ODYSSEUS DATA SERVICES

THE HEALTHCARE DATA EXPERTS

ODYSSEUS
DATA SERVICES INC

MIMIC IV to OMOP

Scope & Approach

Agenda

1. MIMIC III – IV differences
2. Environment
3. Waveforms
4. Focus / Interests
5. Documentation

MIMIC III - IV

Tables and fields

Core
Hosp
ICU
ED (*)
CXR (*reports , ?images)

*optional, maybe community effort

Environment

GCP

Managed by PhysioNet

BigQuery

Waveforms

Continuous measurements

Blood pressure, Heart rate, Saturation

Vector data

EKG

Waveform challenges

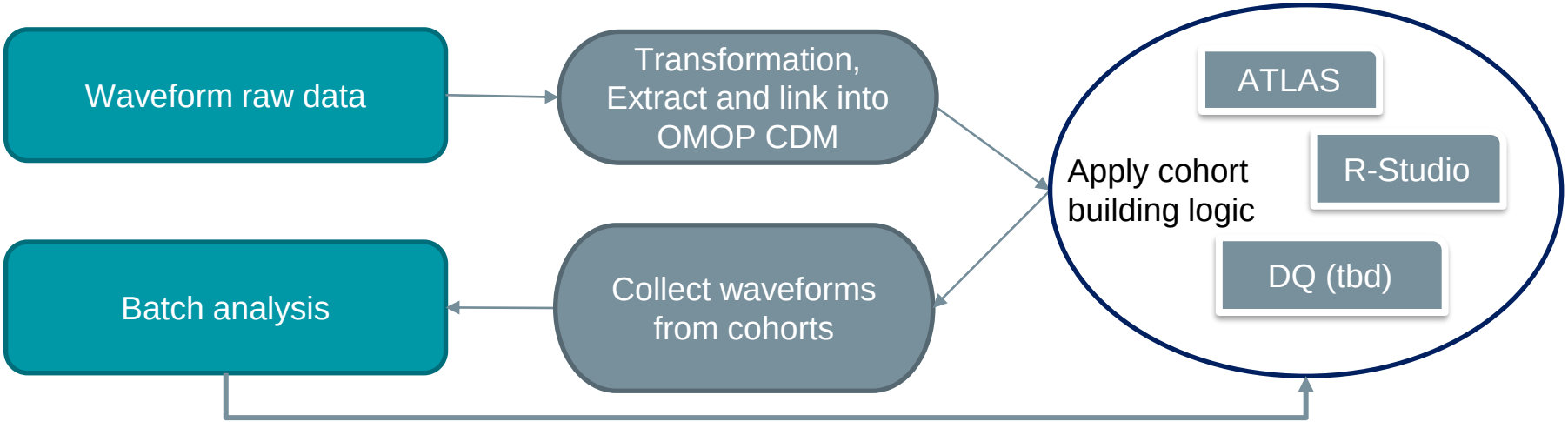
Technical challenge

- storage, extract metadata, link to findings

Use case / application challenge

- cohort building based on patient clinical data and data derived from waveforms

Waveform process



Waveform POC model

- Respiration Rate & EKG => check against MIMIC logic

Set of Analysis procedures / tools

- Apply Neural Network / ML

Waveform formats

“Waveform Database” (WFDB)

“Medical waveform Format Encoding Rules” (MFER)
(ISO/TS 11073-92001, dating from 2003)

…

Actual Waveform data in MIMIC

EKG

Bedside monitor data stream:

Respiration Rate

No ventilator data (?)

…

Waveform analysis

Machine Learning

Multimodal analysis

Tools for batch analysis / Applicable Formats ?

