Overview of TA5 work plan and upcoming deliverables

PUNCH4NFDI TA5 – WP3 Workshop

DESY-Zeuthen 18/01/2024

A. Redelbach

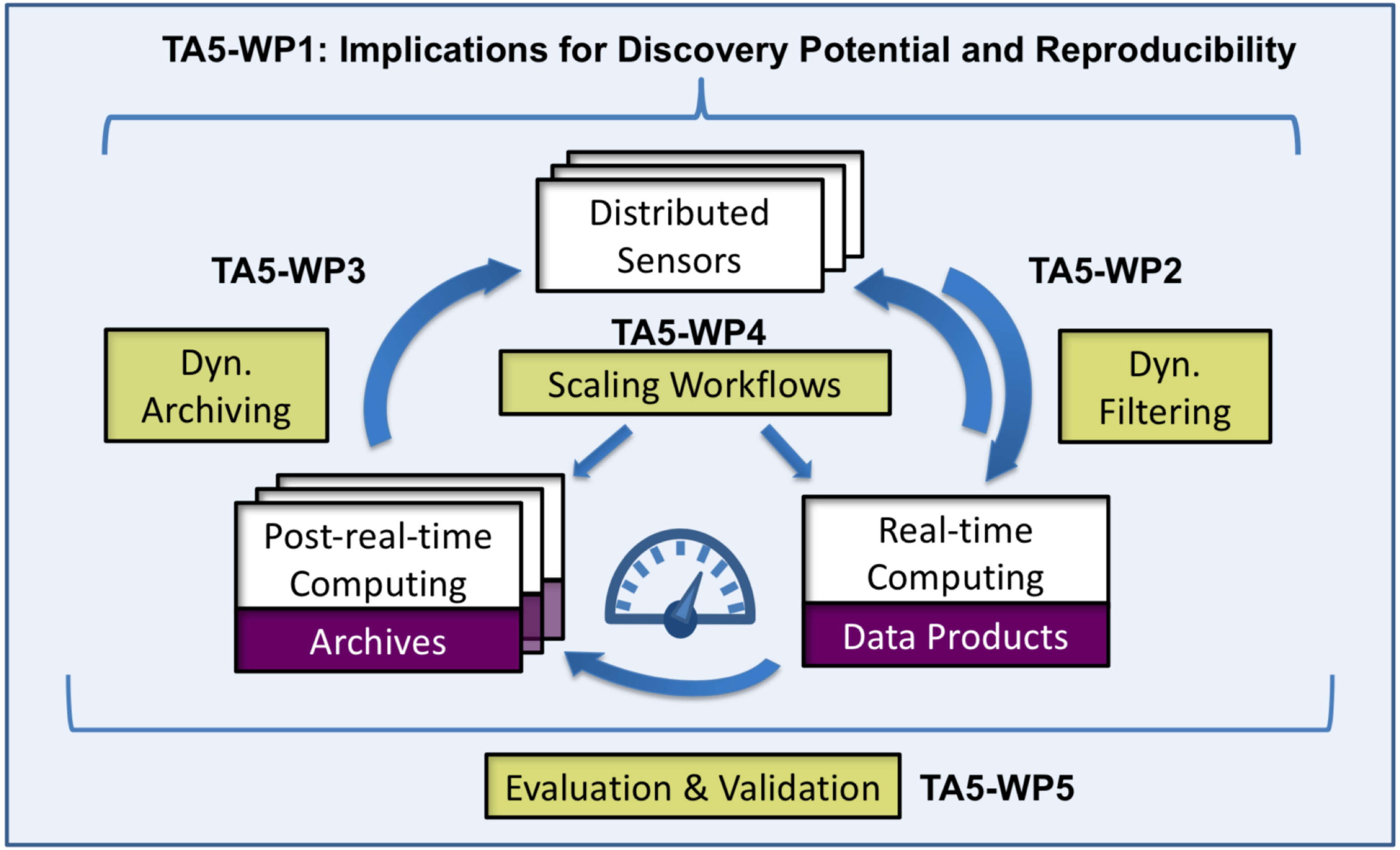
FIAS Frankfurt Institute for Advanced Studies

