EVENFLOW

Project use cases on robust learning and reasoning

Athanasios (Thanasis) Poulakidas

Project coordinator

netcompany

intrasoft

AIoD Technical Contributor Board

Teleconference, 23 February 2024

Funded by the European Union

This document is part of a project that is funded by the European Union under the Horizon Europe agreement No 101070430. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the Commission can be held responsible for them.

Quick project overview

EVENFLOW info card

- Project title: **Robust Learning and Reasoning for Complex Event Forecasting**
- Project type: Horizon RIA, in topic HORIZON-CL4-2021-HUMAN-01-01 "Verifiable robustness, energy efficiency and transparency for Trustworthy AI: Scientific excellence boosting industrial competitiveness (AI, Data and Robotics Partnership)"
- Duration: 1 October 2022 – 30 September 2025
- Budget/funding: about 4M€ (EU + UK)

Project Partners

netcompany intrasoft

DEMOKRITOS

ATHENA Research & Innovation Information Technologies

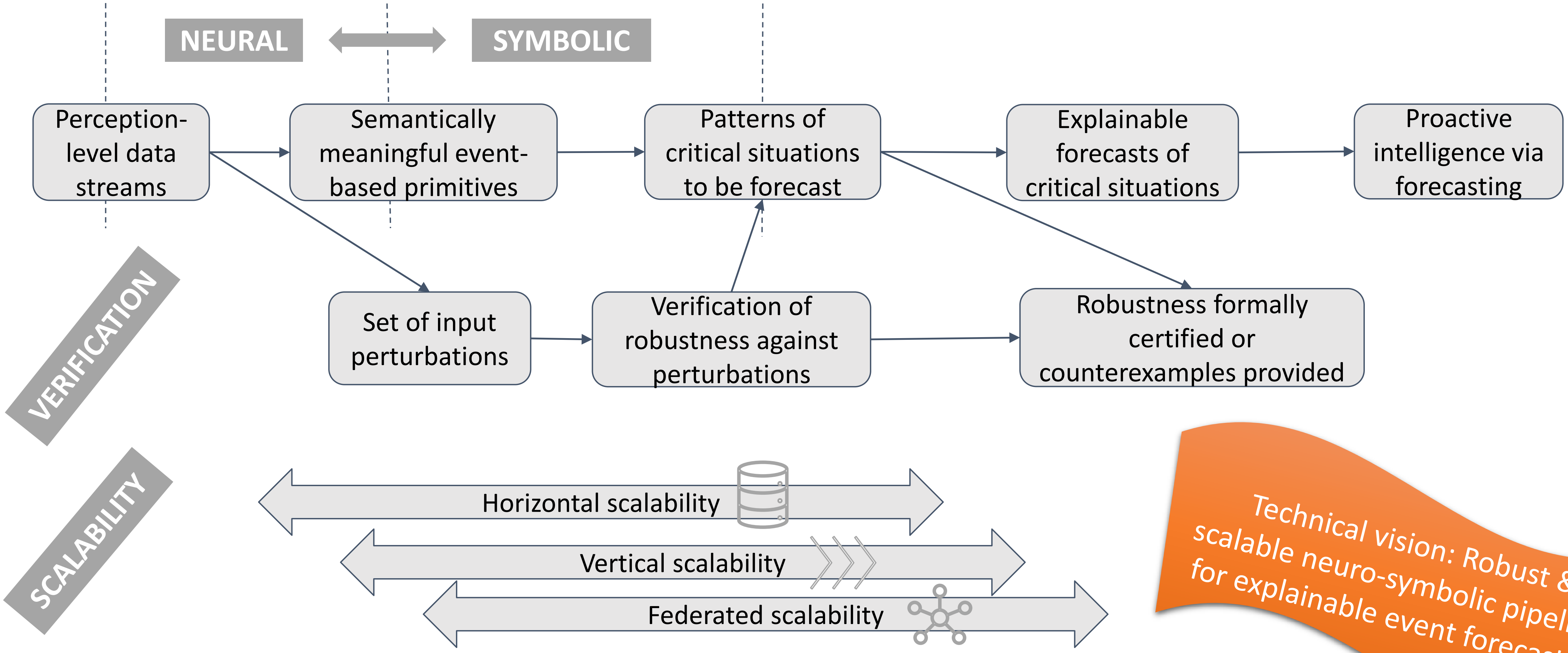
Imperial College London

Barcelona Supercomputing Center Centro Nacional de Supercomputación

dfki ai

Ekso SERVIZI E TECNOLOGIE NO-DIG

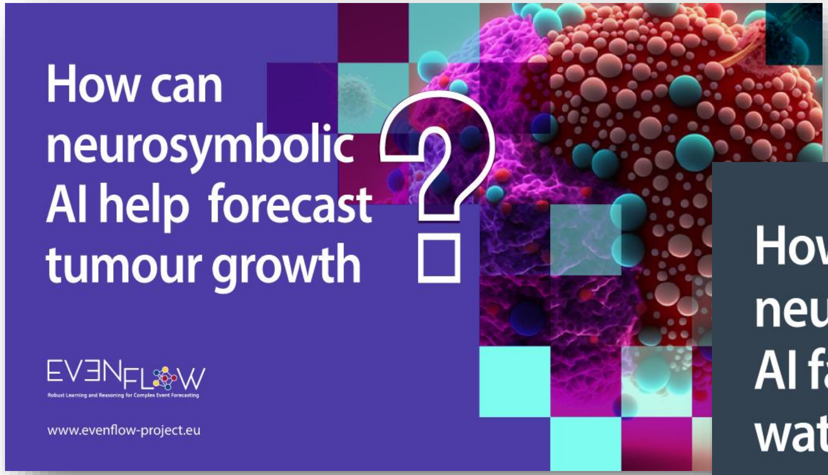
Motivation and vision



Technical vision: Robust & scalable neuro-symbolic pipelines for explainable event forecasting