Cataloguing ways to derive knowledge / features of interest from waveform data…

Waveform metadata and findings extraction

Link metadata and findings to waveform

- retrieved from device while capturing
- created by Machine Learning algorithm
- added during ETL

Use cases…

- Predictive analytics
- Train algorithms (e.g. for alerts)

- **First : Find measurement (QT time) and drug exposure to build cohorts with linked waveforms**

Waveform storage and linking

Waveform capturing event

- Capturing time and duration
- Base entry to link to waveform file location
- Attributes of waveform linked to event

Visit Occurrence ID

- Visit table / Visit details
- Notes table (link to waveform storage?)
- Measurements
- Conditions / Clinical findings
- New Type concept (waveform derived)?

Visit detail approach

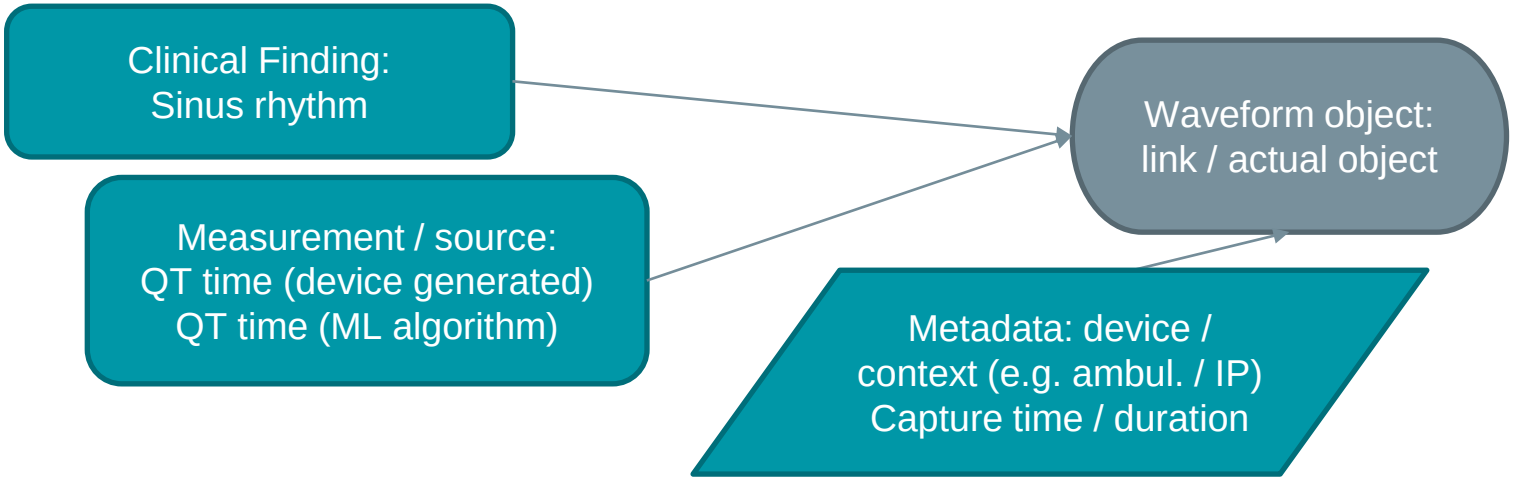
Waveform capturing event as visit_detail

