HORIZON

EUROPE BRIDGING THE ATLANTIC

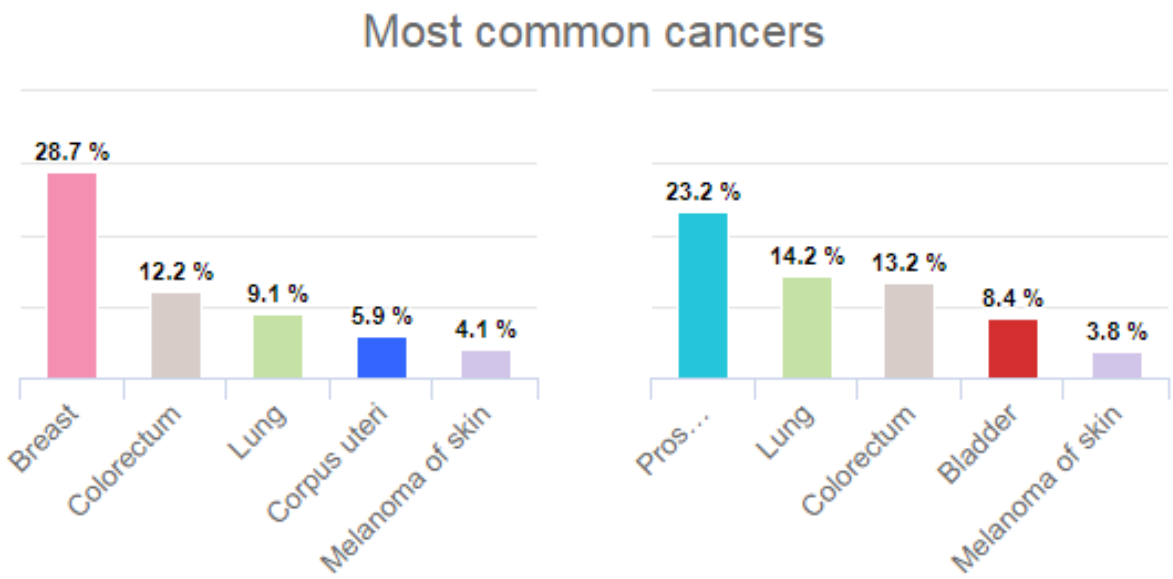
Motivation (MULTIR)