Overview

- Deliverables documented – just briefly listing/structuring them here
- Update of relevant workflows in connection to use cases
- Some achievements related to TA5
- Some ideas for this workshop

Overview of TA5

Aerial perspective



Dynamic Life Cycle Model

TA5 deliverables

WP1-1 (Q4/2023): Report on impact of on-line filtering on discovery potential
WP1-2 (Q3/2025): Report on impact of on-line filtering on FAIR principles
WP1-3 (Q3/2026): Concepts towards a general protocol on capturing the decisions made by, and status of, real-time sensors, as a basis for a future demonstator

WP4-1 (Q4/2024): Porting common off-line packages to a memory-based computing prototype
WP4-2 (Q1/2025): Standard software (e.g. CASA) compatible with Gen–Z.
WP4-3 (Q2/2025): Caching strategies for processing of benchmark files
WP4-4 (Q2/2026): Definition and initial implementation of an efficient real-time data processing framework
WP4-5 (30 Sep 2026): Scaled feedback interfaces between off-line software and real-time processes using MeerKAT data

WP3-1 (Q2/2023): Specifying the concept of a dyn. archive
WP3-2 (Q1/2025): Present a framework in which queries to dyn. archives can be transformed into a dyn. filter (as used by sensors), and vice versa
WP3-3 (Q3/2026): Present methods by which queries to dyn. archives also return an estimate on potential of information loss

WP2-1 (Q4/2022): Curation & metadata schemes for dynamic filtering
WP2-2 (Q1/2023): Strategy concept for identifying highly complex (multi-parametric) signals in huge data streams.
WP2-3 (Q1/2024): Test environment for identifying highly complex (multi-parametric) signals in huge data streams.
WP2-4 (Q3/2024): Generic tools to both convert trained neural networks into efficient HLS/VHDL FPGA firmware
WP2-5 (Q1/2026): Algorithms for massively parallel real-time sorting, clustering and pattern recognition on specialised hardware
WP2-6 (Q1/2026): Algorithms/ML methods for filtering and selecting relevant transient/anomalous signals
WP2-7 (Q3/2026): Pipeline for anomalous signal detection with low false-alarm probability for multi-messenger follow-up

WP5-1 (Q3/2024): Development of ML prototypes for anomaly detection and predictive maintenance
WP5-2 (Q3/2024): Interference recognition and mitigation schemes for transient discovery leading to a robust triggering system
WP5-3 (Q3/2026): Expansion of the concept to a generalized toolkit for predictive maintenance and anomaly detection
WP5-4 (Q3/ 2026): Evaluation of the ML approaches by analyzing false-alarm rates and online feedback

TA5 deliverables

With strong connection to dyn. archive

WP1-1 (Q4/2023): Report on impact of on-line filtering on discovery potential
WP1-2 (Q3/2025): Report on impact of on-line filtering on FAIR principles
WP1-3 (Q3/2026): Concepts towards a general protocol on capturing the decisions made by, and status of, real-time sensors, as a basis for a future demonstator

WP2-1 (Q4/2022): Curation & metadata schemes for dynamic filtering
WP2-2 (Q1/2023): Strategy concept for identifying highly complex (multi-parametric) signals in huge data streams.
WP2-3 (Q1/2024): Test environment for identifying highly complex (multi-parametric) signals in huge data streams.
WP2-4 (Q3/2024): Generic tools to both convert trained neural networks into efficient HLS/VHDL FPGA firmware
WP2-5 (Q1/2026): Algorithms for massively parallel real-time sorting, clustering and pattern recognition on specialised hardware
WP2-6 (Q1/2026): Algorithms/ML methods for filtering and selecting relevant transient/anomalous signals
WP2-7 (Q3/2026): Pipeline for anomalous signal detection with low false-alarm probability for multi-messenger follow-up

Dynamic archive

WP4-1 (Q4/2024): Porting common off-line packages to a memory-based computing prototype
WP4-2 (Q1/2025): Standard software (e.g. CASA) compatible with Gen–Z.
WP4-3 (Q2/2025): Caching strategies for processing of benchmark files
WP4-4 (Q2/2026): Definition and initial implementation of an efficient real-time data processing framework
WP4-5 (30 Sep 2026): Scaled feedback interfaces between off-line software and real-time processes using MeerKAT data

