



HiPEAC 2024

Anomaly Detection In EVEREST

Automated Anomaly Detection to improve Security in a High- Performance Heterogeneous Environment

TOM SLOOFF

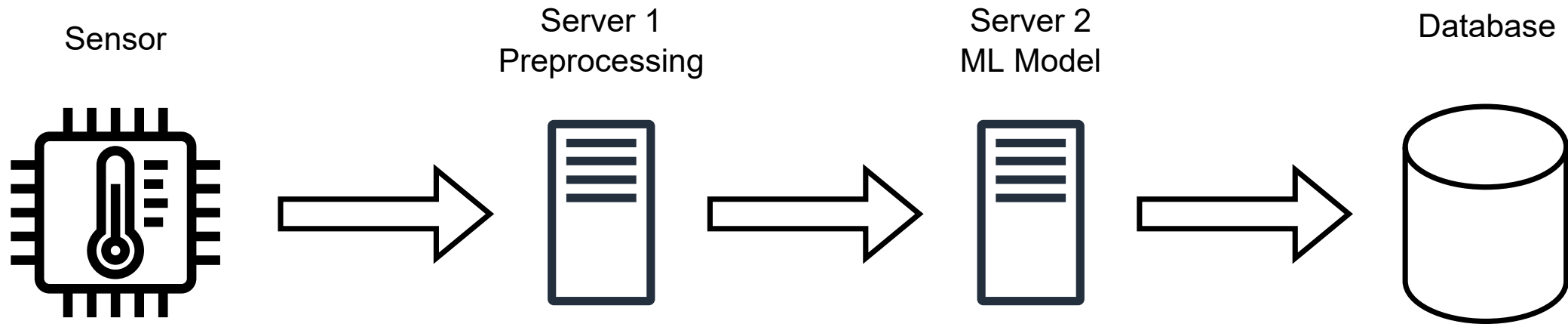
Università della Svizzera italiana, Lugano, Switzerland

tom.slooff@usi.ch

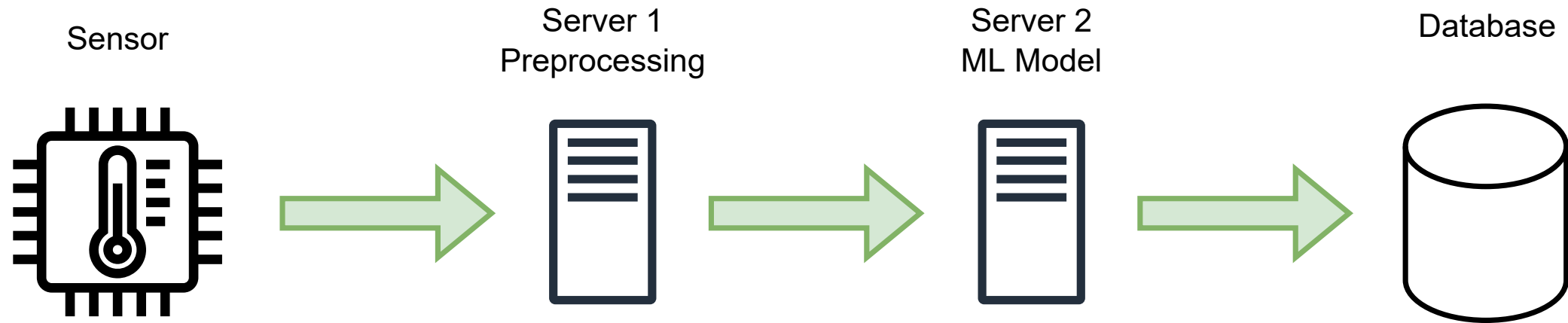
Anomaly Detection

- Detecting abnormal datapoints in datasets or data streams.
- Security?