- visit_detail_source_value to store the link
- Record start and end time

- Determine matching visit occurrence from waveform metadata

- Create all other depending entries with link to visit_detail

Waveform relationships

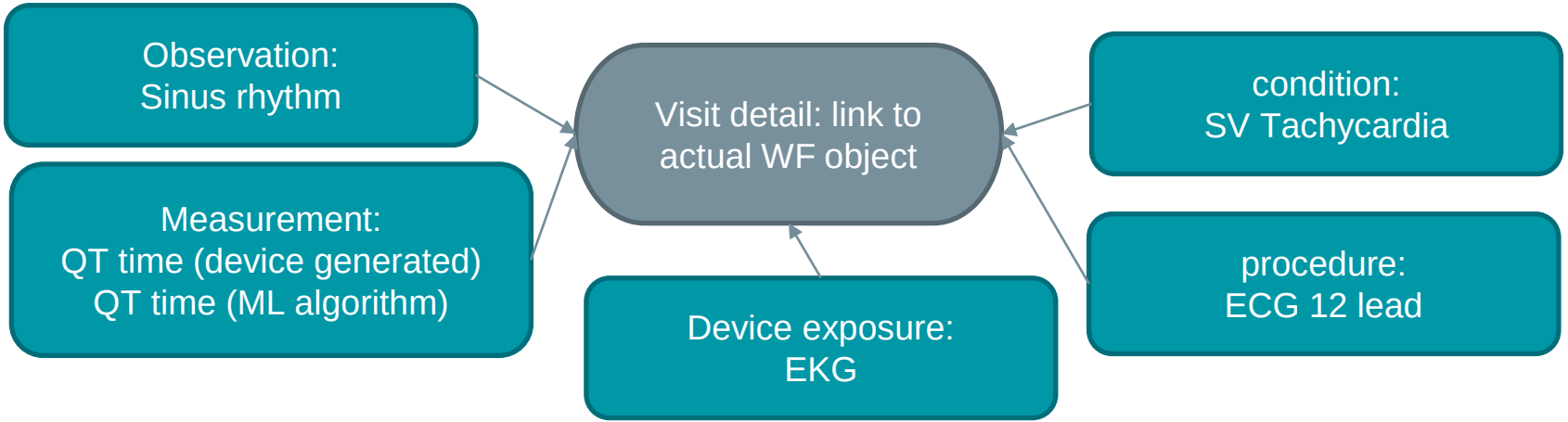


Acquire / extract additional data

- Store with relation to waveform object
- Extract more information subsequently (Crawler?)

…

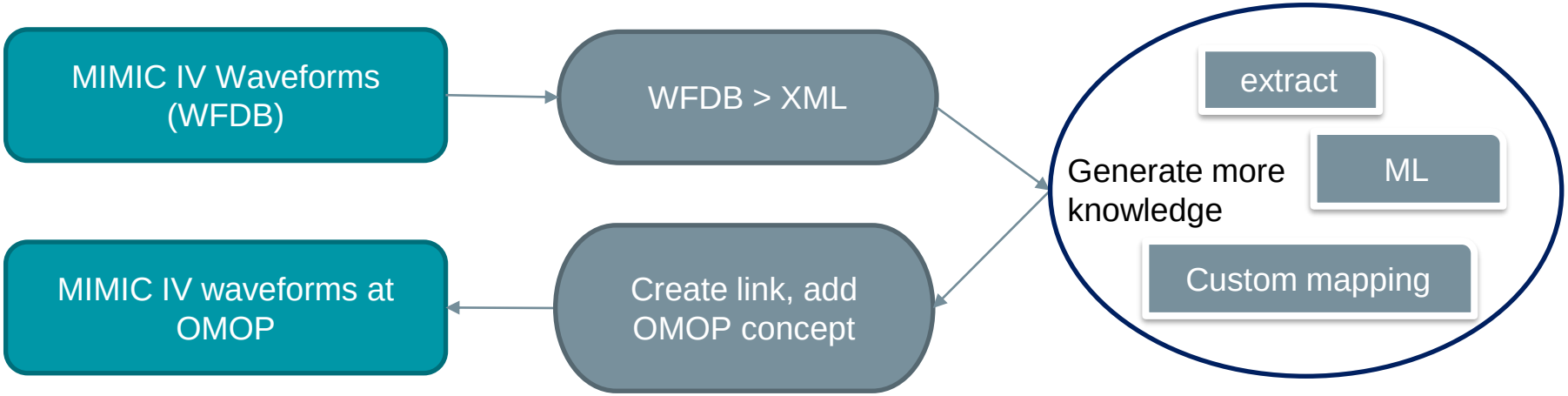
Waveform representation



Visit detail as central node for waveforms

- Identify visit associated with waveform
- Create visit detail entry with waveform storage link
- Parse extracted information in xml and create respective OMOP CDM entries

