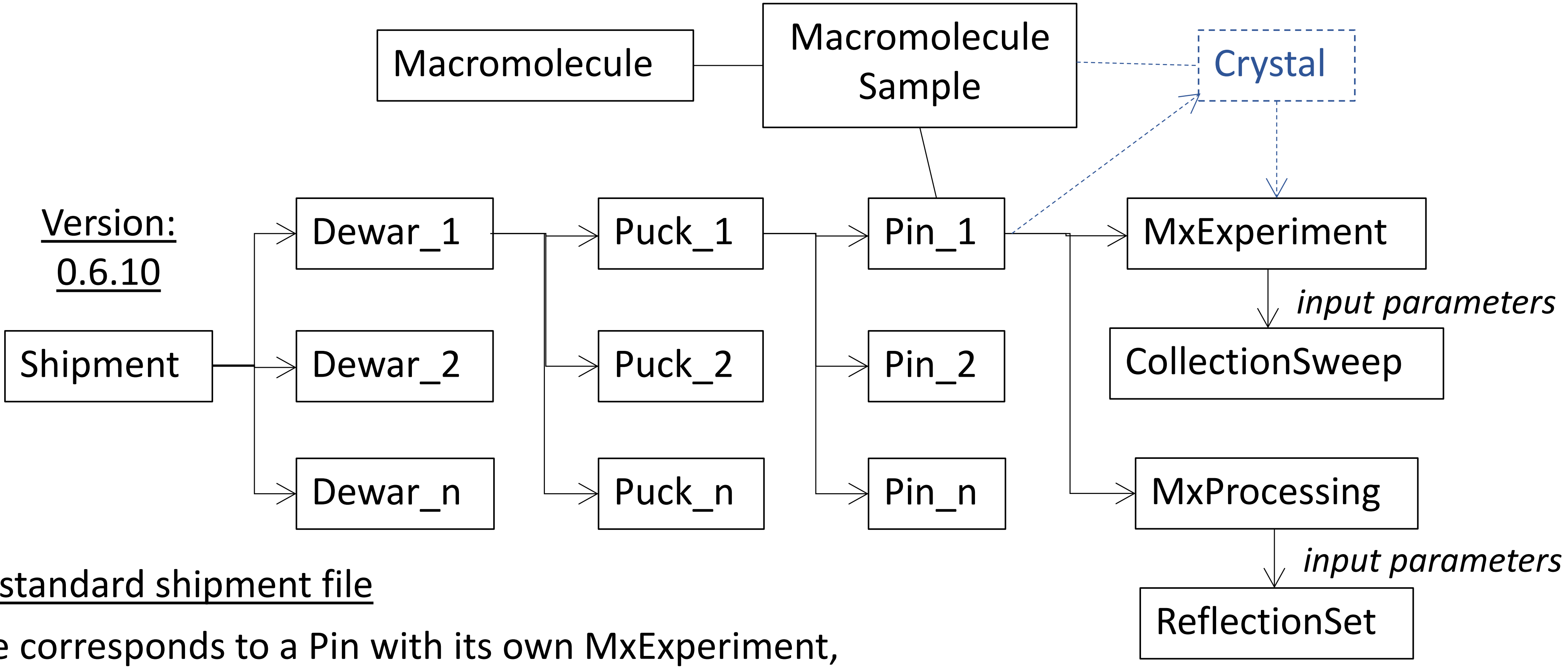
API-only links (in red) supported by foreign-key attributes **(in bold)**