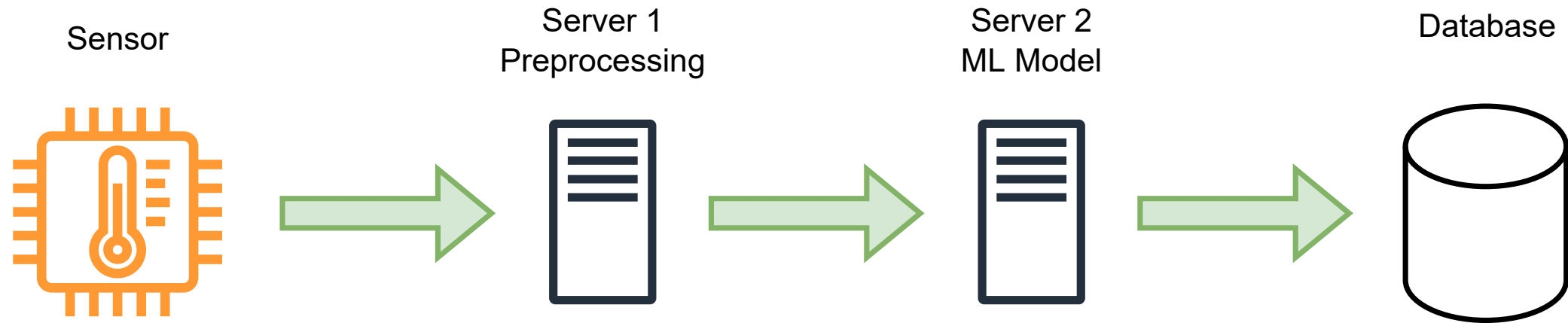
Example Workflow



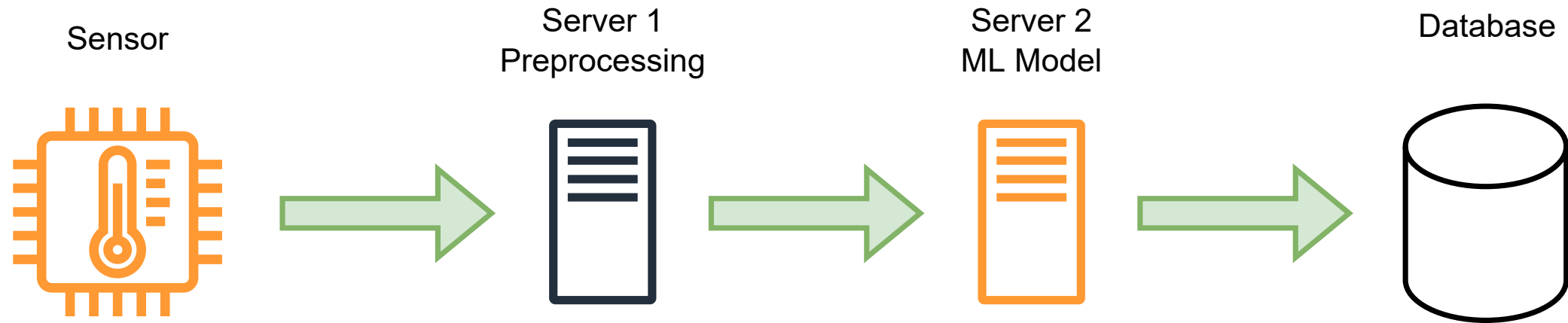
Example Workflow



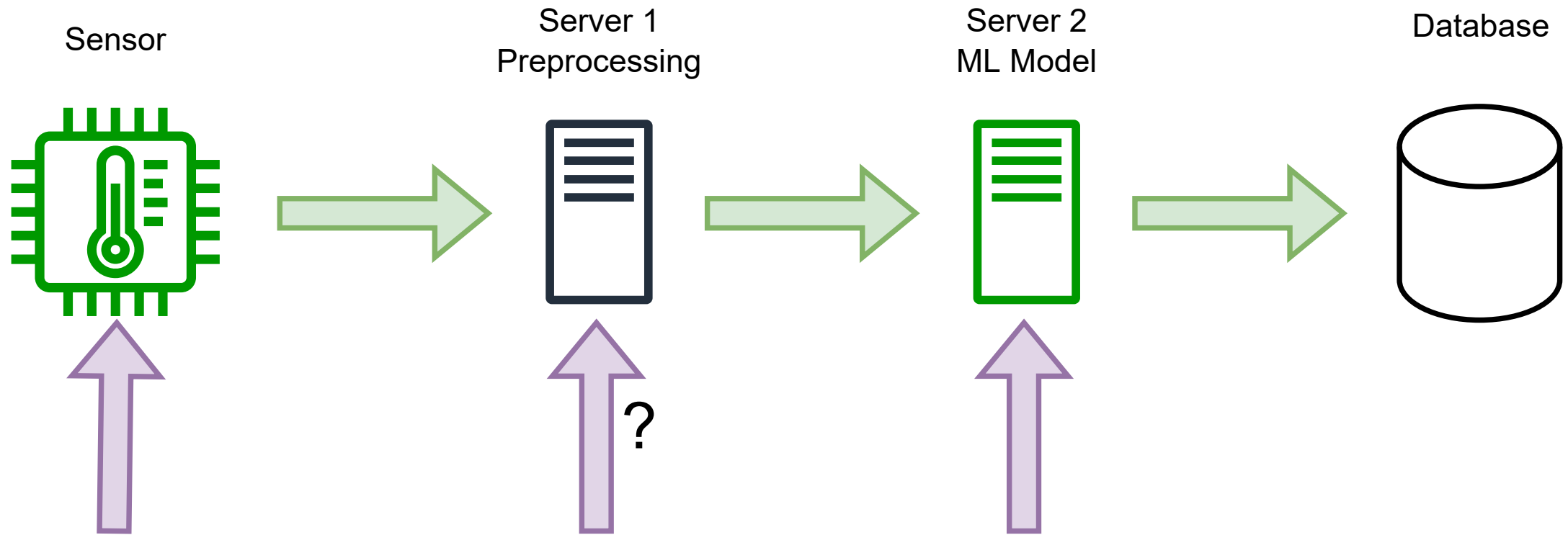
Example Workflow



Example Workflow



Where to deploy AD



interface

Data Collection