Waveform handling



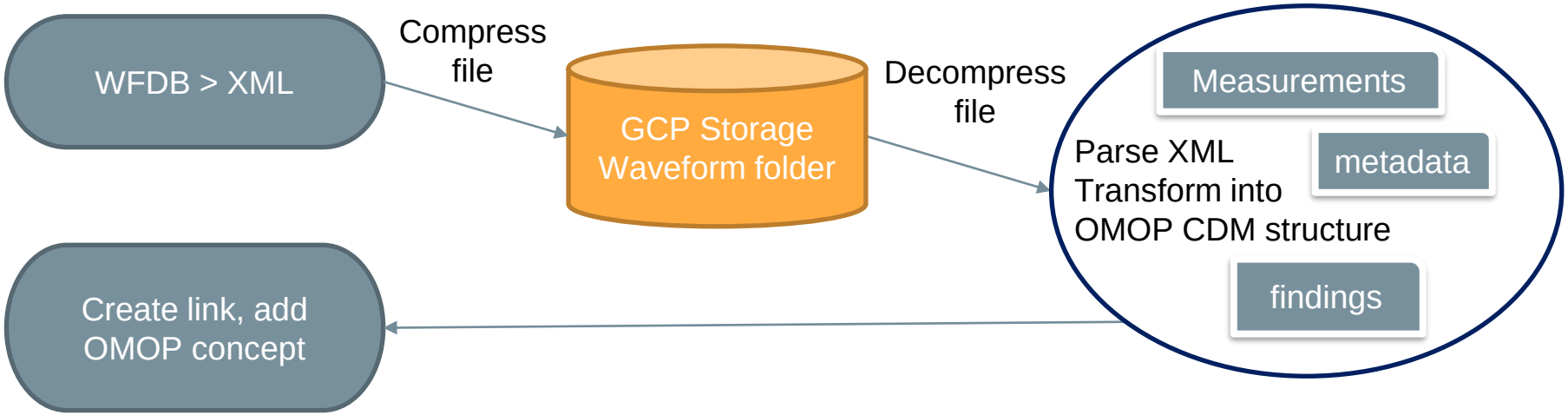
Development environment

- Sample set

Final environment

- Full duplication?

Waveform handling



Automated process - conversion

- Extract meaningful information
- Convert wfdb format into XML
- Compress file
- Save to storage location

Automated process - parsing

- Retrieve files by patient ID
- Decompress file
- Parse XML
- Transform into OMOP concepts