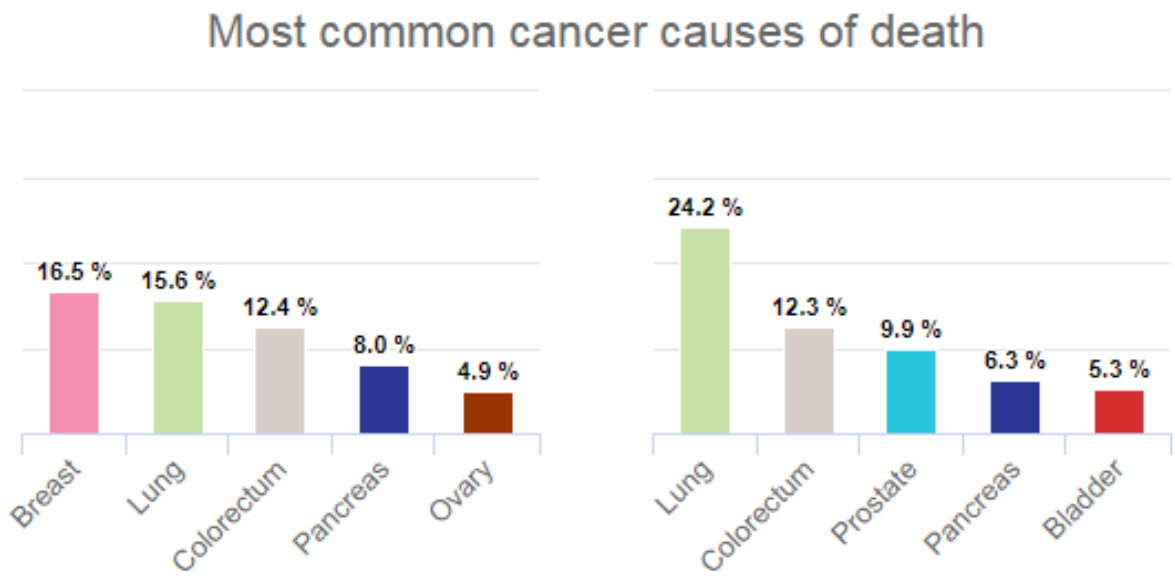
1 237 009 new cases

1 444 949 new cases

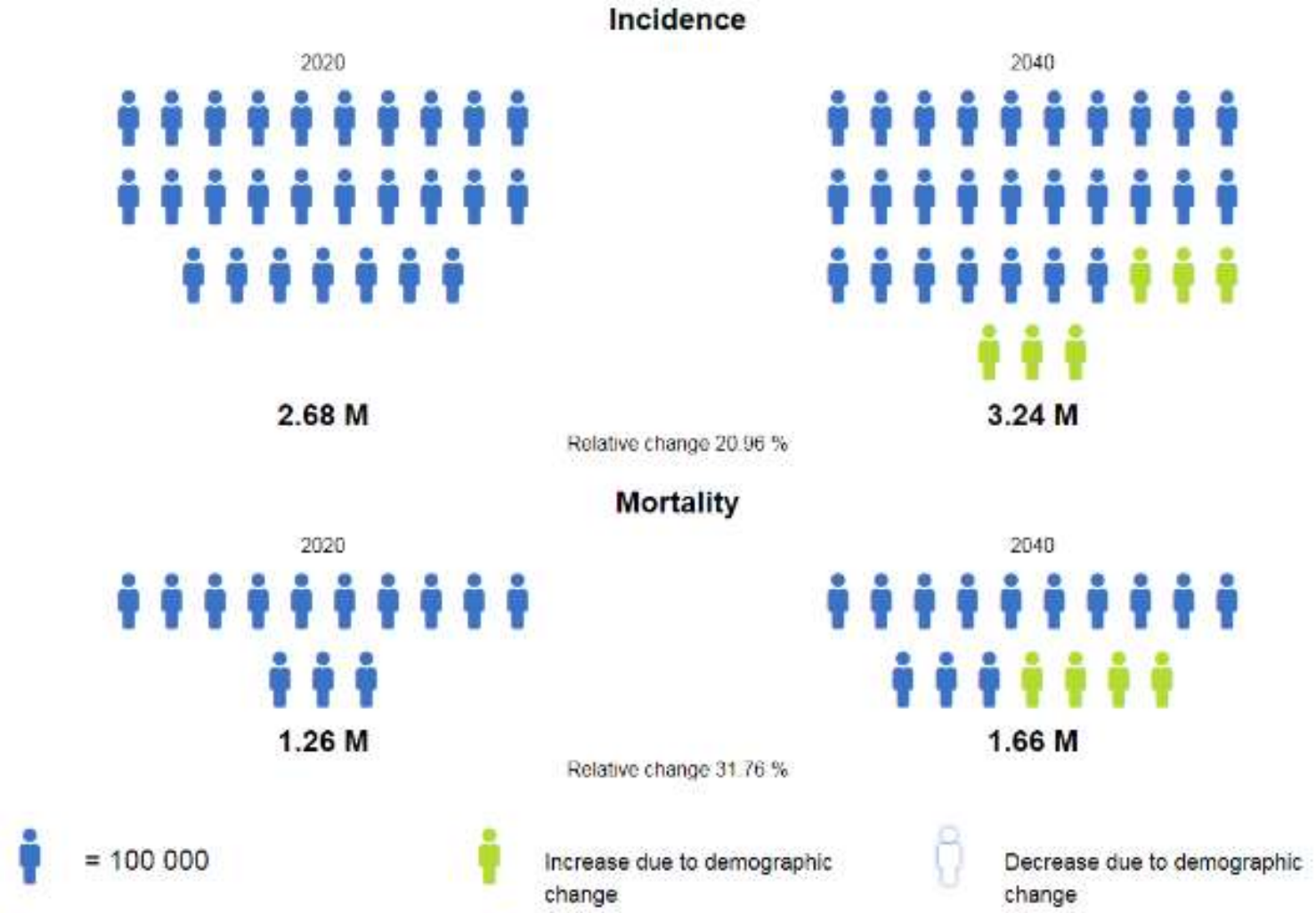
555 650 deaths

706 072 deaths

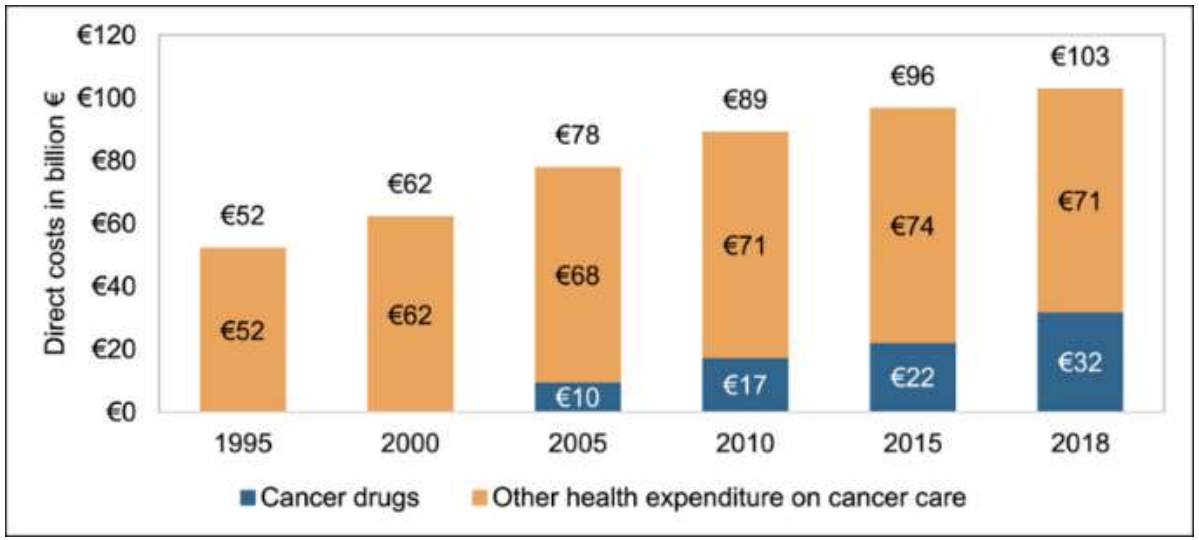


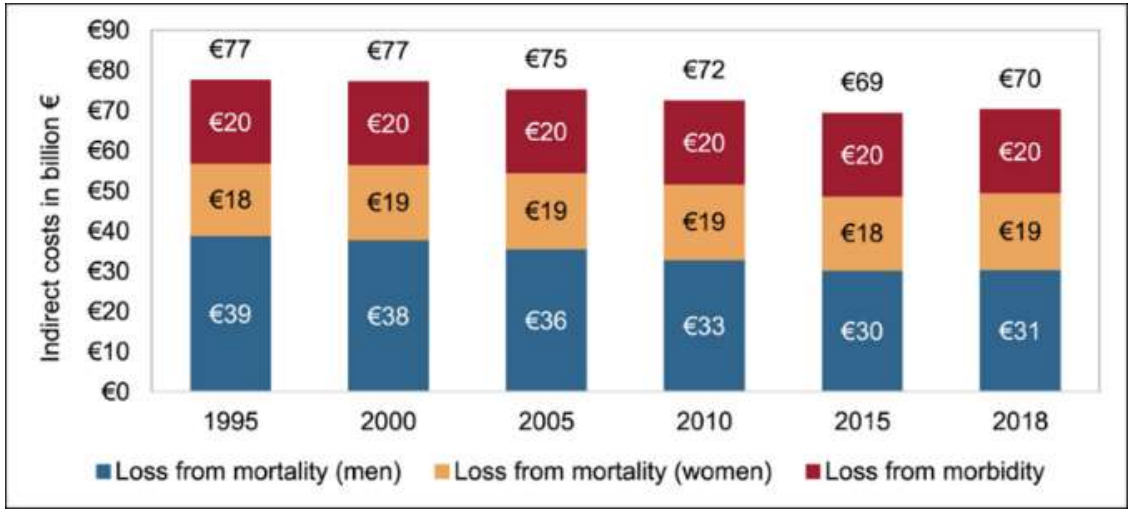


Incidence vs Mortality within EU-27

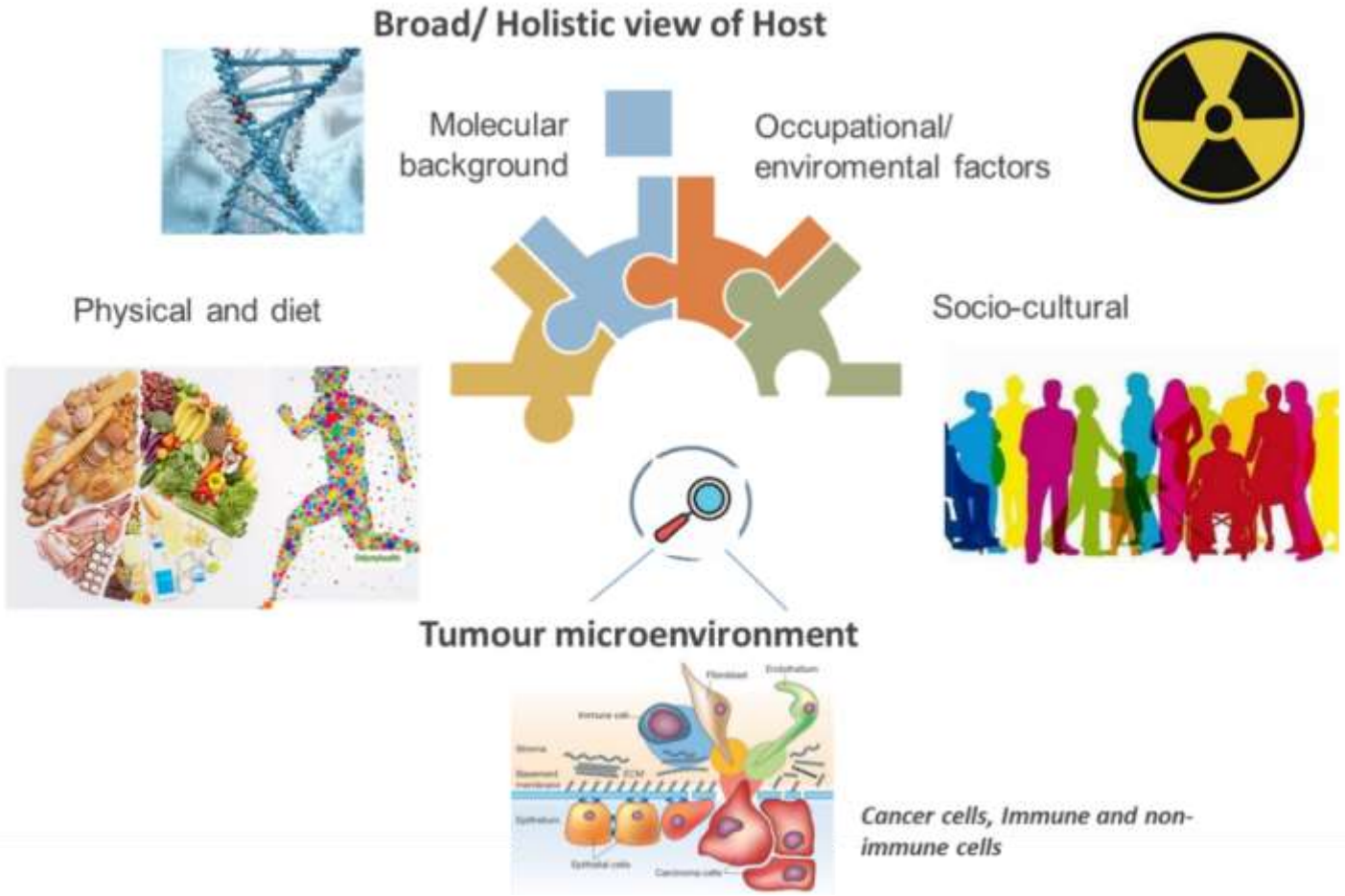


Costs of cancer in EU in 1995–2018

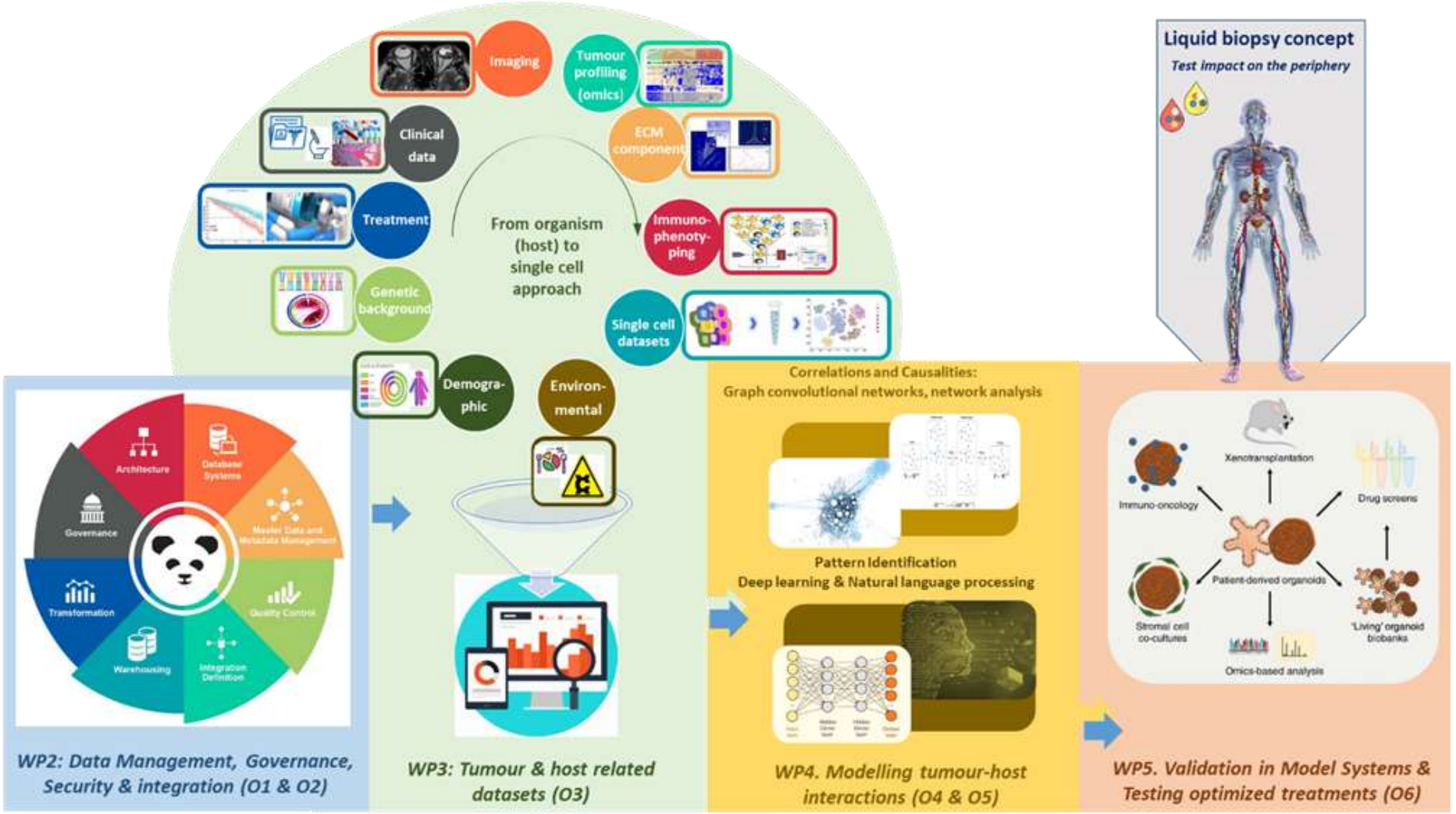




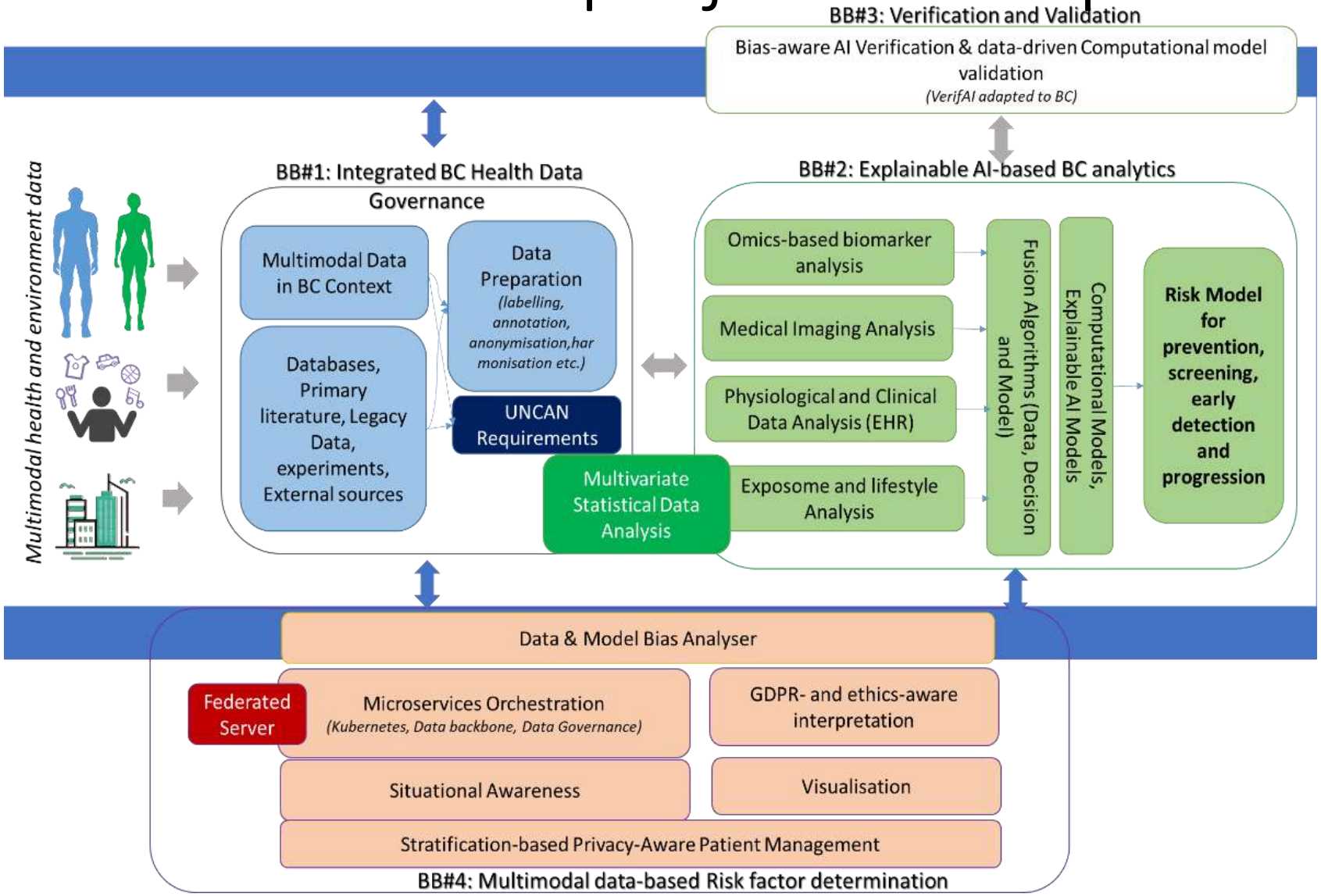
Rationale for holistic consideration of tumor-host interactions



MULTIR project concept



MULTIR – project concept



MULTIR

Funded by the European Union

Consortium

Funded by
the European Union

McGill University

Funded by the European Union

Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

McGill University (MGU)

Logan Walsh, PhD Professor, Group leader

- McGill University (MGU) is an expert in spatial proteomics and will generate and analyze highly multiplexed imaging data.
- They also have consented well over 1500 patients with various stages of lung cancer and other rare thoracic malignancies.
- The most unique and scientifically versatile aspect of our biobank effort has been the development of a living tumour bank which includes a large array of extensively characterized patients derived xenografts and organoid models.
- The biospecimen collections are accompanied by an extensive clinical annotation and comprehensive clinical grade next generation sequencing data.