WP5-1 (Q3/2024): Development of ML prototypes for anomaly detection and predictive maintenance
WP5-2 (Q3/2024): Interference recognition and mitigation schemes for transient discovery leading to a robust triggering system
WP5-3 (Q3/2026): Expansion of the concept to a generalized toolkit for predictive maintenance and anomaly detection
WP5-4 (Q3/ 2026): Evaluation of the ML approaches by analyzing false-alarm rates and online feedback

Workflows in connection to use cases

(More or less) documented use cases as starting point:

- Data from **tracking** in high-energy physics
- Data from ground-based **air-shower observations**
- Identification of **pulsars/fast radio bursts including metadata**
- Concepts for related data from **simulations**
- Detection of **anomalies in detector response**
- Identification of **particle decays in software-based triggering**

Already identified relations to **dyn. archiving**:

- Distributed sensors: Outside triggers and sending triggers
- Astronomical transients detected in optical surveys
- Radio transients in real-time

Also: Definition of **information measure** (WP1)

Link to documents: https://gitlab-p4n.aip.de/punch/intra-docs-content/-/tree/master/files/TA5/Documents_deliverables

→ Re-shaping the work in terms of more functional blocks/diagrams
→ Some of the functionalities of a dynamic archive are also relevant for HEP and other fields of data intensive science

Workflows in connection to use cases

Metadata document

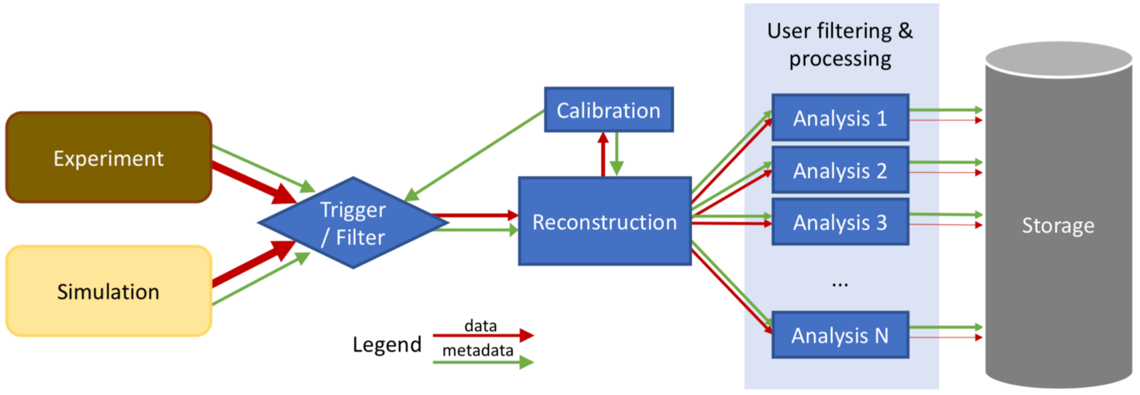


Figure 2: General data processing graph for particle and astroparticle experiments. Variations of the data flow and triggering scheme are possible. The arrow width qualitatively indicates the data rate.