Project steps

Create github repository as collaboration space

- https://github.com/OHDSI/MIMIC => provide github user names

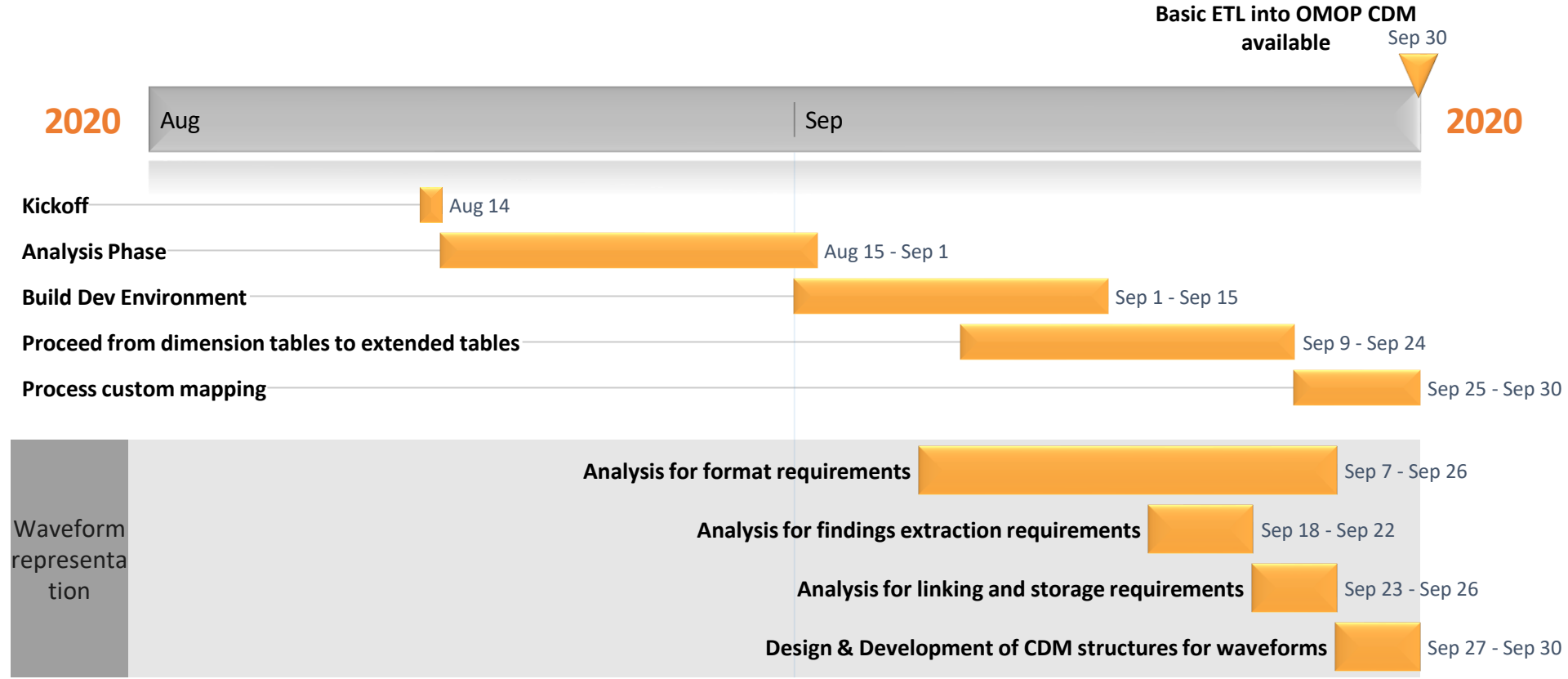
Assemble Odysseus team – do PhysioNet Training

- Anna Tsetkova – ETL development / Architecture
- Dmitry Dymshyts – Medical Specialist
- Michael Kallfelz – Project Manager / Analyst