ModelSelection



Detection



WRF

Dataset with metadata (D1)

Dataset with metadata (D2)

AD Model M1 (D10)

Dataset with metadata (D3)

D1 AD Index
DDI (iRODS)

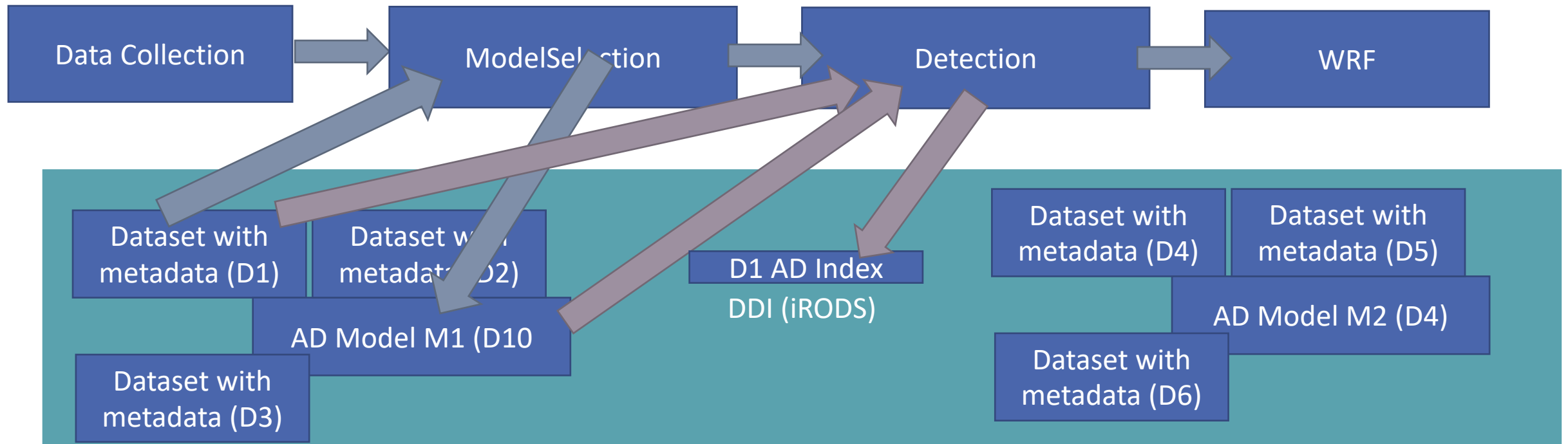
Dataset with metadata (D4)

Dataset with metadata (D5)