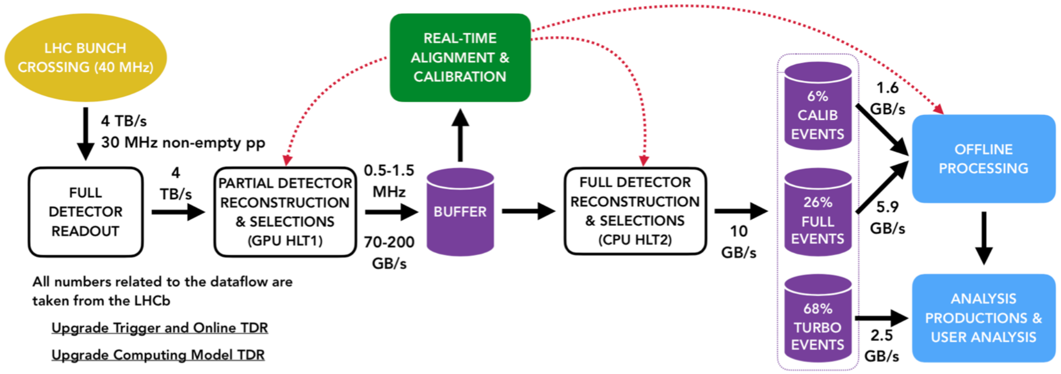
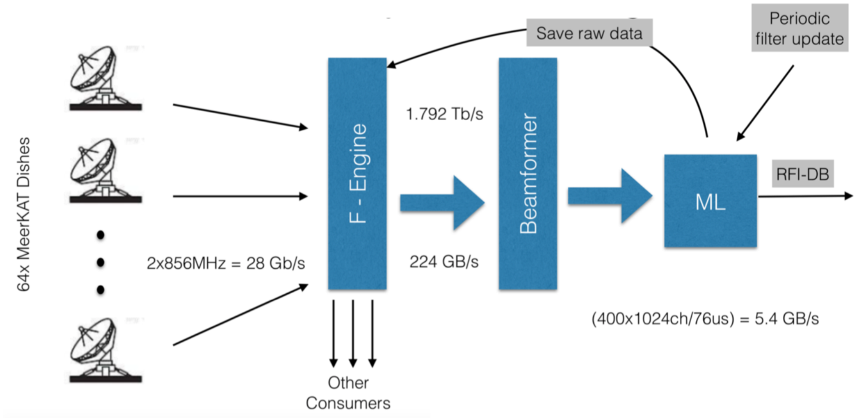


Figure 3: Current data processing pipeline of the LHCb experiment for proton-proton collisions [6, 7]. Arrows indicate data flow, which are annotated with event and data rates.