Core Model



Global Phasing Limited

MXLIMS model - Dewar shipment



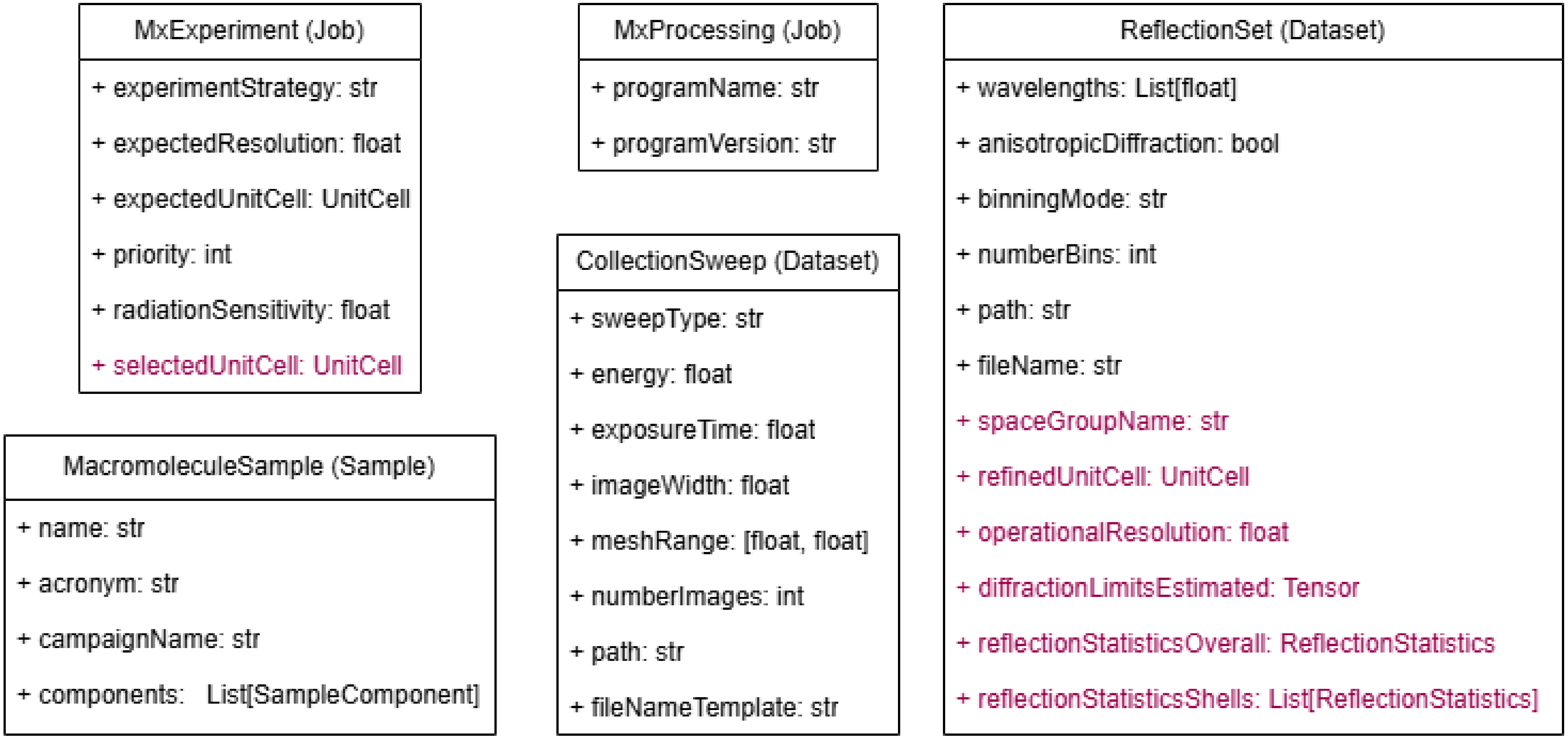
Object structure for standard shipment file

Each line in a .csv file corresponds to a Pin with its own MxExperiment, MxProcessing, and template CollectionSweep and ReflectionSet

Crystal objects are only generated during the experiment to distinguish multiple crystals

Global Phasing Limited

Model data - examples



MxExperiment (Job)

+ experimentStrategy: str
+ expectedResolution: float
+ expectedUnitCell: UnitCell
+ priority: int
+ radiationSensitivity: float
+ selectedUnitCell: UnitCell

MacromoleculeSample (Sample)

+ name: str
+ acronym: str
+ campaignName: str
+ components: List[SampleComponent]

MxProcessing (Job)

+ programName: str
+ programVersion: str

CollectionSweep (Dataset)

+ sweepType: str
+ energy: float
+ exposureTime: float
+ imageWidth: float
+ meshRange: [float, float]
+ numberImages: int
+ path: str
+ fileNameTemplate: str

ReflectionSet (Dataset)

+ wavelengths: List[float]
+ anisotropicDiffraction: bool
+ binningMode: str
+ numberBins: int
+ path: str
+ fileName: str
+ spaceGroupName: str
+ refinedUnitCell: UnitCell
+ operationalResolution: float
+ diffractionLimitsEstimated: Tensor
+ reflectionStatisticsOverall: ReflectionStatistics
+ reflectionStatisticsShells: List[ReflectionStatistics]

GΦL

Global Phasing Limited

- Where we are now
- Model, structure, and background
- **Future development**

Global Phasing Limited

Ready for full integration

- Direct integration to LIMS systems via API (or file I/O)
  - Protype work done at
    - IceBear (Ed Daniel)
    - SLS (Dennis Stegmann, May Sharpe)
    - Diamond (Guilherme de Freitas)
  - Including support for model versioning
  - API access points for software-driven upload/download, separate from web interface
- Test and improvement of result upload

Global Phasing Limited

Immediate improvements

- Integration also for richer metadata than the .csv files
  - Already supported in MXLIMS model
- Harmonisation and end points for access tokens, requirements etc.
  - More in next talk, by Guilherme de Freitas
- Improved in-memory implementation (Pydantic)?
- Test and extend MX model
  - Plate samples, multipins, **DONE**
  - X-ray centring **DONE**
  - Results and processing

Global Phasing Limited

Possible later expansion

- Expand coverage of model
  - More processing programs? Maps? Structure files?
  - Sample contents, production and crystallisation?
  - Serial Crystallography samples and methods ???
- noSQL database implementation???

Global Phasing Limited

Getting organised

- MXLIMS is not a private project, but a nascent community standard

  As new participants join, the project will be reorganised:
  - Shared, public repositories
  - Procedures for accepting changes, making new versions etc.
  - Governance by stakeholders

- MXLIMS is ready for you to join in!

Global Phasing Limited

Acknowledgements

- Thanks to the many, many people who have participated in the MXLIMS working group, the MXCuBE automation WG, or contributed information, discussions and feedback.
- For work on integration, synchrotron mapping and testing and making the MXLIMS model, particular thanks to Guilherme de Freitas and Karl Levik (Diamond), Daniele de Sanctis, Didier Nurizzo, Marcus Oscarsson, Olof Svensson, and Yan Walesch (ESRF), Leigh Carter, Clemens Vonrhein and Peter Keller (Global Phasing), Ed Daniel (IceBear), Ezequiel Panepucci (MAX IV), Dennis Stegmann and May Sharpe (SLS), and Martin Savko (SOLEIL)