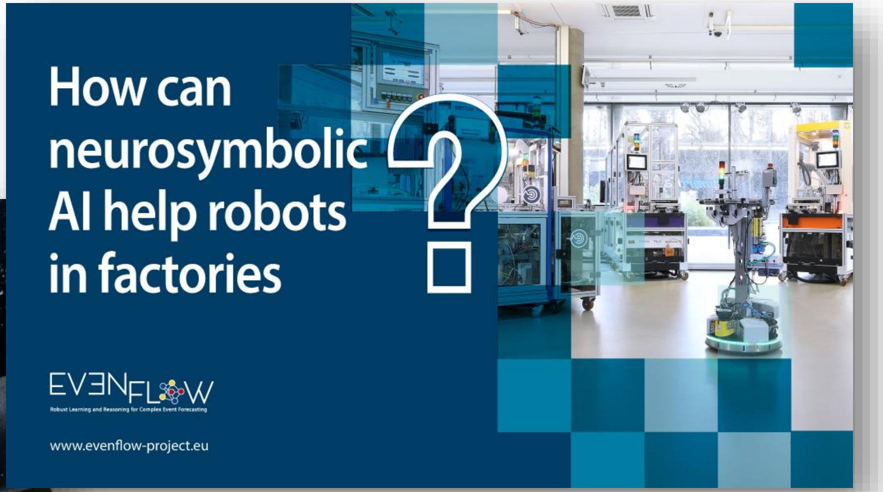
Use cases

Personalised Medicine



Infrastructure Life Cycle Assessment

Industry 4.0



EVENFLOW

Personalised Medicine Use Case

Led by: Barcelona Supercomputing Center

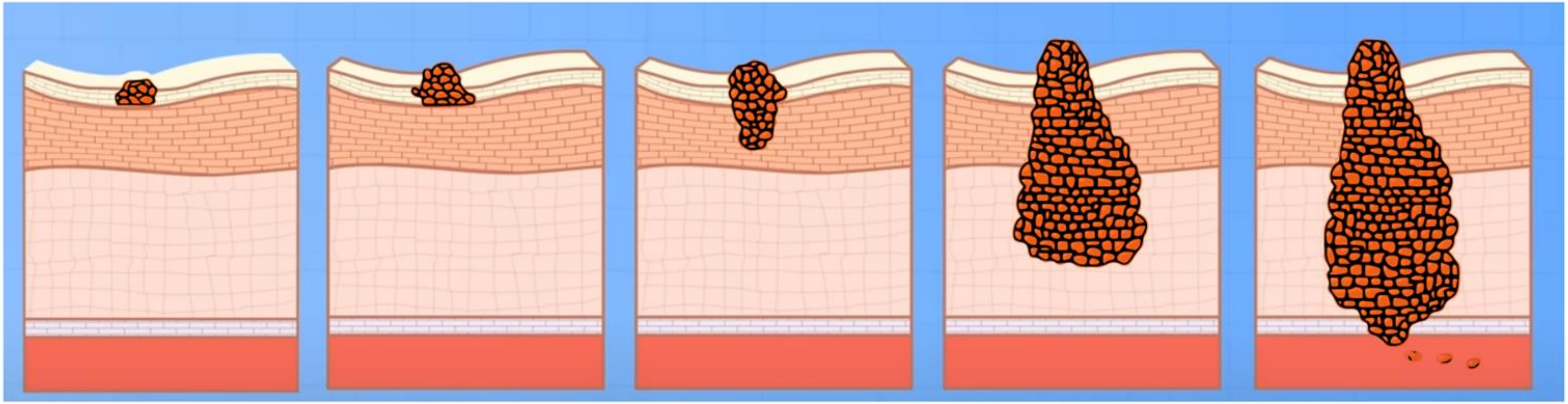
Centro Nacional de Supercomputación

Use case description

Cancer stages

benign tumor

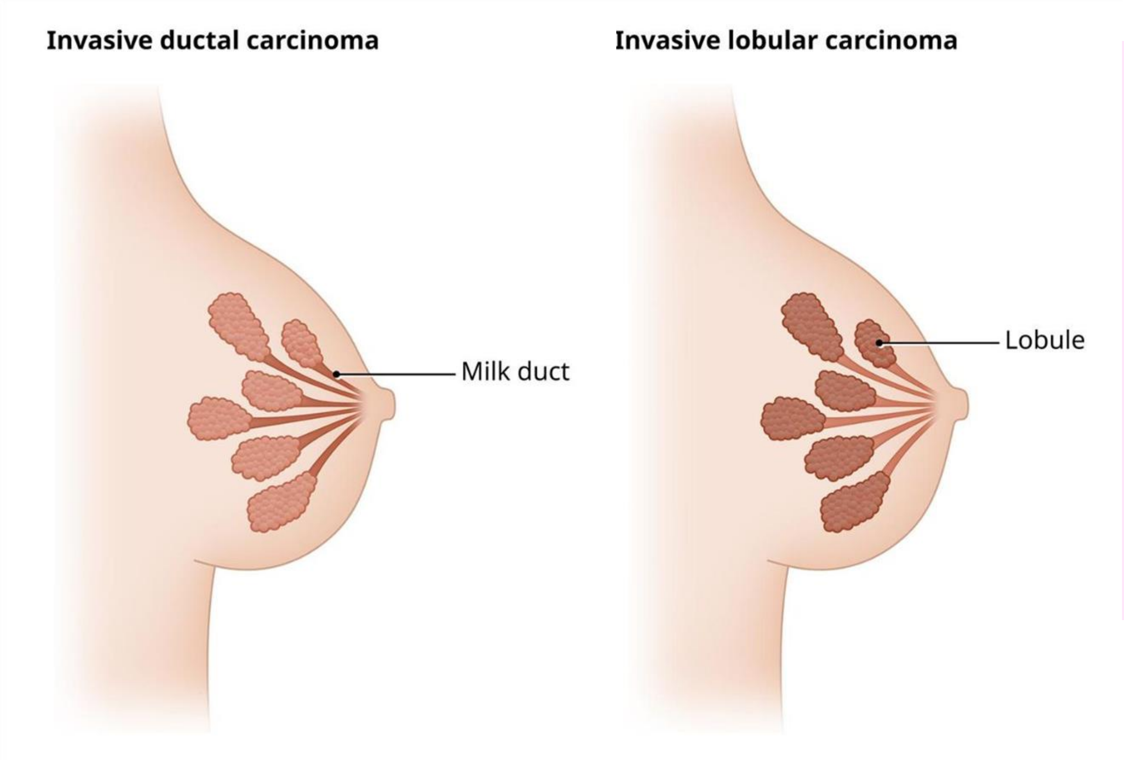
metastasis

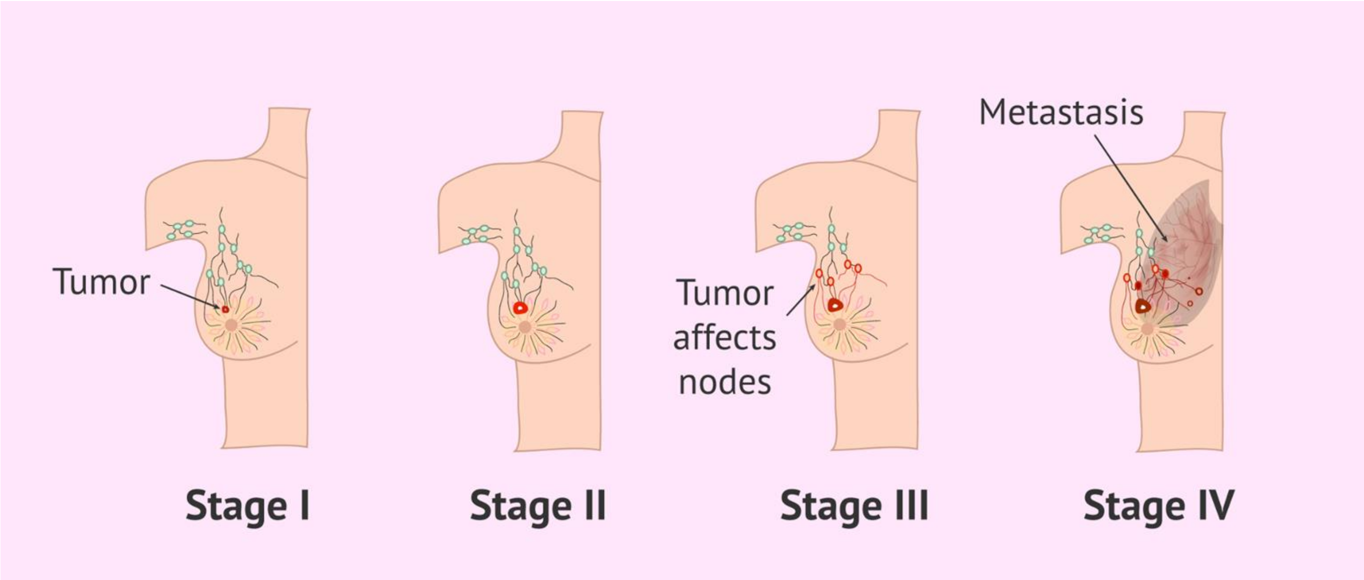


Stage 0 Stage I Stage II Stage III Stage IV

Use case description

Breast cancer





100% 5-year survival

90% 5-year survival

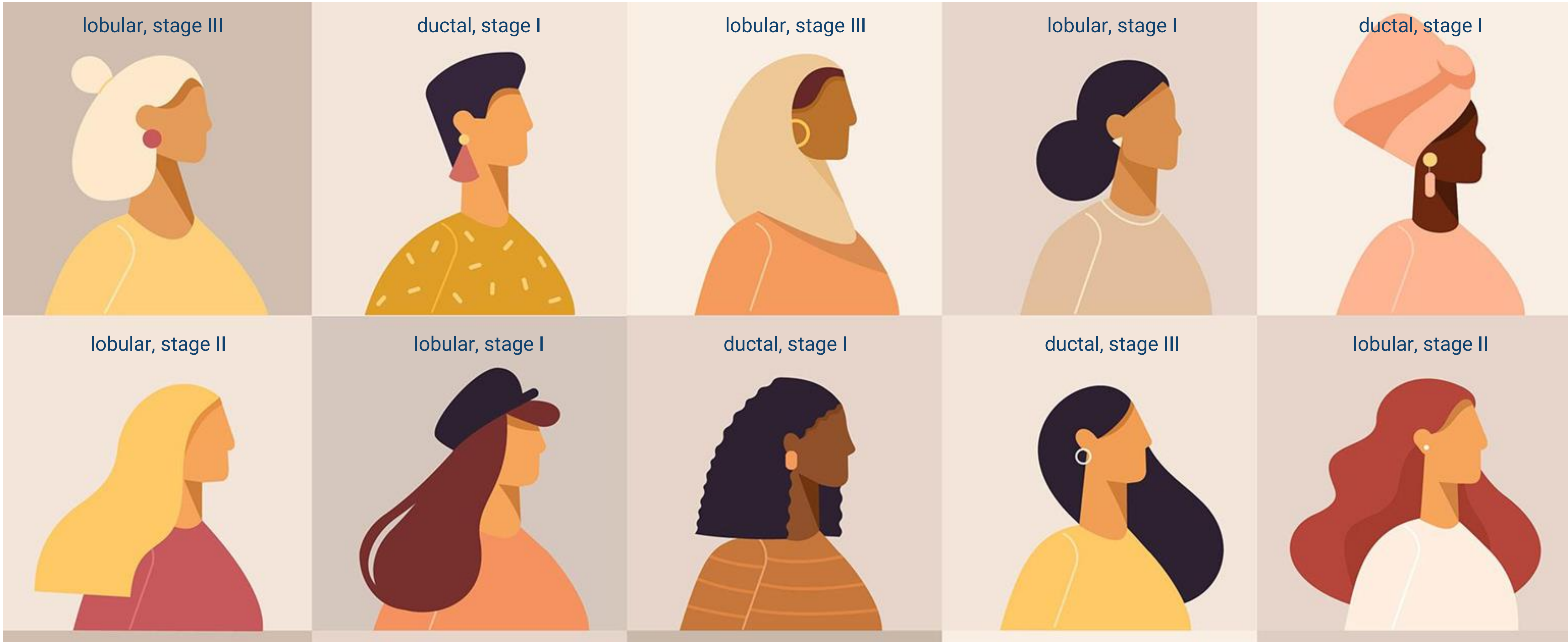
70% 5-year survival

25% 5-year survival

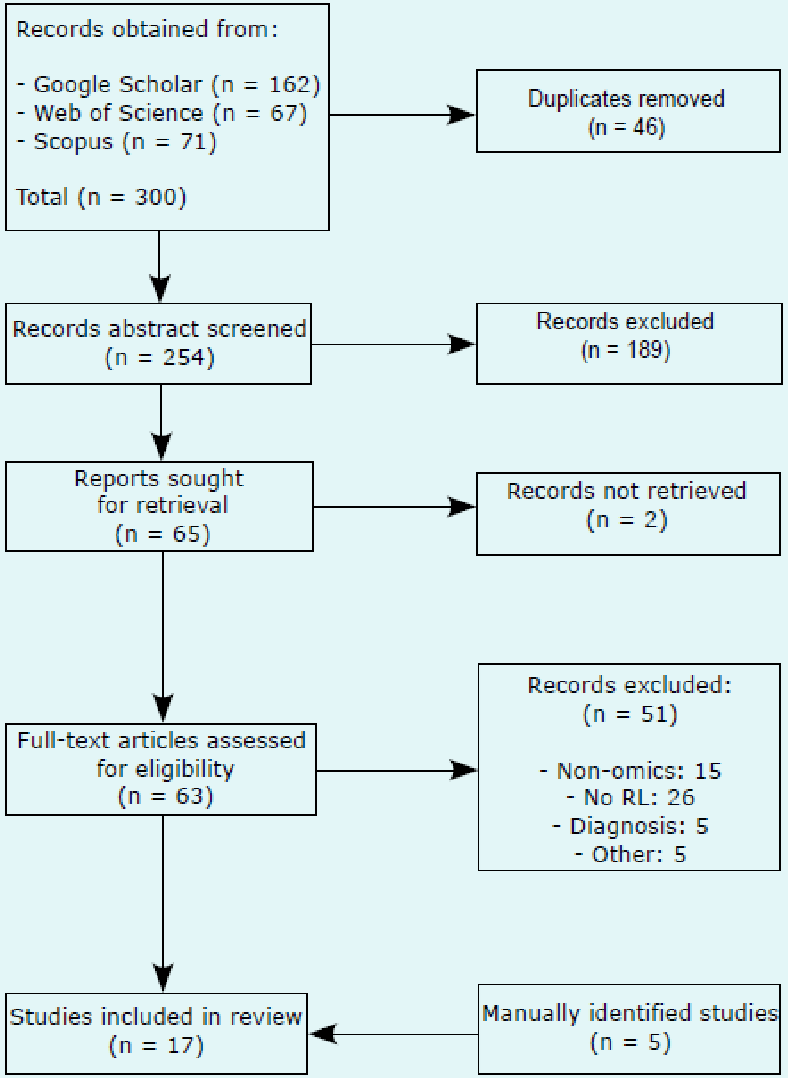
TCGA
The Cancer Genome Atlas

~900 women

gene expression data (RNA-sequencing)



Systematic Literature Review



Studies included in review

1. Franco et al., 2021. DOI: 10.3390/cancers13092013
2. Wang and Wang, 2019. DOI:10.1186/s12859-019-3130-9
3. Islam et al., 2020. DOI:10.1016/j.csbj.2020.08.005
4. Zhang et al., 2022. DOI: 10.3389/fgene.2022.855629
5. Song et al., 2022. DOI: 10.1093/bib/bbab398
6. Hira et al., 2021. DOI: 10.1038/s41598-021-85285-4