Funded by the European Union

Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

WP2 Tasks

T2.1

Development of a data security architecture based on AI PANDA

T2.2

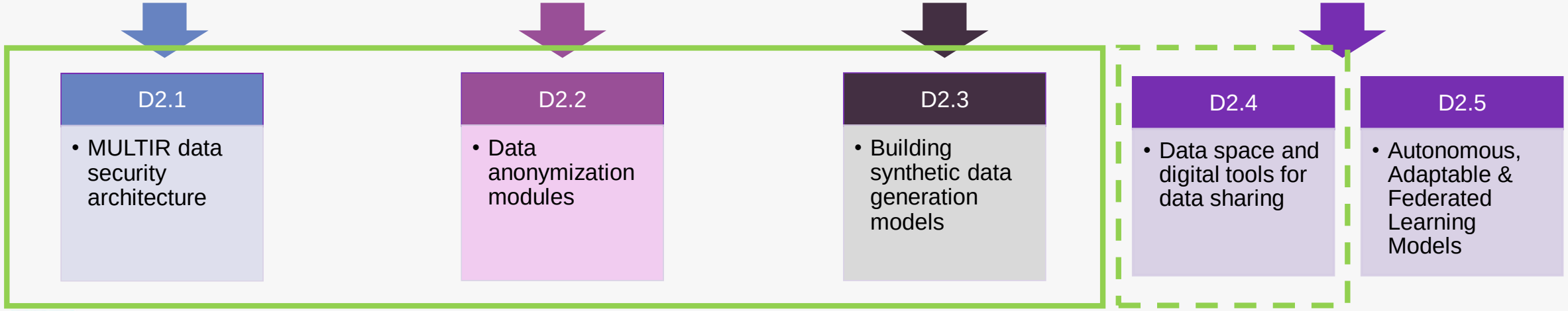
Developing the healthcare data space for data sharing and curation

T2.3

Developing AI powered digital tools for sensitive data, generation of synthetic data

T2.4

Integration of the digital tools and federated learning



Funded by the European Union

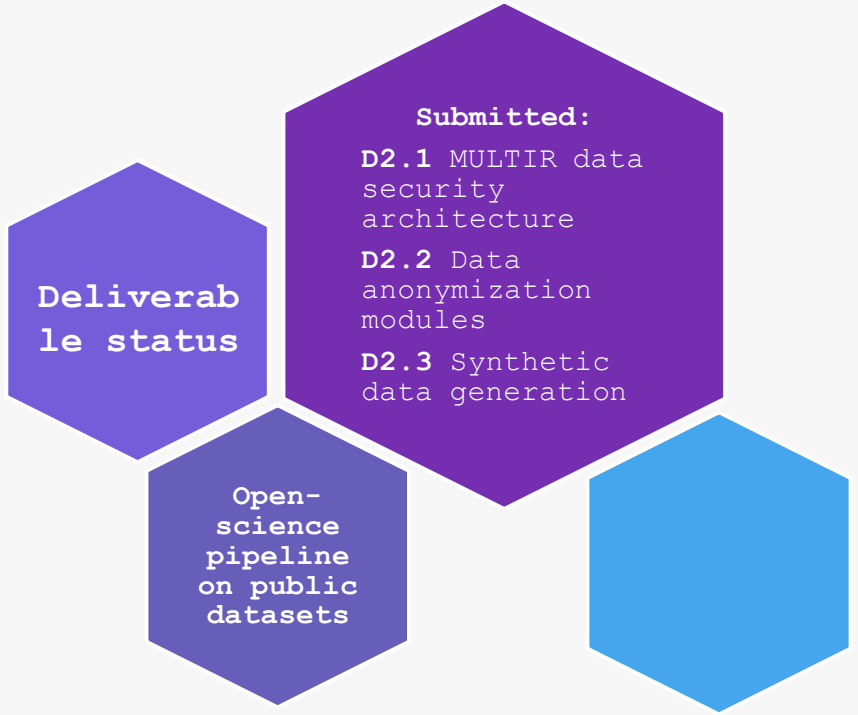
Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

WP2 Review (to date...)

- **Scope:** Build and operate the MULTIR healthcare data space underpinning all downstream analytics.
  - **T2.1 Security architecture (ATMC):** AI PANDA-based, defense-in-depth, IAM, encryption, secure pipeline, audit.
  - **T2.2 Data space for sharing & curation (ATMC):** IDSA-compliant connector, policy-driven data handling, metadata catalog.
  - **T2.3 Synthetic data generation (ATMC+MOS):** Privacy-preserving data for development/validation