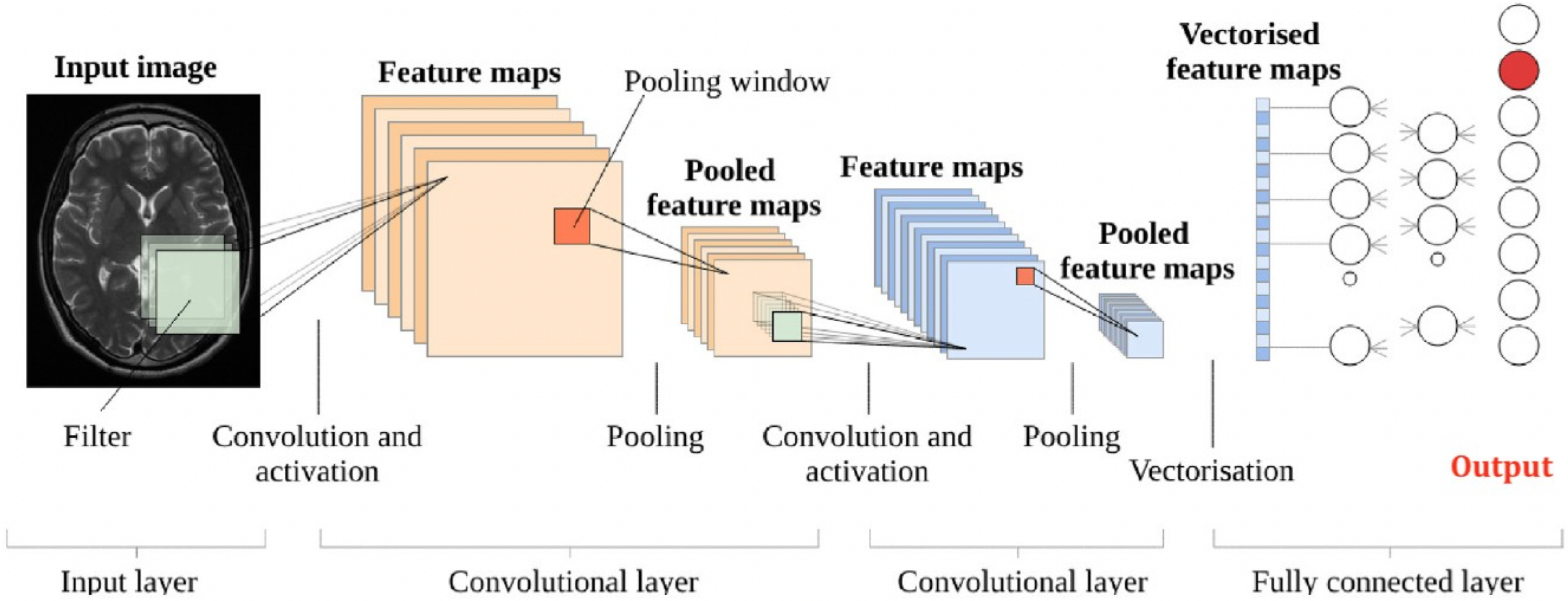
Workflows in connection to use cases

Examples: Dynamic filtering

Example for MeerKAT (MPIfR):



From Ramesh et al.



Pivotal for upcoming deliverable „Test environment for identifying highly complex (multi-parametric) signals in huge data streams“

Workflows in connection to use cases

ML-PPA

Repository: https://gitlab-p4n.aip.de/punch/ta5/wp4/ml-ppa

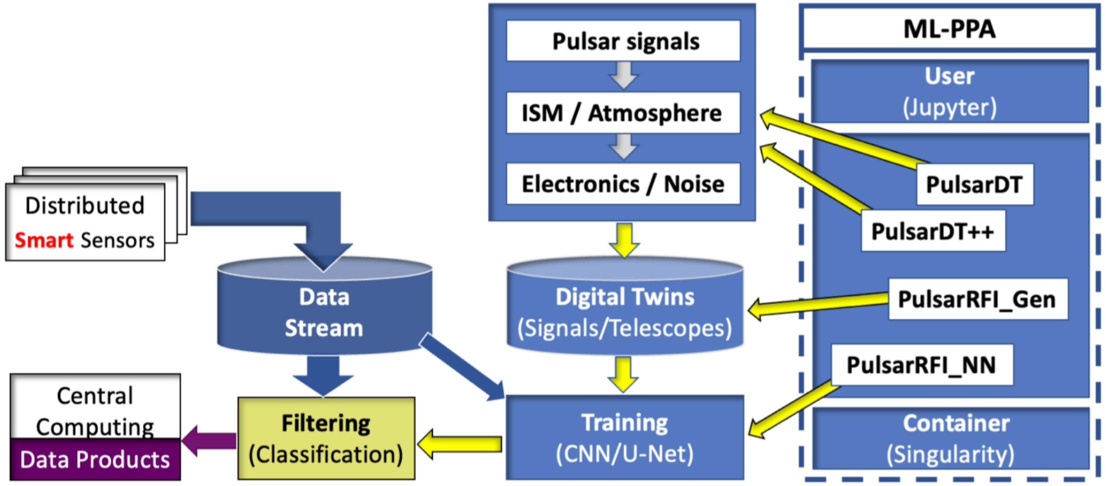


Figure 3: Integration of ML–PPA into the Dynamic Life Cycle Model.