7. Xu et al., 2019. DOI: 10.1186/s12859-019-3116-7
8. Mondol et al., 2022. DOI: 10.1109/TCBB.2021.3066086
9. Madhumita and Paul, 2022. DOI: 10.1016/j.compbiomed.2022.105832
10. Ryu et al., 2021. DOI: 10.1007/978-981-33-6757-9_52
11. Zhang et al., 2020. DOI: 10.1007/978-3-030-60796-8_38
12. Yang et al., 2021. DOI: 10.1093/bioinformatics/btab625
13. Chaudhary et al., 2018. DOI: 10.1158/1078-0432.CCR-17-0853
14. Guo et al., 2019. DOI: 10.1016/j.neucom.2018.03.072
15. Cristovao et al., 2022. DOI: 10.1109/TCBB.2020.3042309
16. Caravagna et al., 2018. DOI: 10.1038/s41592-018-0108-x
17. Sun et al., 2017. DOI: 10.1093/nar/gkx003

Flow diagram of the systematic literature review process, following PRISMA guidelines. RL: representation learning

State of the Art

Prol-Castelo et al. 2024
[in preparation]

- Google Scholar, Scopus, WoS
- 2017 - 2022
- Results:
  - Only AI-based studies on cancer subtyping or molecular characterization of tumor progression
  - **No AI-based studies on cancer progression through time**

Challenges and achievements

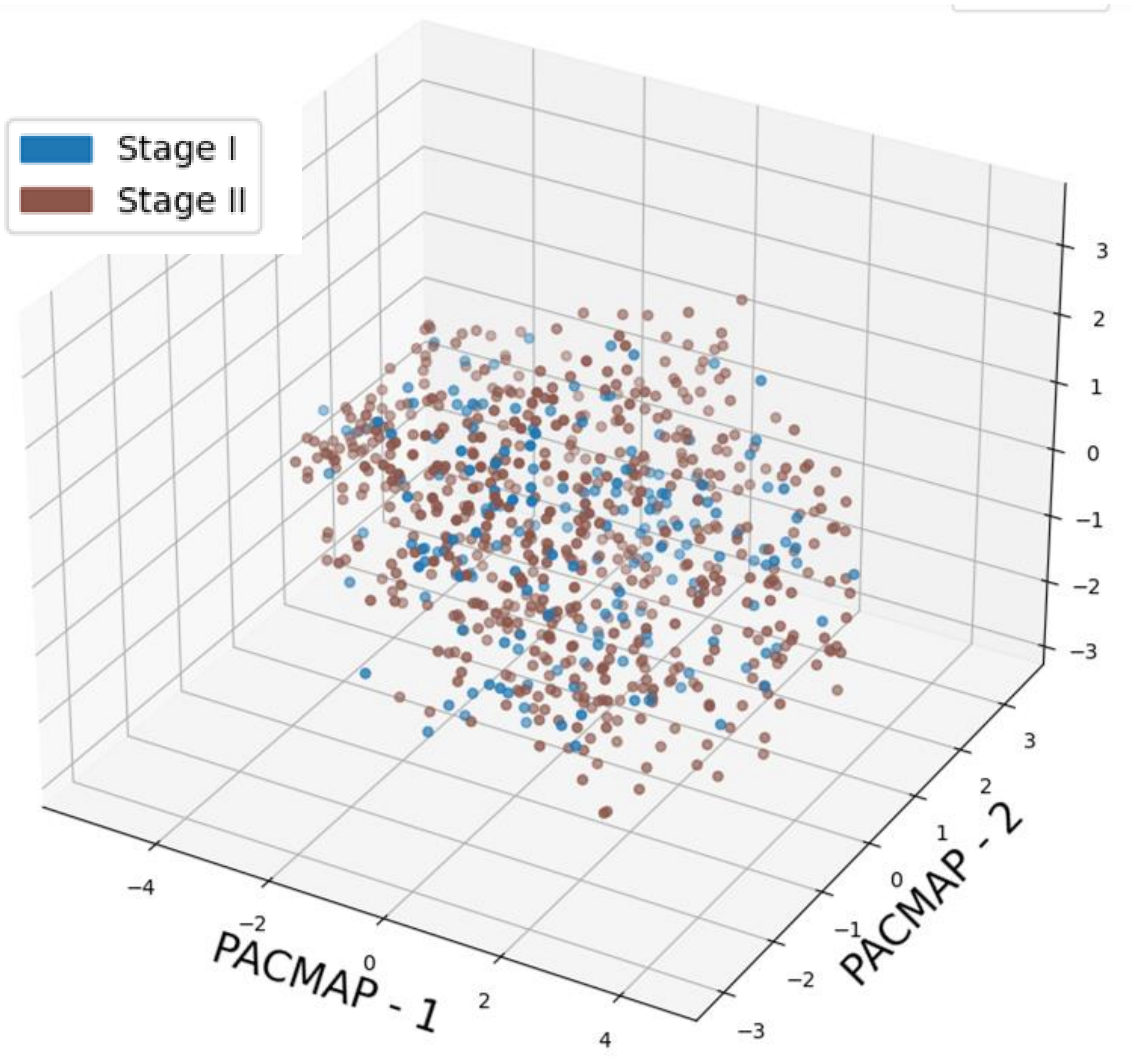
Challenges:

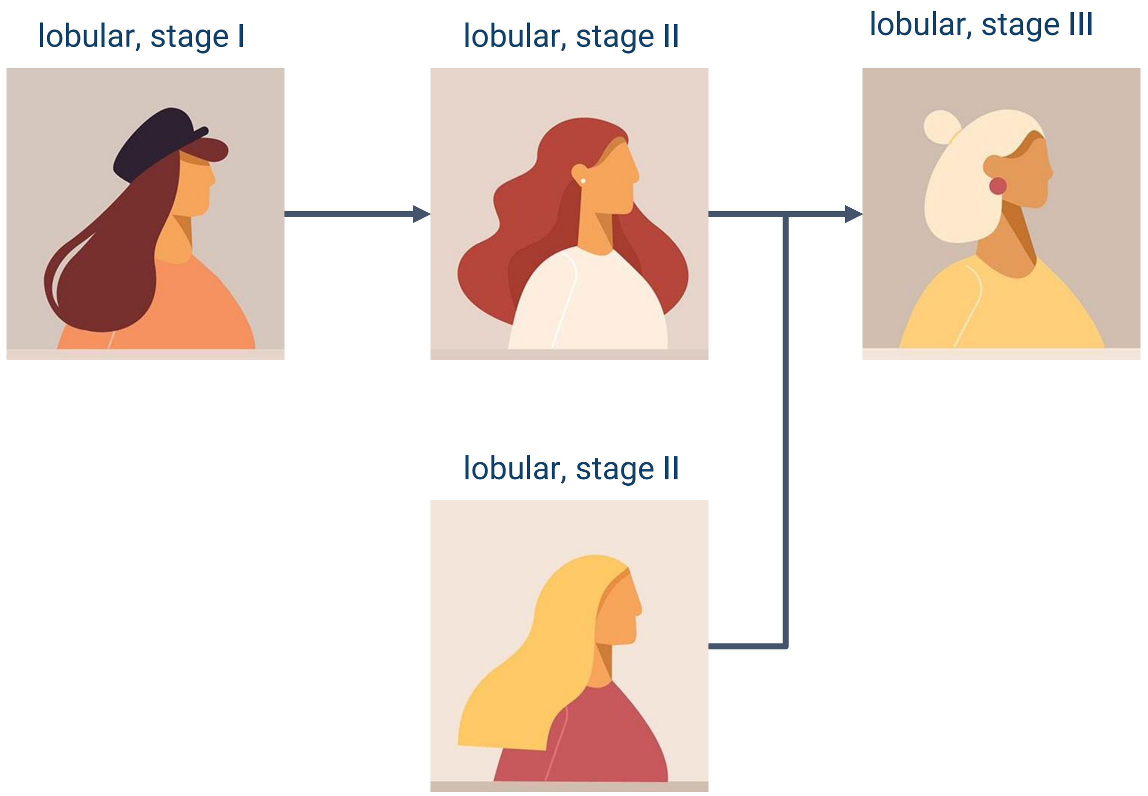
1. Reconstruct dynamic processes from "**snapshots**" without complete trajectories or a known time frame.
2. Model data with **high dimensionality** and low separability
3. Forecast clinically-relevant **molecular changes**

Achievements:

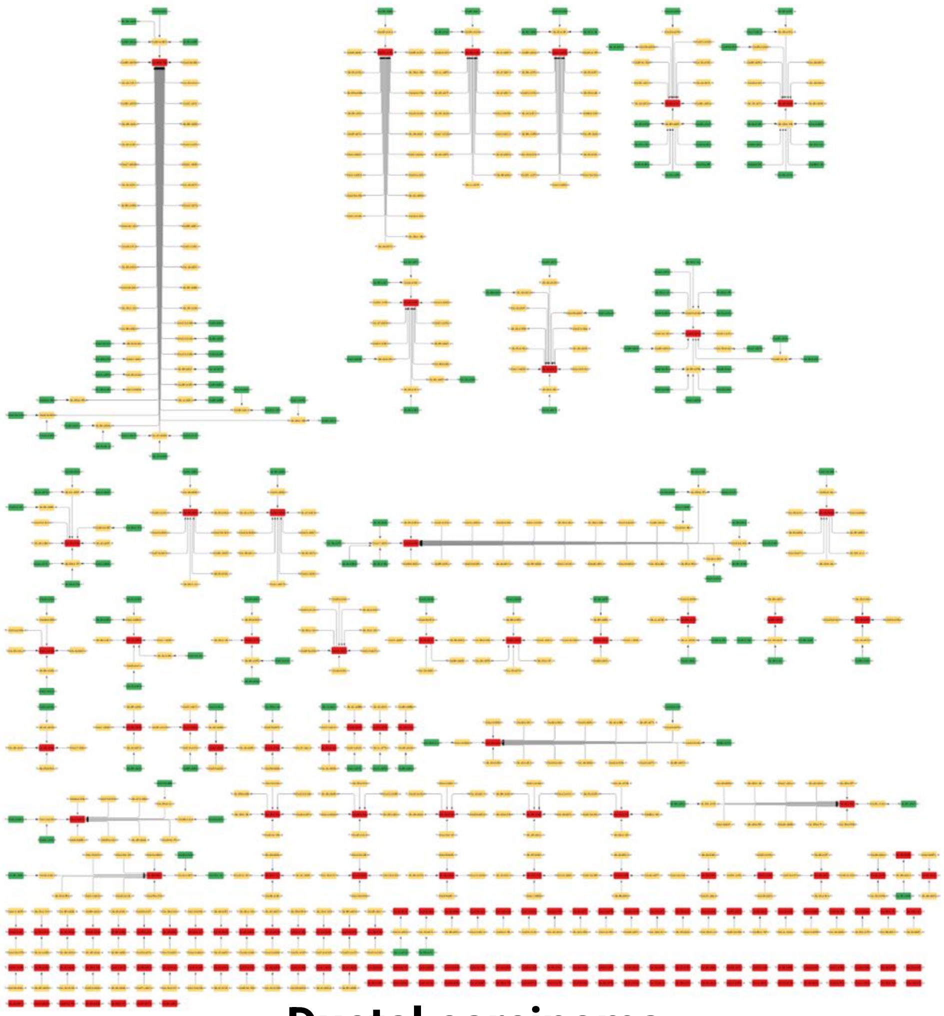
1. Define **pseudo-trajectories** based on patient similarity
2. Quantify **pathways activation** from gene expression
3. Generate **synthetic data** between stages