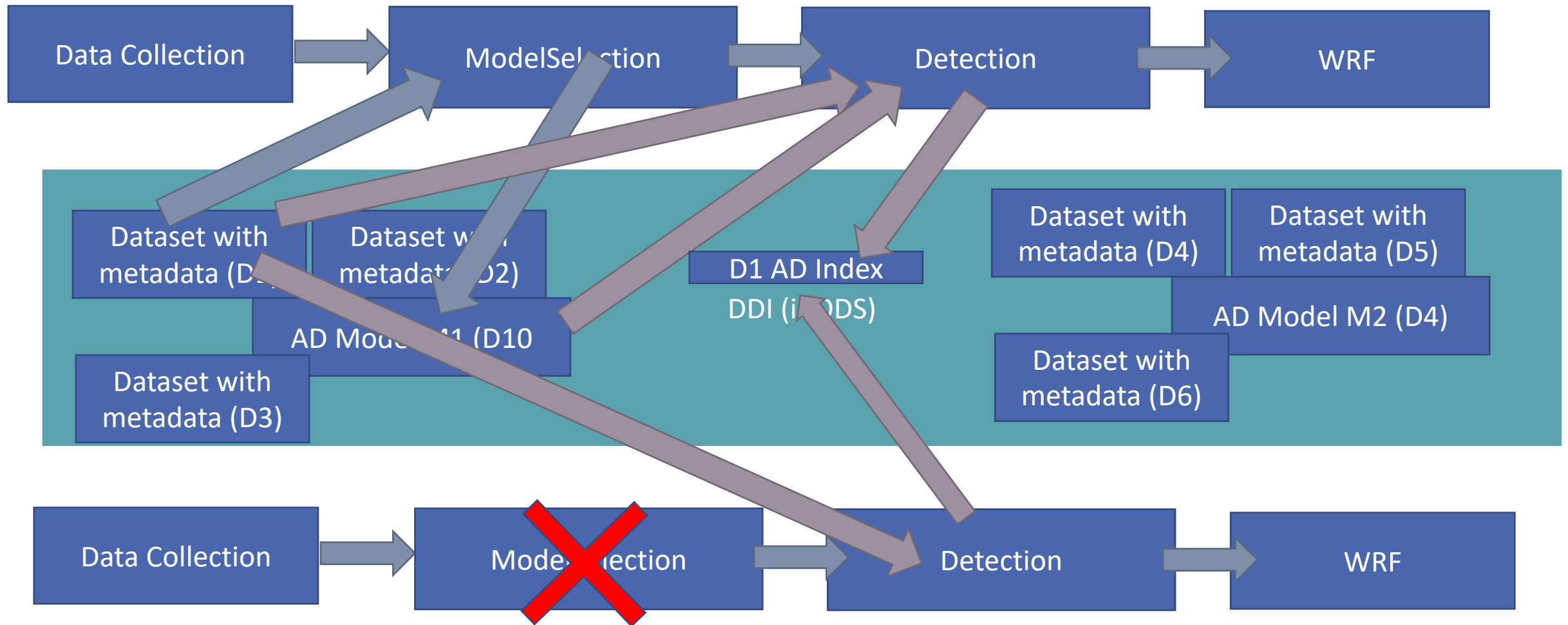
AD Model M2 (D4)

Dataset with metadata (D6)

interface



interface



Interface of library

```
(ad_env) tom@pro everest_ad_script % python adlib/model_selection/run.py -h
usage: run.py [-h] [-f] [--tag TAG] [-s] [--timeout TIMEOUT] [-db] [-p]
```

optional arguments:

```
-h, --help            show this help message and exit
-f, --fpga            This flag indicates the found model should be runnable on FPGA.
--tag TAG, -t TAG    Name of the experiment, will be used to name database. Defaults to name of input data file.
-s, --skip           This flag skips the optimization and causes immediate exit.
--timeout TIMEOUT, -to TIMEOUT
                    Time limit for the optimization in minutes. After timeout current trials will finish before exiting.
-db                  Indicates the database used by the optimizer should be kept after completion. This database can be reused.
-p                  This flag disables the progress bar for each trial.
```

Data is taken from ./data , with possibly required config file
Model stored in ./model

Available as python script, docker available

General	
slice_columns	Columns to process as list of indices, e.g. [0, 2, 3] for 2D data or [(0,0), (0,1), (1,2)] for 3D data. If not given all data will be processed.
TODO: resolution	Resolution per column: the smallest significant difference for a column. E.g. for a celcius temperature 0.5. Expected as list, currently not supported yet.
TODO: groupby	In case of grouped data, this indicates which feature the data is grouped by. E.g. if you have multiagent data, this will indicate which feature identifies the agent.
dtype	Data type per column. As a list of numpy datatype strings. For valid strings see numpy.sctypeDict NOTE: if you use numpy.tofile(), all shape information is also lost so this should also be included in the datatype description if it is important. I would recommend to use numpy.save() instead, as this works much better.
For csv / excel	
csv_delimiter	e.g. ";" (default), ",", ...
header	None if the csv does not include a header, otherwise row index of header
for netcdf	
netcdf_variable	variable to take for the anomaly detection

Interface of library

```
(ad_env) tom@pro everest_ad_script % python adlib/detection/run.py -h
usage: run.py [-h] [-c CUTOFF] [-p]

Anomaly Detection algorithm applied to all data in data folder.

optional arguments:
  -h, --help            show this help message and exit
  -c CUTOFF, --cutoff CUTOFF
                        cutoff value for anomaly detection. anomaly scores are in [0,1]. You can specify higher cutoff values, in v
                        will be detected.
  -p                    This flag disables the progress bar when parsing the files.
```

Data is taken from ./data , no config file necessary as it is stored in model

Model taken from ./model

Produces anomalies.json file in ./data containing anomalous indexes for all processed data files

Conclusion

- Anomaly Detection can be used to improve security
- We utilise a simple interface within EVEREST to have the opportunity to expand in the future

Future Work

- Accelerating models
- Constrained optimisation



POLITECNICO
MILANO 1863



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 957269

Thank You!