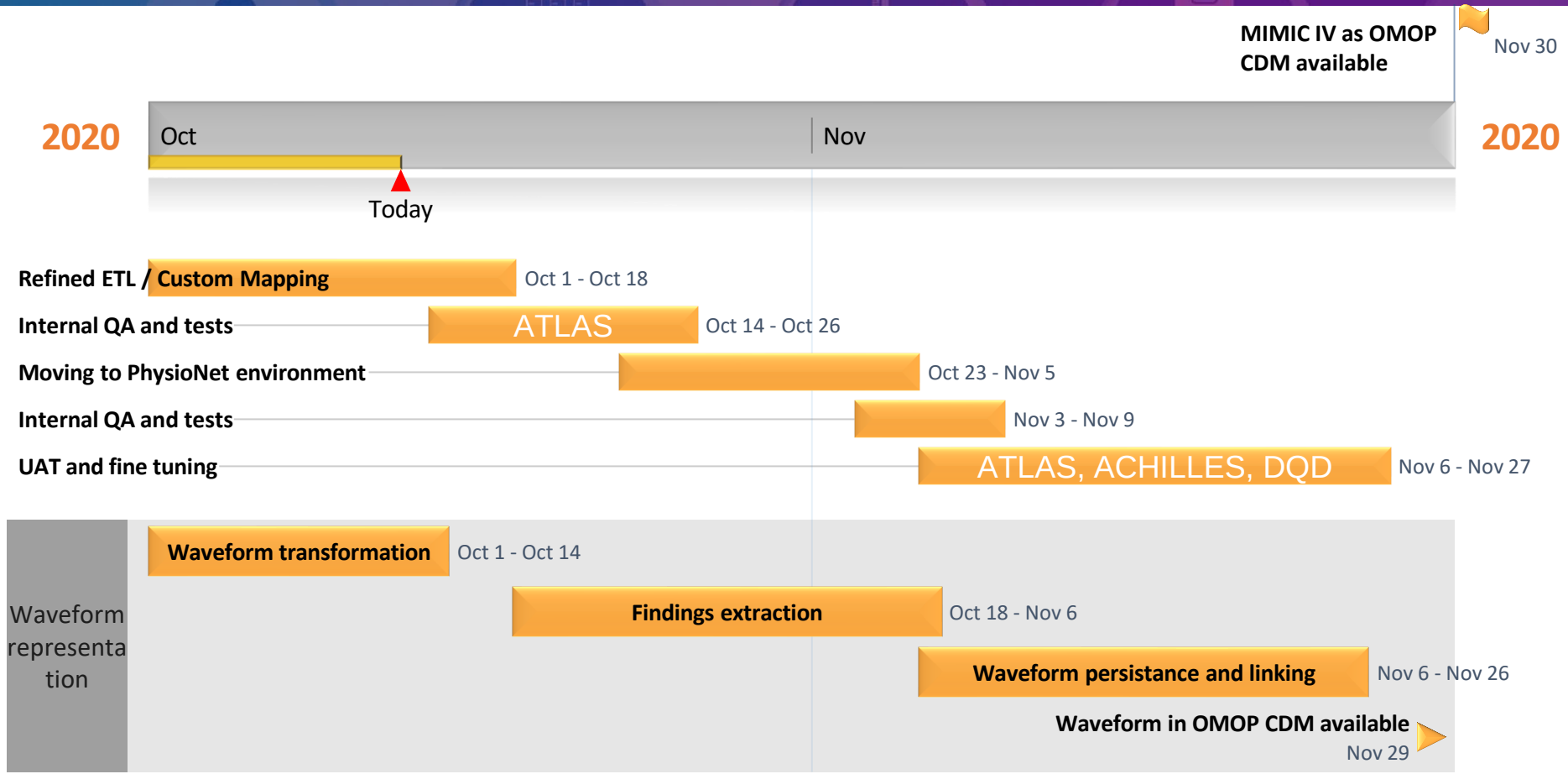
Work packages

- Define Architecture for GCP Environment – restrictions / opportunities
- Delta MIMIC III > IV
- Determine Custom Mapping needs
- Agree on Waveform representation