Pseudo-trajectories based on patient similarity

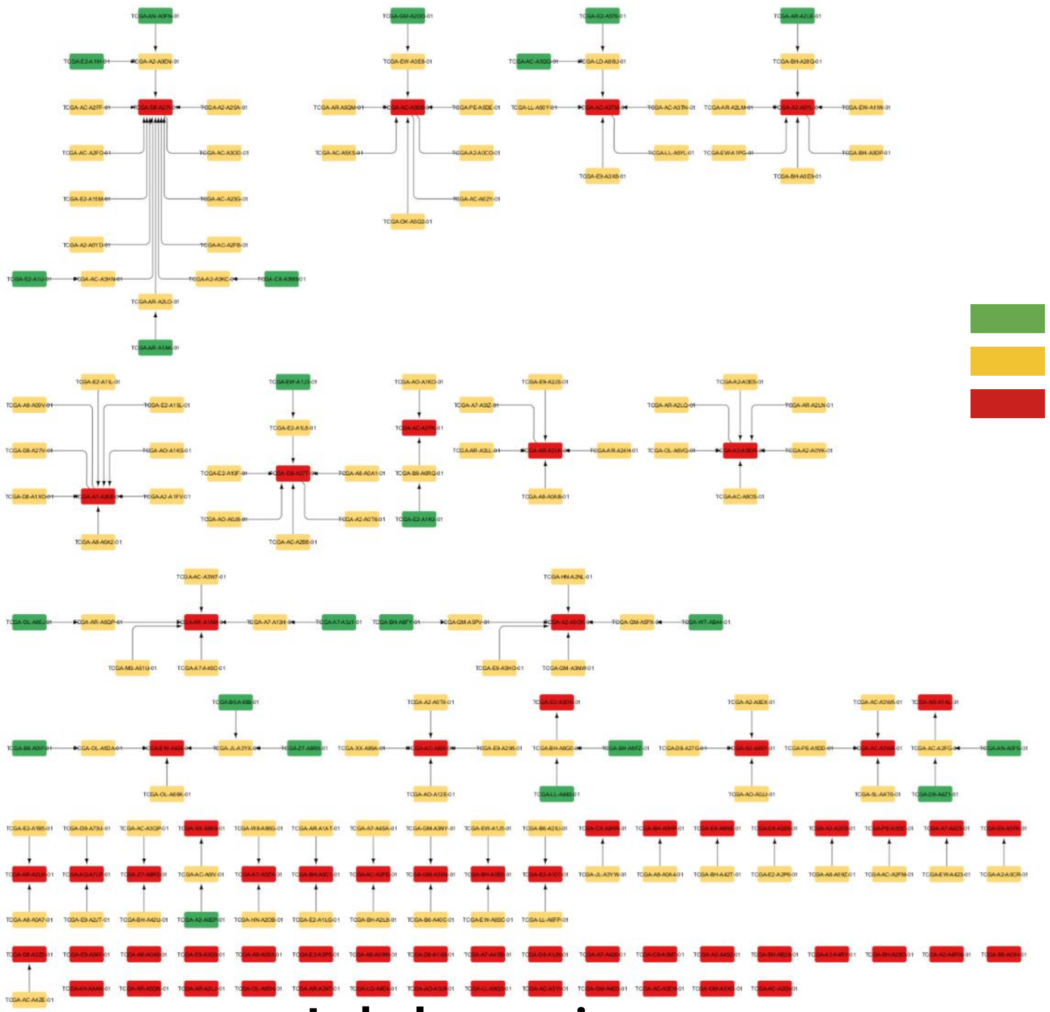




Pseudo-trajectories based on patient similarity



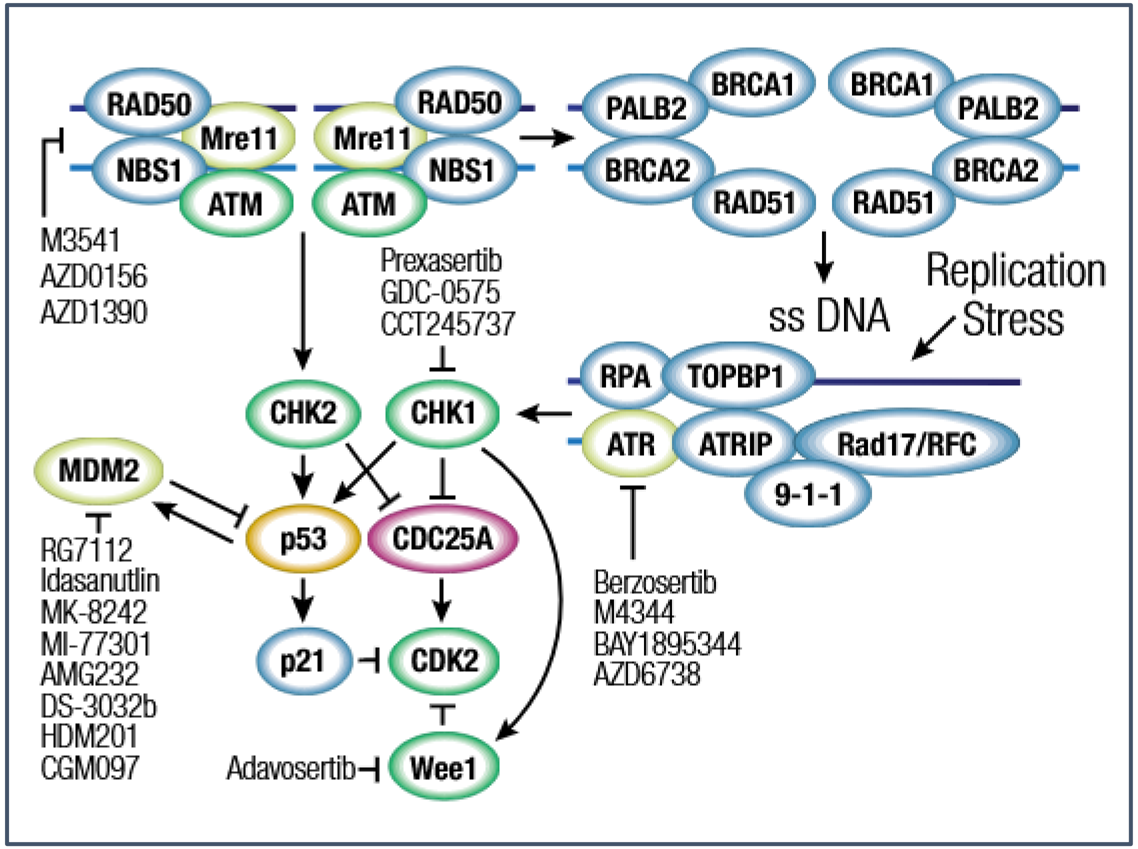
Ductal carcinoma patients



Lobular carcinoma patients

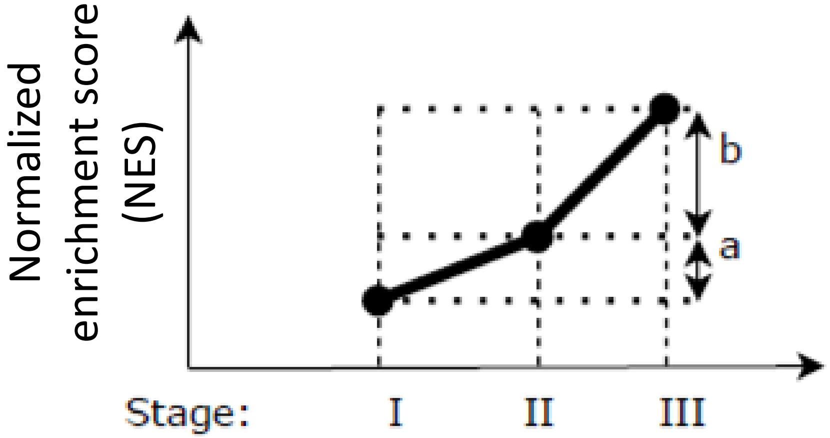
STAGE I
STAGE II
STAGE III

Harnessing biological pathways

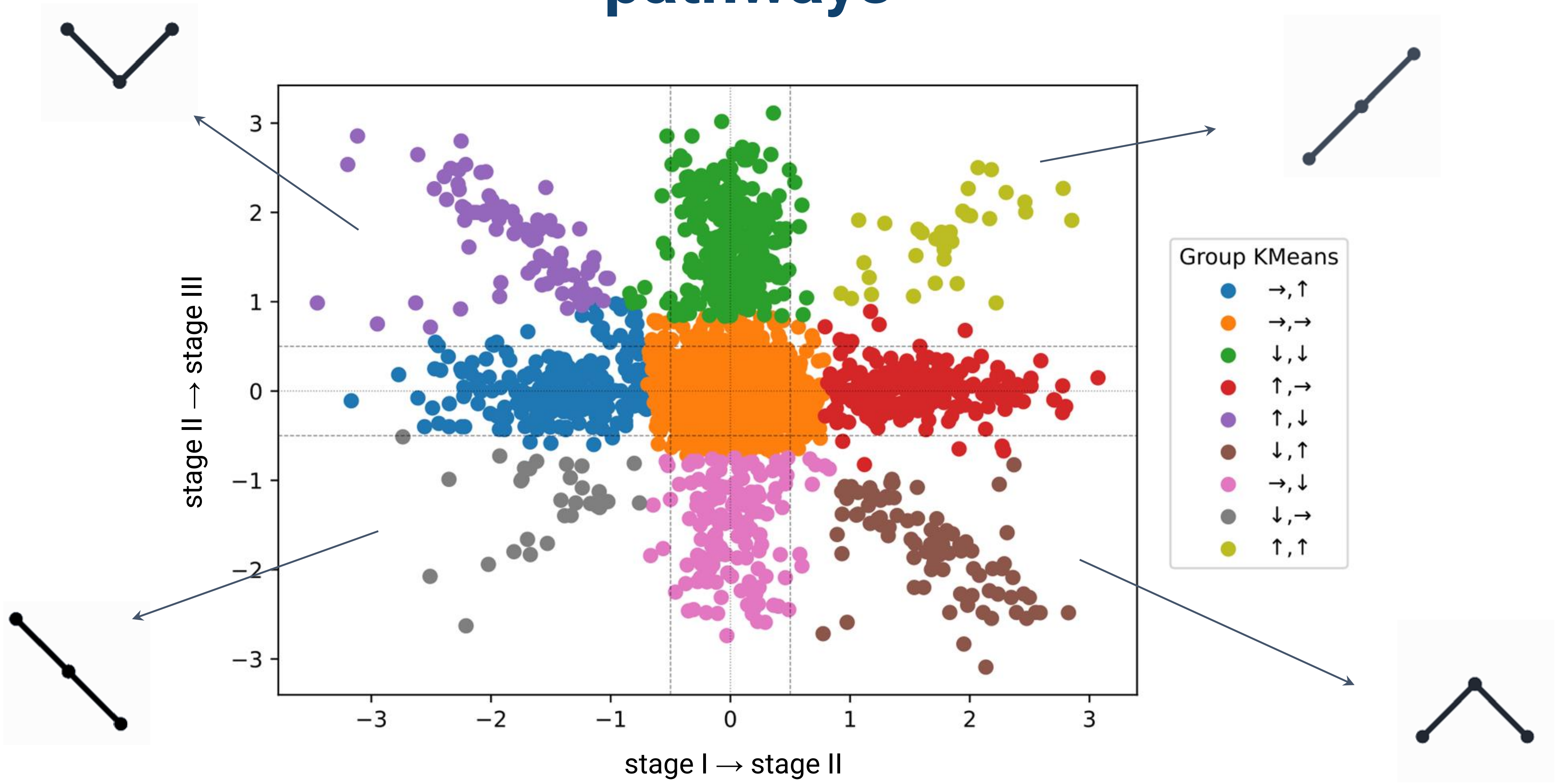


Example of a gene pathway (DNA damage repair)