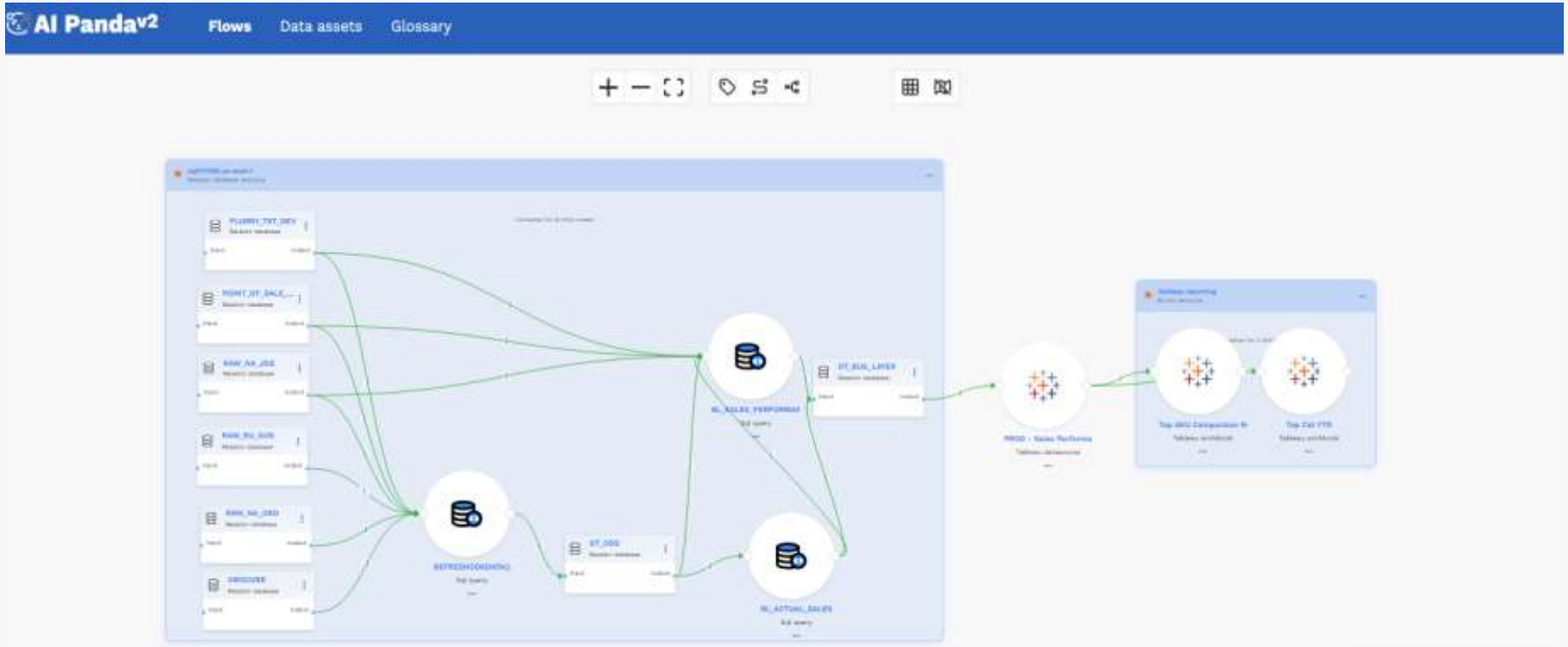
Operating principles: GDPR-by-design, ISO 27001 alignment, API-first, FAIR data, EOSC/UNCAN.eu interoperability.

Execution mode M1–M18: Isolated ATMC environment, no real cohort data processed (JCA/DPA not yet signed).



Funded by the European Union
Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

Summary of Activities



D2.1 (Security architecture):

Deployed **AI PANDA** in isolated ATMC DC (Kubernetes).

IAM: Azure AD integration + AI PANDA user/role model, MFA **to-be** enforced.

Encryption: TLS in-transit, AES-256 at rest, secrets management; key rotation.

Network security: Segmented namespaces, allow-lists, hardened ingress.

Monitoring & audit: Centralized immutable logs, alerting rules,

Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

Summary of Activities

D2.1 (Security architecture):

Deployed **AI PANDA** in isolated ATMC DC (Kubernetes).

IAM: Azure AD integration + AI PANDA user/role model, MFA **to-be** enforced.

Encryption: TLS in-transit, AES-256 at rest, secrets management; key rotation.

Network security: Segmented namespaces, allow-lists, hardened ingress.

Monitoring & audit: Centralized immutable logs, alerting rules.

D2.2 (Anonymization modules):

Metadata-level pseudonymization/generalization/suppression ruleset.

Policy engine integrated with AI PANDA profiling/classification.

IDSA-compliant connector design for future partner exchanges.

D2.3 (Synthetic data):

Built **Gaussian copula** pipeline, validated on public peptidomics datasets.

Benchmarks: PCA/MDS overlap, SVM parity, preserved clinical correlations (e.g., eGFR).

MIT-licensed repo, joint manuscript with MULTIR members (MOS, etc.).

Governance & QA:

ISO-aligned SOPs, **internal audits,** DR tests, evidence packs compiled.

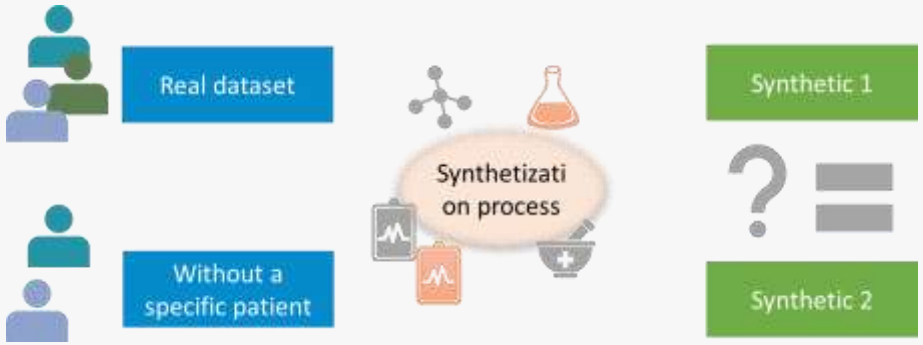
Constraint:

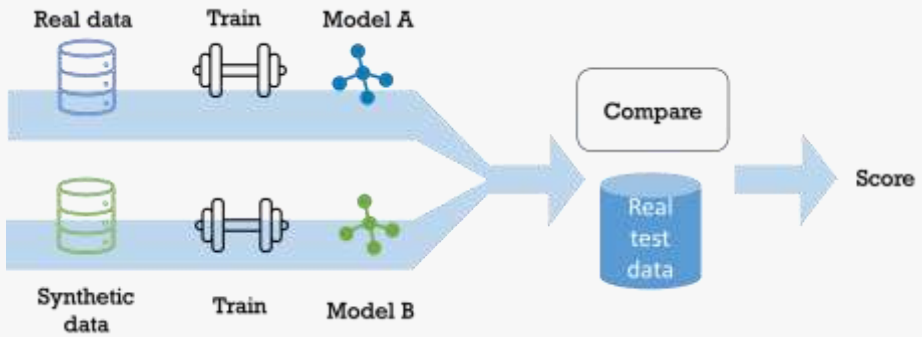
No internal cohort ingestion until **JCA/DPA** are signed (technical onboarding runbook prepared).