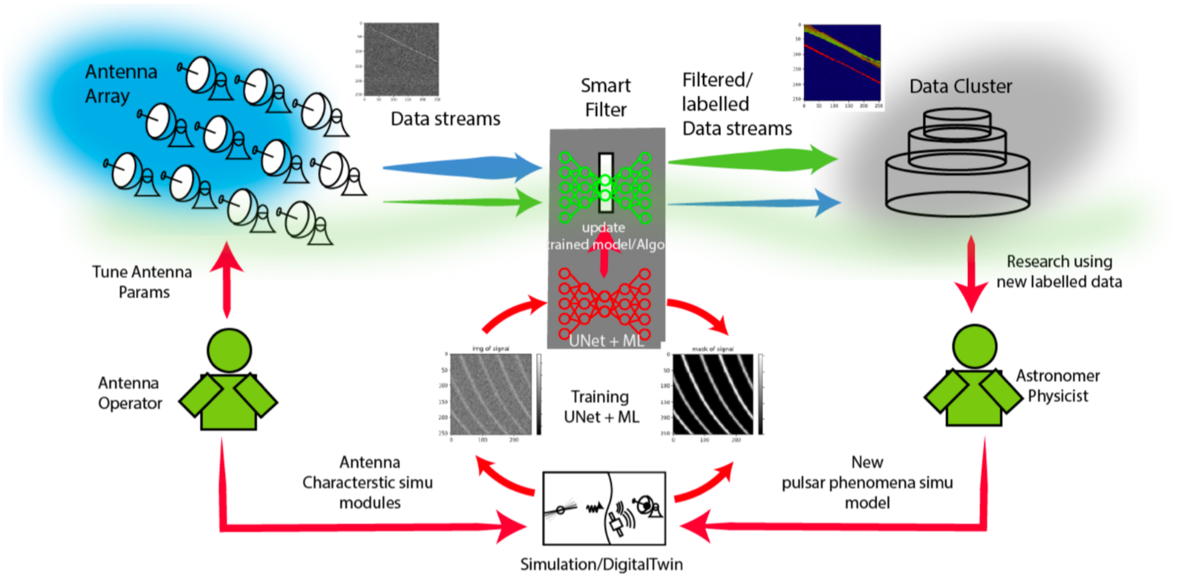


Figure 28: **Summary of the project goal** : *Illustrating the data flow from antenna array to data archiving centers through smart filtering. Efficiently managing the vast data streams generated by radio antennas is crucial, as only a small fraction contains valuable information. To address the challenge of data irreversibility, intelligent filtering methods employing sophisticated algorithms and neural networks play a vital role. Astronomers and physicists analyze labeled signals to gain insights and can integrate new theories into simulation modules.*

Workflows in connection to use cases

Examples: Dynamic archiving document

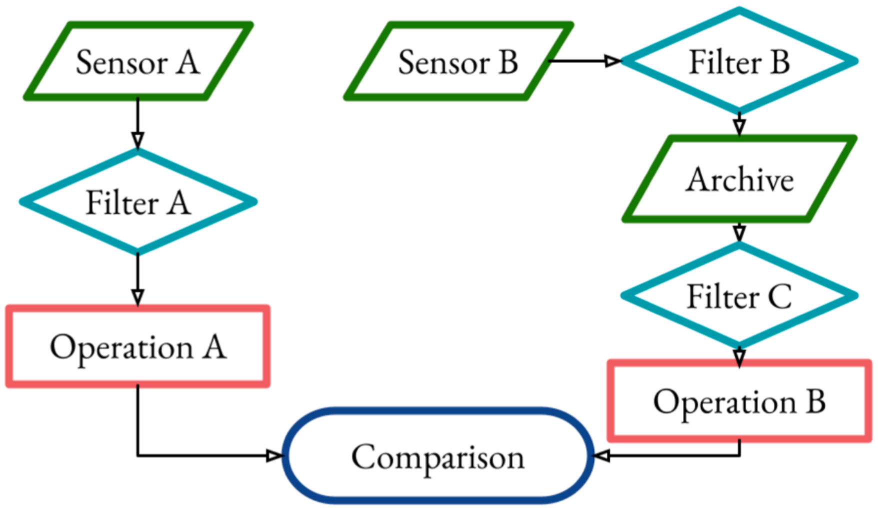


Figure 4: Goal of a dynamical archive: A dataset could have been assembled through the real-time data workflow presented through pipeline A. Alternatively, a dataset could have been produced based on operations on an archive (in turn generated from some other data stream).