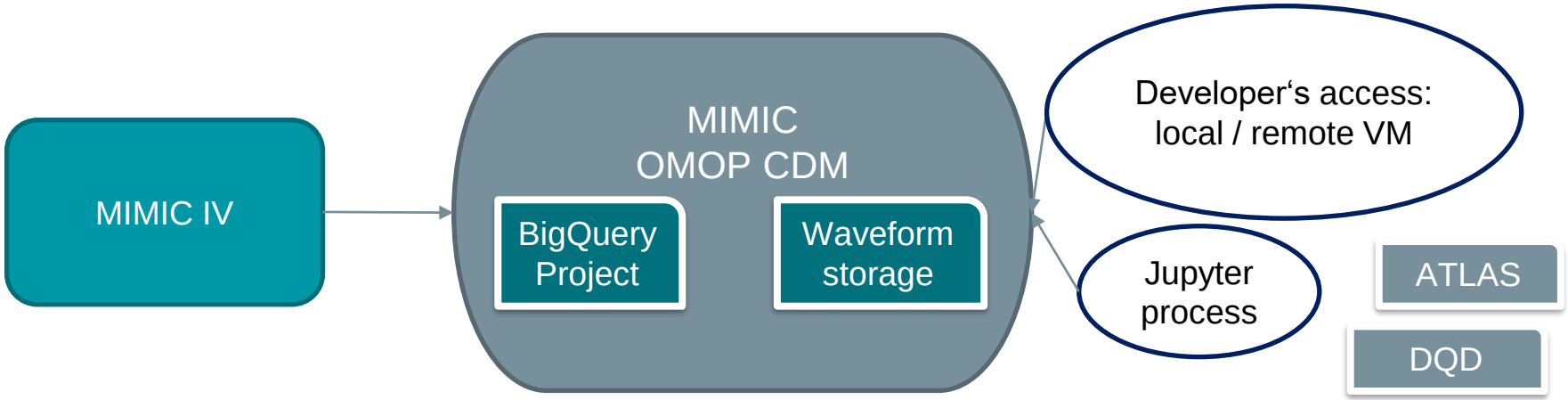
Project steps – first part



Project steps – second part



Odysseus GCP environment



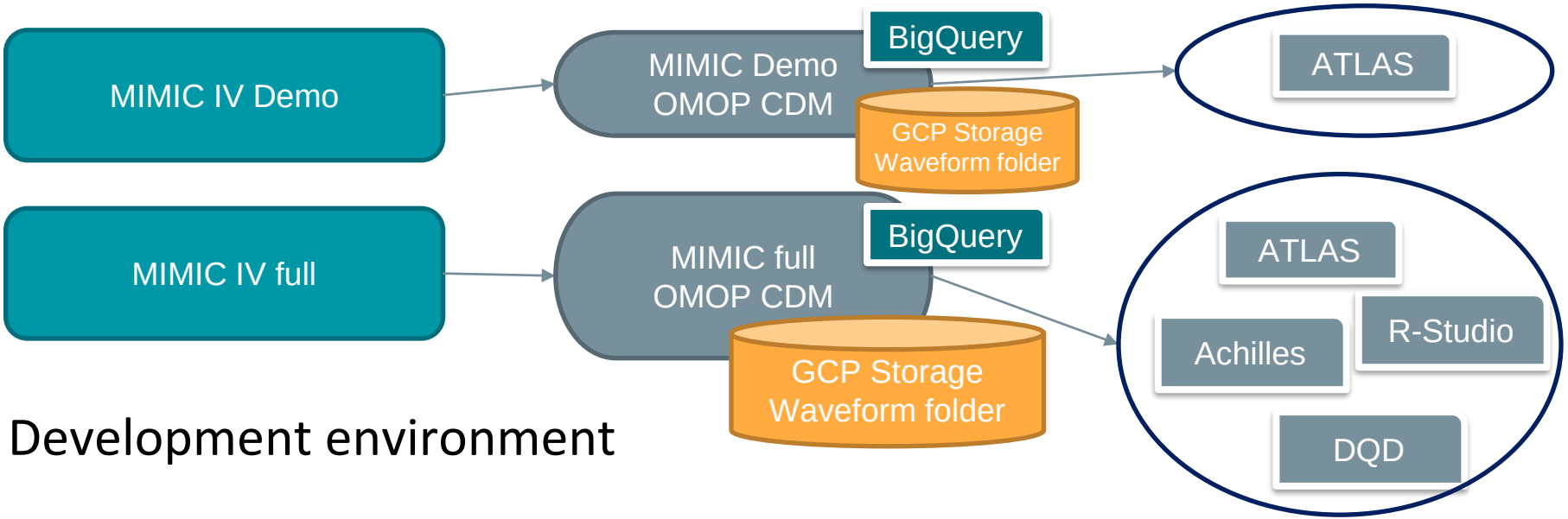
Development environment

- Project / Billing ID set up separately by Odysseus
- BigQuery, access to PhysioNet MIMIC IV BQ instance
- Python Scripts for ETL

optional

- DataFlow: integration with Jupyter / Visualization
- Jupyter process for reporting

Possible OMOP CDM @ Physionet GCP



Development environment

Final environment

- Transfer of finalized OMOP CDM and ETL logic to PhysioNet
- OHDSI tools like ATLAS => automated **authentication** against PhysioNet user (google ID)