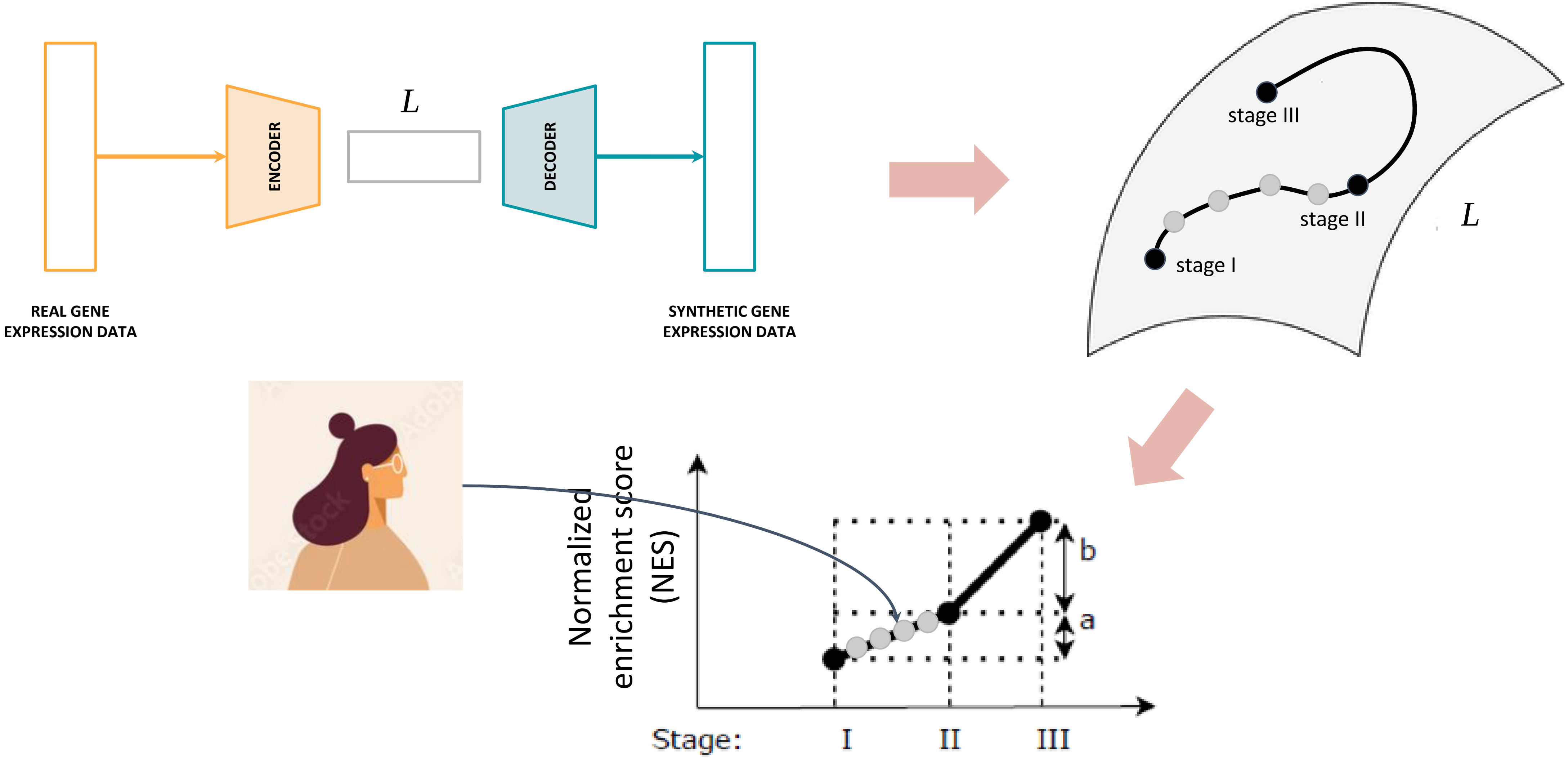
Quantification of pathways activation across the stages



Harnessing biological pathways



Generating synthetic data between stages



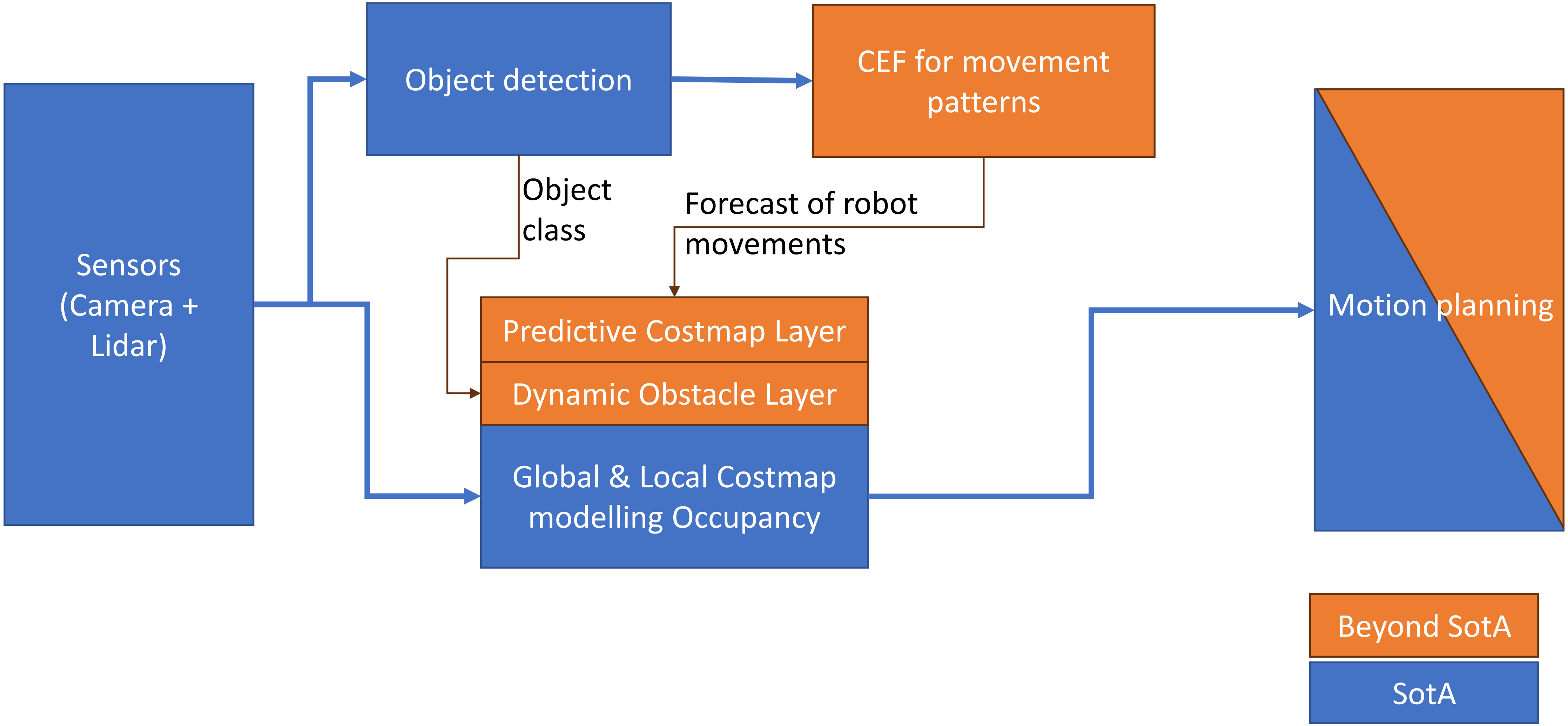
Industry 4.0 Use Case

Led by: dfki ai

Use case description

- Use case in intralogistics using autonomous mobile robots (AMRs)
- Forecasting of delays and collisions in a modular production environment