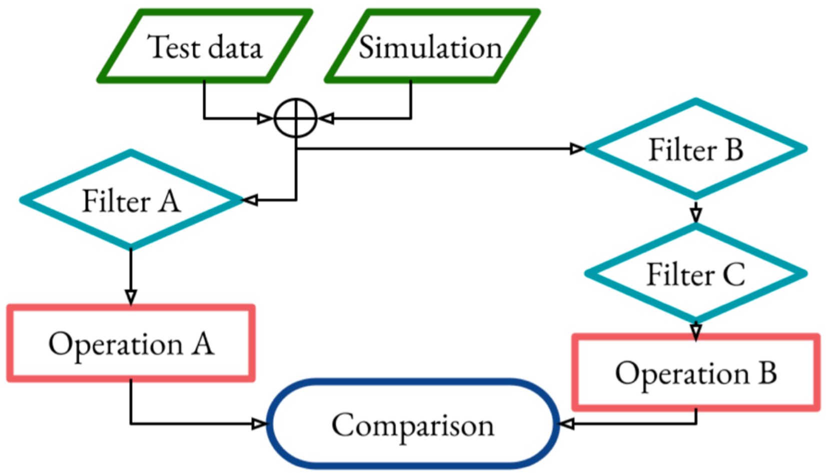


Figure 5: The incompleteness of an archive can be explored using test samples or simulated data. Here the same data is propagated through the filters used both in the real-time and archive setups. The difference is used to estimate the archive incompleteness *for these particular filters, test data and comparison statistic.*

Summary and outlook

Need for more detailed/technical description of workflows in connection to use cases
→ Including definition of interfaces

Focus on synergies between prototypes and also the strong connection to dyn. filtering

Need for more open (and usable) test data sets

Quite some achievements within 1.5 years:

- Collaboration established
- Concepts documented
- ML-PPA as a working example
- ML on FPGAs working group
- Test environments

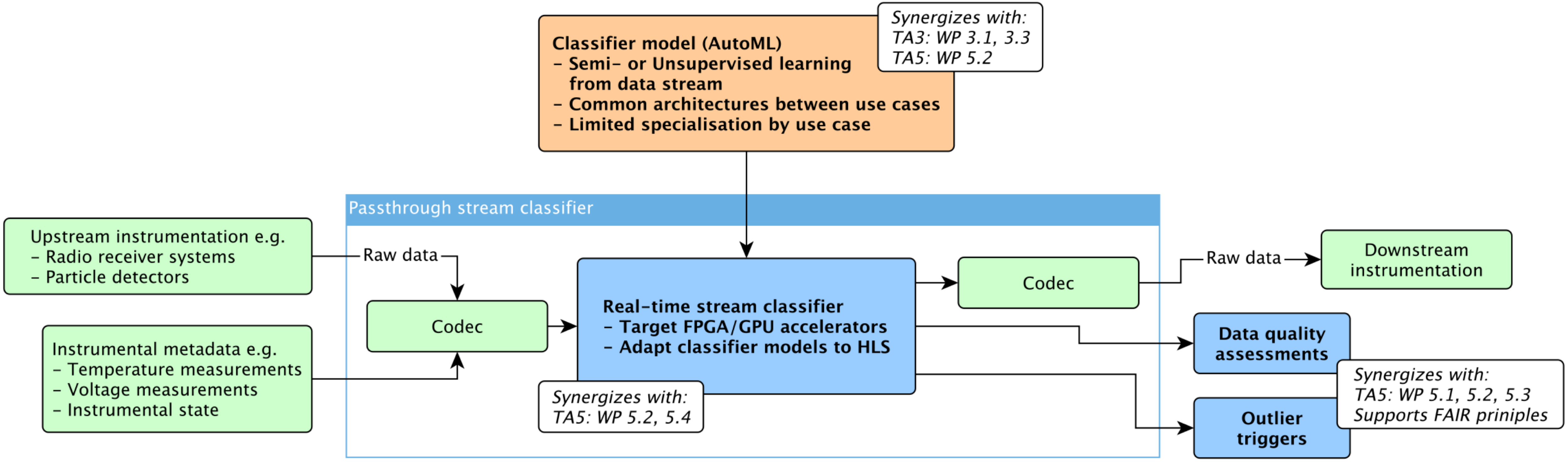
Solutions and pioneering concepts for real-time data challenges as “Alleinstellungsmerkmal”

Towards end of this workshop: Summary and discussion of next steps, next meetings

BACKUP

WP5 Evaluation and validation of instrument response

Workflow developed back in 2021



- WP aims to develop prototype ML systems for real-time anomaly detection and predictive maintenance
- Targets high-performance accelerator architectures: FPGAs and GPUs to handle high-data-rate streams (e.g. radio telescope voltages)
- Requires standardised codecs for interoperability between varied instrument types
- Will produce rich high-resolution quality assessments of the data stream for filtering and archiving (supporting reusability principle of FAIR)