Funded by the European Union

Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

Results Achieved — Synthetic Data Validation (D2.3)





Why synthetic now:

- Enables development/validation **without PHI** while JCA/DPA are pending, reduces re-ID risk.

Method:

- **Gaussian copula** for high-dimensional tabular/omics; stable, interpretable dependence modeling.

Co-development of copula-based synthetic pipeline

Validation leadership on **public datasets**:

- peptide data (MS), additional diseases for generalization.

Evidence produced:

- PCA/MDS cluster preservation
- **SVM performance parity** (train-synthetic/test-real and vice versa)
- **Clinical association** preservation (e.g., eGFR).

Reproducibility:

- Contributed scripts, experiment configs, figure generation; coordinated analysis narrative.

Dissemination:

- Co-authored manuscript
- open-science alignment
- readiness to extend to MULTIR cohorts post-JCA/DPA.

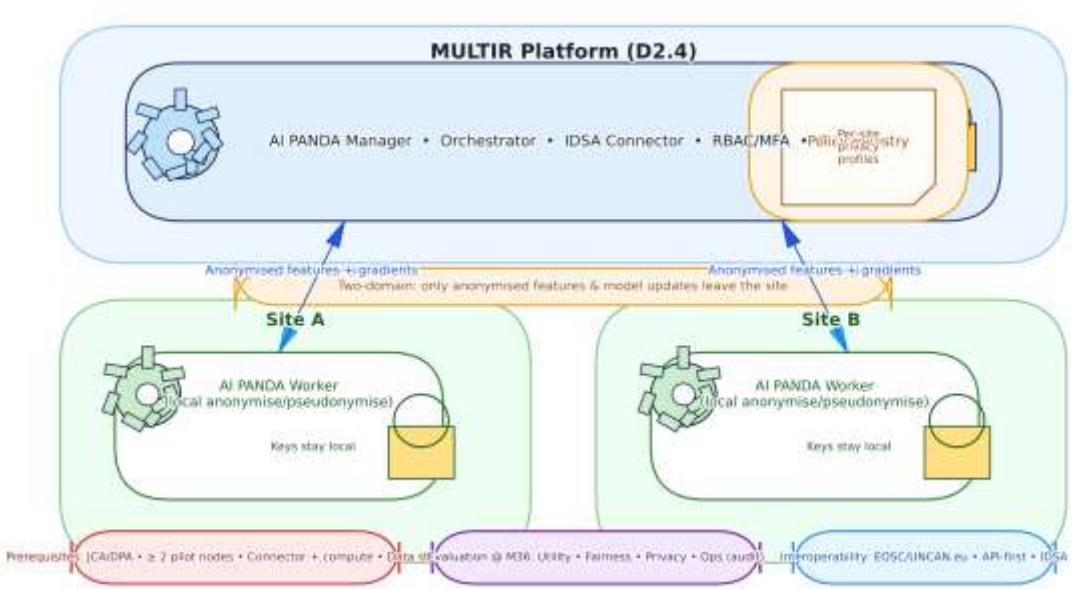
Feedback loop:

- Provided requirements for rare-class handling; roadmap to hybrid deep generative components.

Funded by the European Union

Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

D2.5 Autonomous, Adaptable & Federated Learning Models (M36)



Data access pattern (with Data Shield/Data Broker)

- **Manager/Worker** design: the **ATMC AI PANDA Manager** deploys a **Worker** at each site to pre-process and **anonymise/pseudonymise** locally; only anonymised features and model gradients/updates leave the site.
- Aligns with the **two-domain approach** (keys remain local) and supports sharing only anonymised outputs to the MULTIR platform.

Integration & dependencies

- **Built on D2.4:** uses the governed **IDSA connector**, site identities/attestations, and platform RBAC/MFA.
- **Interoperability:** API-first; prepared for **EOSC/UNCAN.eu** endpoints once cohorts are live; metadata/ontology alignment inherited from D2.4.
- **Prerequisites:** JCA/DPA signed; at least two pilot nodes with connector + basic compute (CPU/GPU) and a nominated data steward.

Evaluation & acceptance (evidence at M36)

- **Utility:** cross-site AUC/F1/PR-AUC; calibration (ECE); survival concordance for time-to-event tasks.
- **Fairness & robustness:** stratified performance (site/tumour/type), sensitivity to missingness, and shift tests.
- **Privacy/security:** DP ε-budgets documented; penetration of secure-agg; audit trail coverage for every round.
- **Ops:** successful multi-round federations with **automatic recovery**, reproducible model artefacts, and signed **model cards**.

Funded by the European Union

Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

Data Shield

Goal:

- process data where it resides,
- minimize data movement
- share only GDPR-compliant outputs

Manager/Worker architecture:

- ATMC AI PANDA Manager orchestrates workload
- Worker performs local anonymization/pseudonymization

Configuration-driven:

- per-site privacy/security policy
- workers pulled on-demand
- ephemeral and auditable

Outcome:

- standardized, repeatable local processing with platform-only receiving anonymized results

Funded by the European Union

Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

Cross-WP links (WP2 to WP3 & WP4)

WP3:

- method development consumes D2.4 APIs
- synthetic data from D2.3 supports method pre-validation

WP4:

- statistical associations (D4.2) rely on curated datasets and audit trails from D2.4
- publication-grade provenance