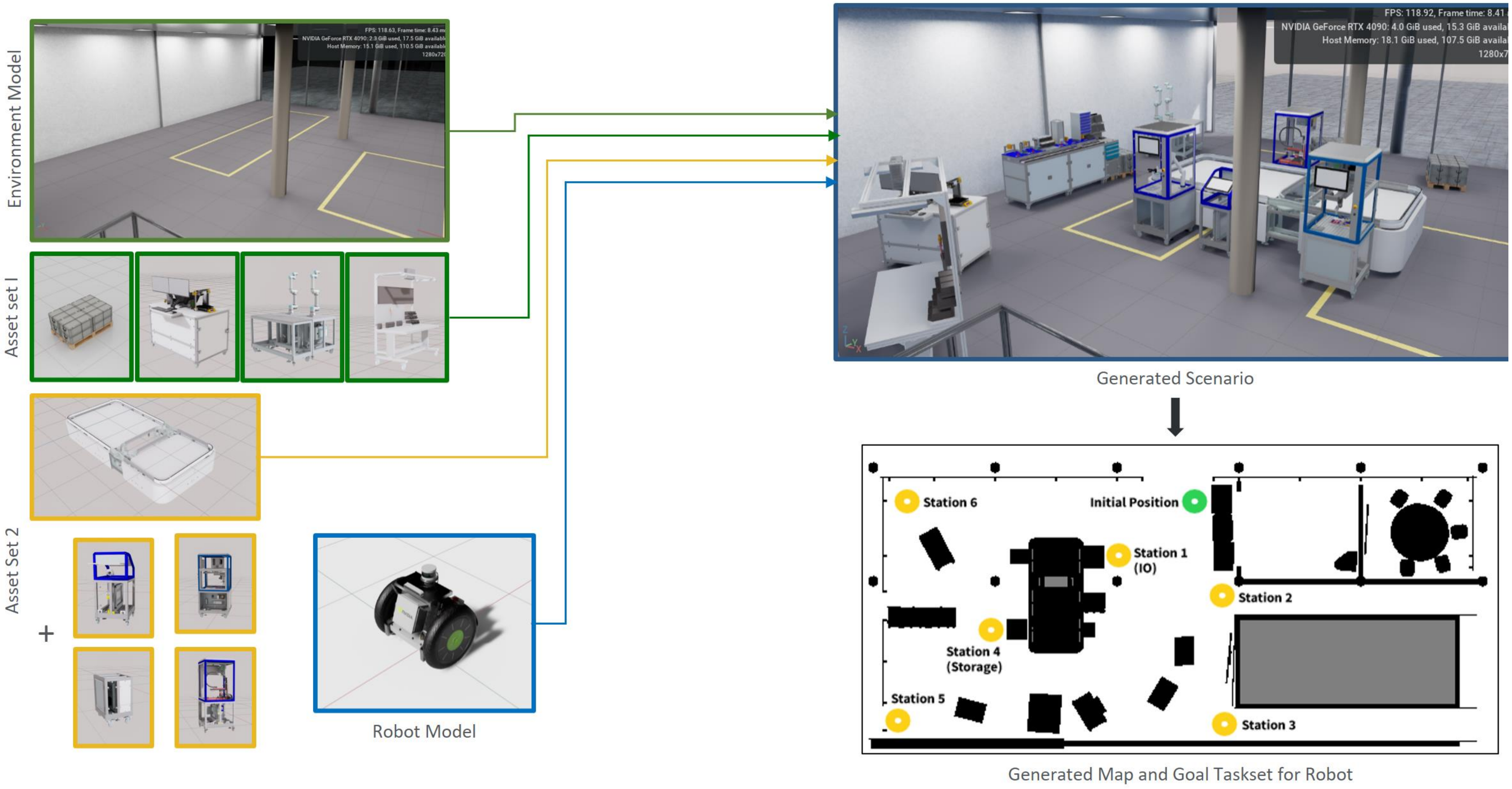
State of the art and beyond AMR Navigation



Use Case I4.0 progress towards its objectives

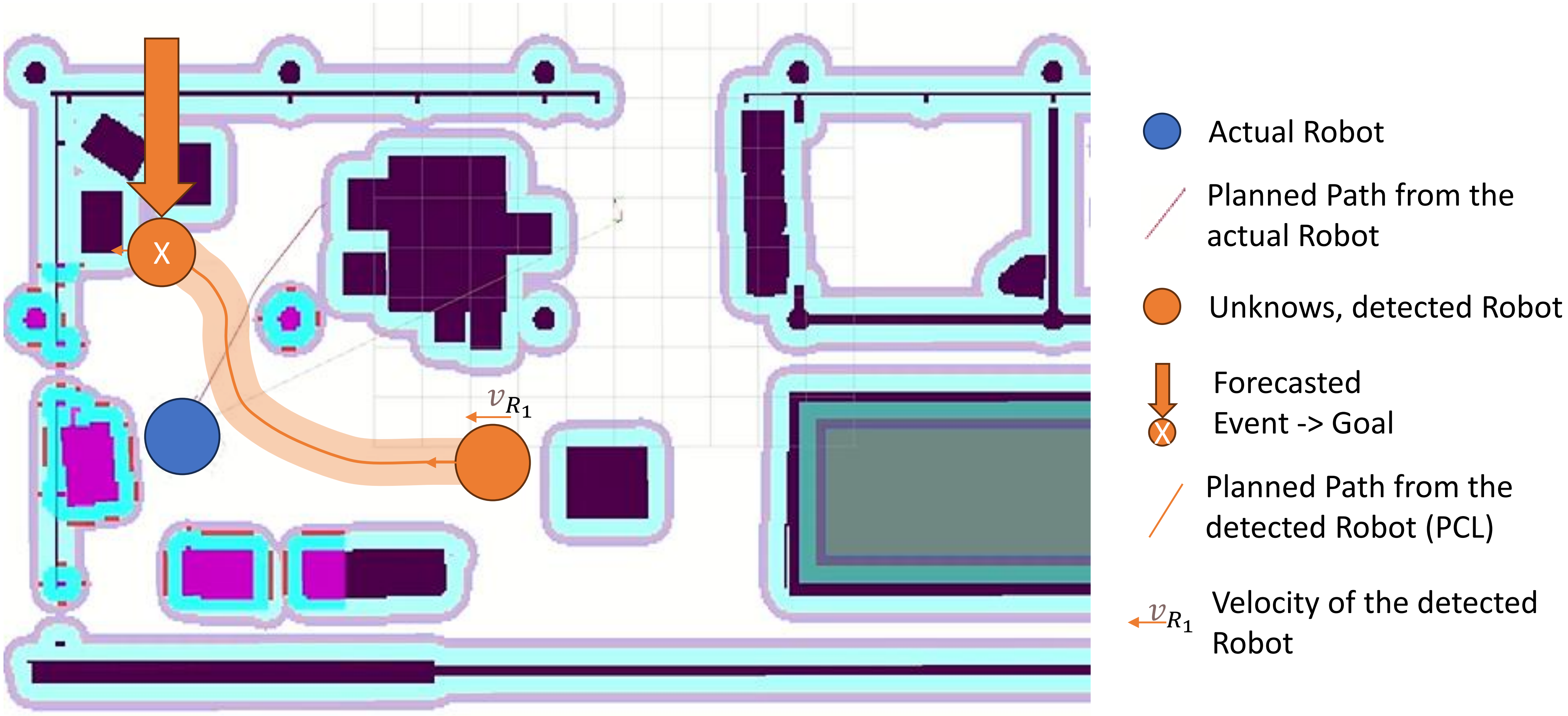
- Creation of dynamic simulation environment to create Training data for CEF
- Automated Domain Randomized Dataset creation
- Setup of state-of-the-art AMR
- Semantic & Predictive Costmap based on ROS2 standards
- MPC-based trajectory generation for AMR

Simulation Environment



Robot navigation – predictive costmap

Example forecasted event



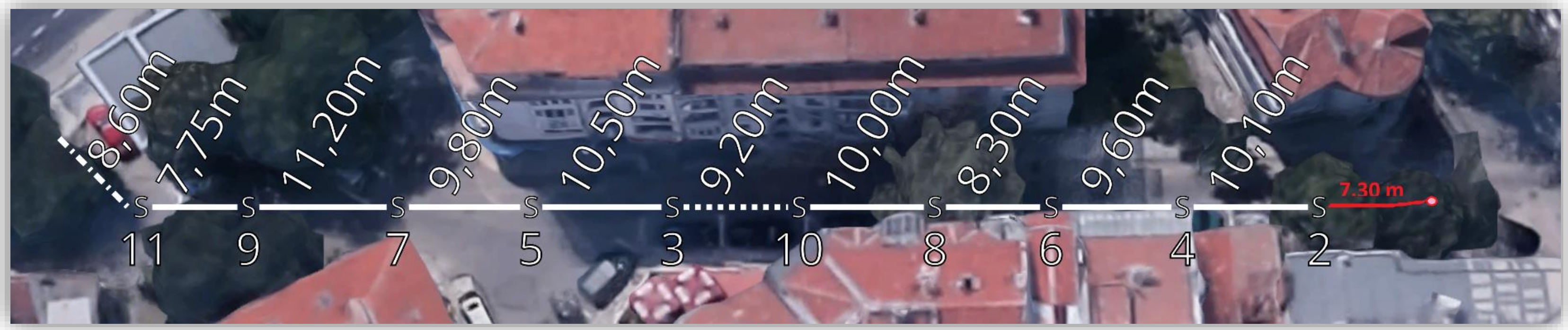
Infrastructure Life Cycle Assessment Use Case

Led by:

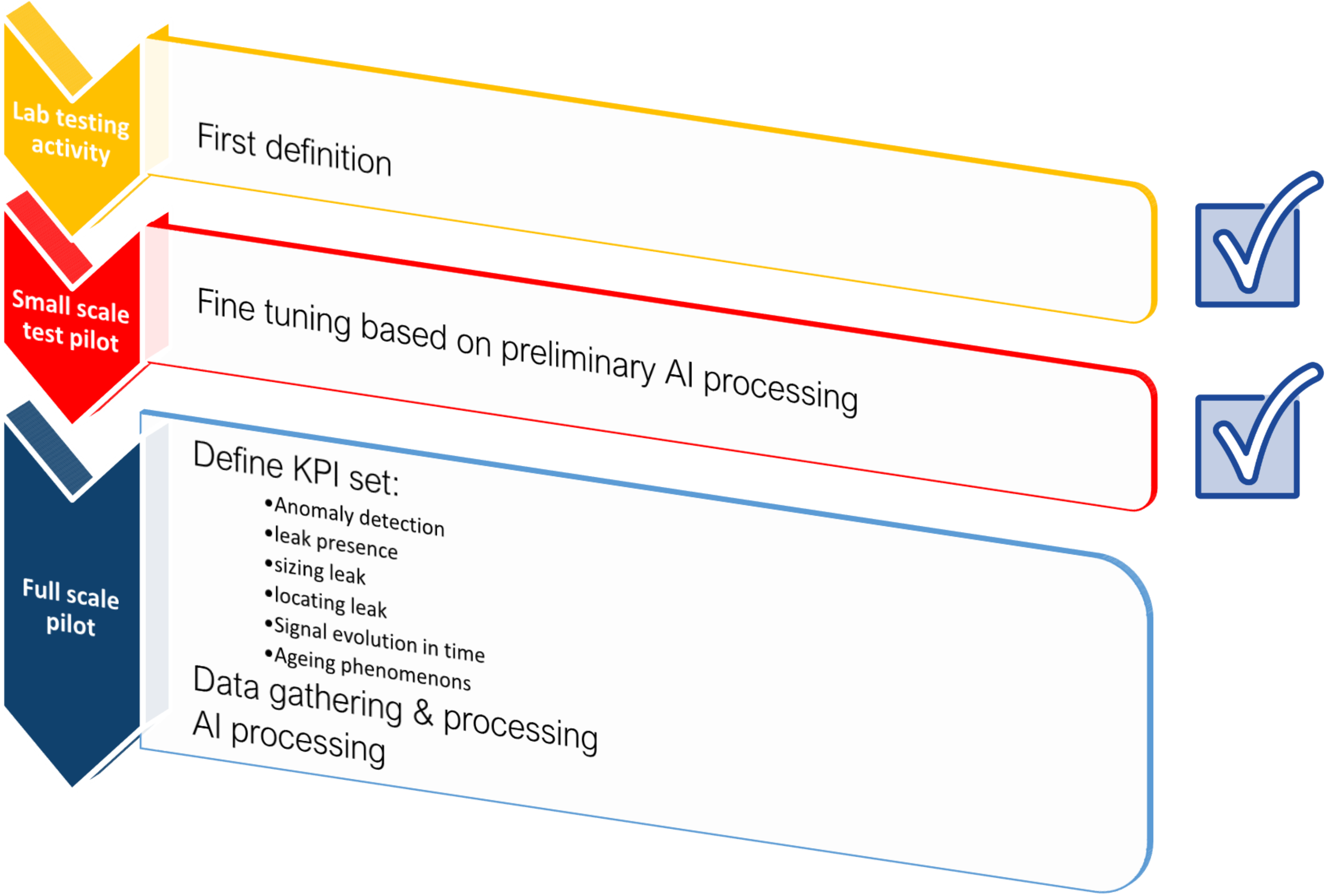
Use case description

EVENFLOW Complex Event Forecasting (CEF)
to produce accurate, reliable, and explainable LCA predictions
on Smart Pipes

- Data source: from MEMS/vibration sensors on the pipe
- Process: Intelligent Condition-Based Monitoring (CBM) (e.g., predictive maintenance) and Life Cycle Assessment (LCA)
- Target: optimizing the operational efficiency of pipes
  - State-of-the-art CBM and LCA use deep learning techniques, which predict parameters like system output flows, performance degradation, or the Remaining Useful Life of the asset (RUL)



KPI definition process



The datasets

Small scale Pilot:

1. Vibration time series from 1 sensor (event labelled) locally registered;
2. Frequency: 6,6 ksps;
3. Magnitude: 800MB on compressed CSV file;
4. Limited time frame measurements: 1 hour;
5. Different simulated leakage in distance and size.

Full scale Pilot:

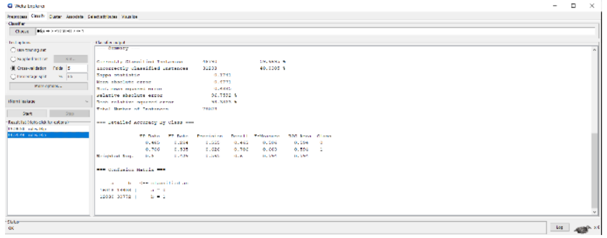
1. Vibration time series from 10 sensors (event labelled) remotely registered;
2. Frequency:1,6 ksps (each sensor-BUS main constraint);
3. Magnitude: 14MB/10min. (all sensors) in Binary format (Numpy zipped)
4. Continuous measurement: 24/7
5. Single leakage simulation.

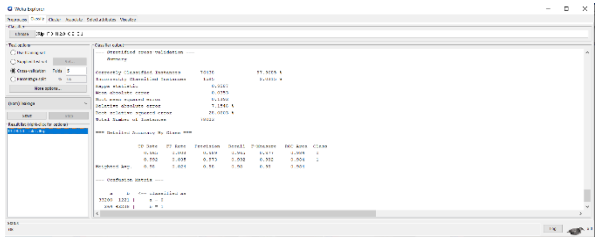
Roadmap and progress

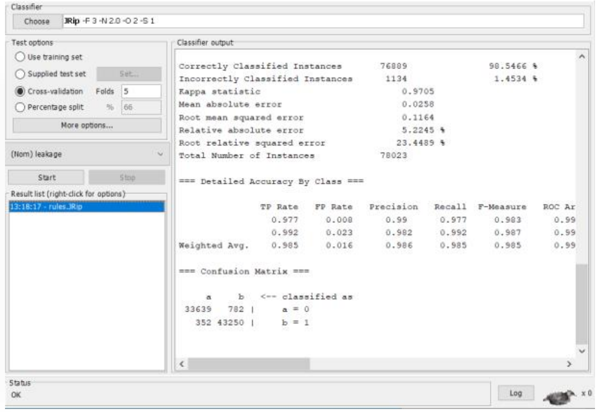
1. Anomaly detection
2. Leak presence
3. Sizing leak: under evaluation
4. Locating leak: under evaluation, through sensors interaction analysis
5. Anomaly evolution on long run: On going analysis in Combination with the expected Pipe life declared by the producer
6. Ageing phenomenons: On hold waiting #

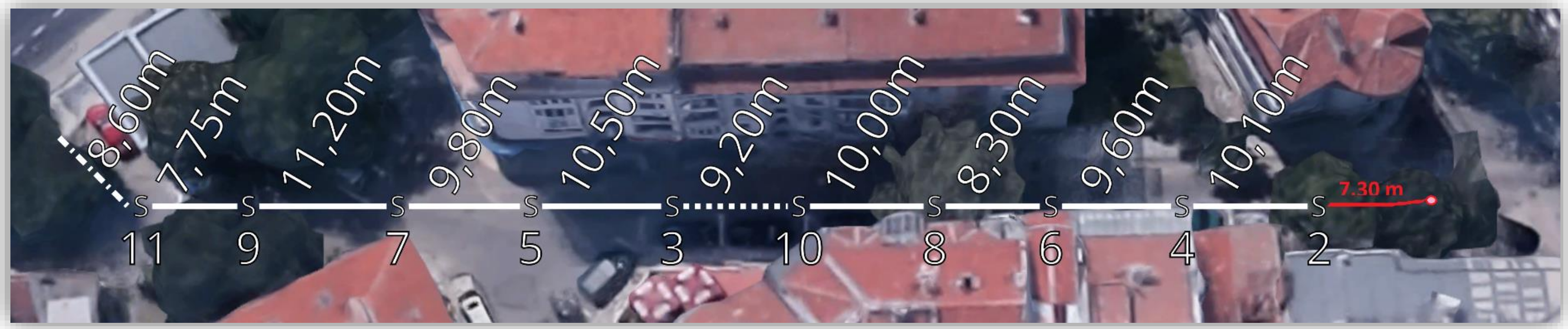
Leak presence - Problems locating where leakage is

- One sensor’s recordings analysis alone are not enough
    - Pre-processing one sensor’s recordings gives accuracy from 60% to 97%
    - Using closest sensor increases accuracy to 98.5%
    - And if we use all 10 sensors, we increase accuracy to 99.8 %









Thank you!

https://evenflow-project.eu

https://twitter.com/EvenflowProject

https://www.linkedin.com/company/evenflow-project/

https://zenodo.org/communities/evenflowproject

https://github.com/EVENFLOW-project-EU

Project Partners