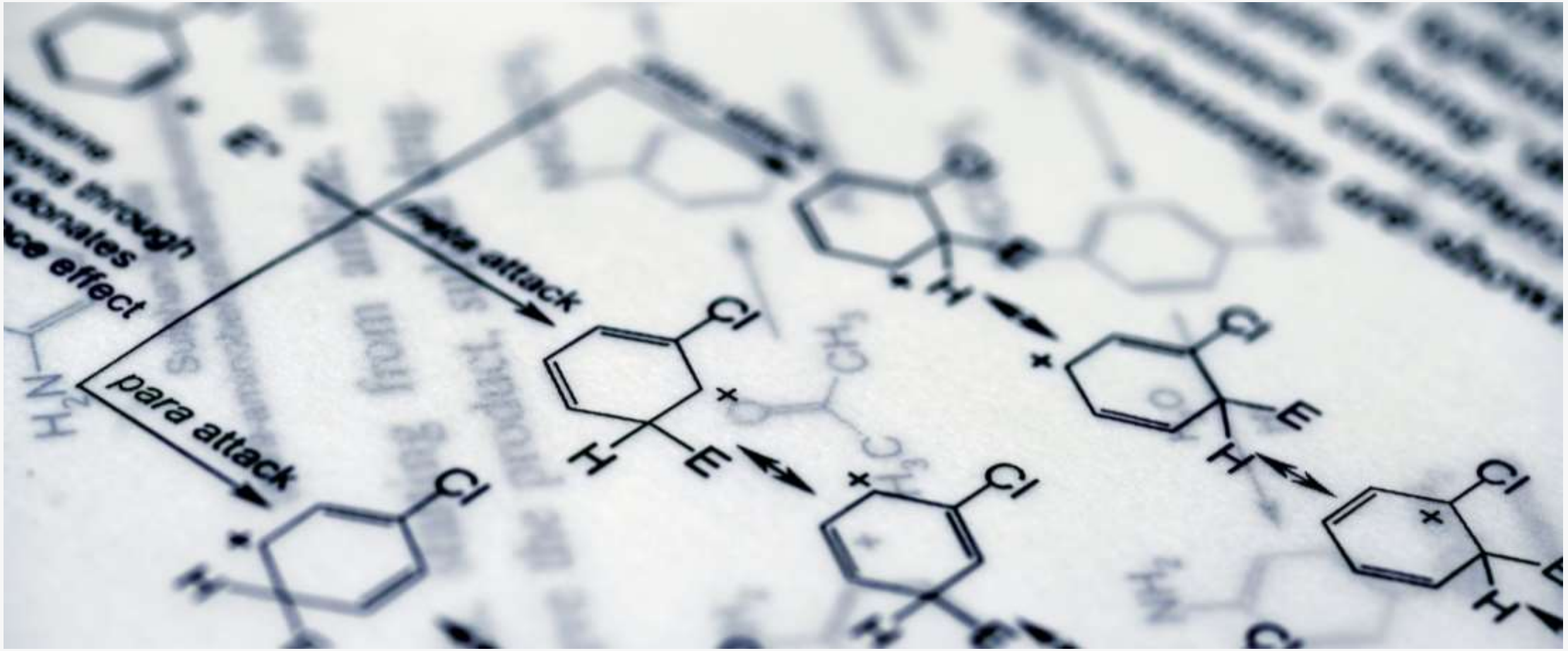
Feedback loops:

- WP3/WP4 findings refine D2.2 policies and D2.5 federated model requirements

Funded by the European Union

Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

WP3 - Clinical, pathology, molecular, immunology and imaging data: integration of existing and newly acquired datasets



Funded by the European Union

Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

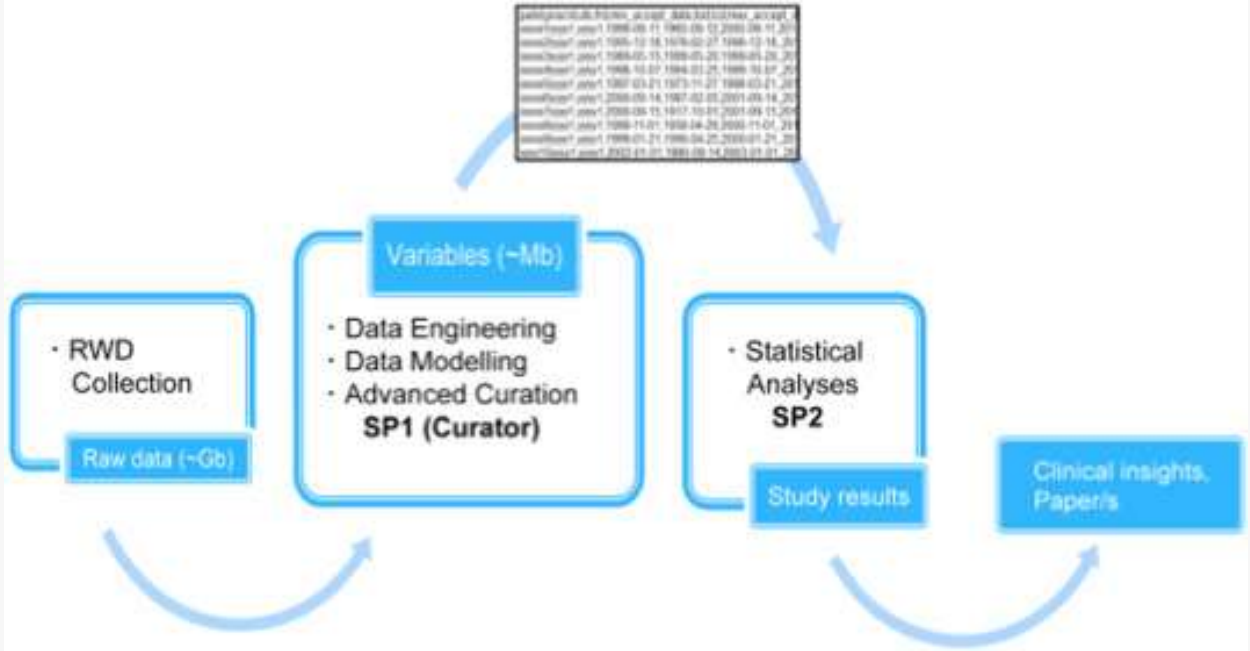
Description
T3.1 - Collection and curation of clinical datasets, treatment and follow-up datasets for lung cancer, bladder carcinoma and melanoma patients [M3 – M12] [Lead – NUH, Participants –HORG, UK-E, UZ, NUH, MGU, MOS, BRFAA, RadboudUMC, AU)
T3.2 - Collection and curation of broader host-related factors (including demographics, germline DNA variation) [M3 – M22] [Lead - RadboudUMC, Participant: AU, HORG]
T3.3 - Collection and curation of tumour molecular datasets (genomics, transcriptomics, epigenomics, proteomics datasets) [M3 – M22] [Lead - BRFAA, Participants AU, UK-E, UZ, HORG, MGU, RadboudUMC, MOS]
T3.4 – Collection and curation of immunophenotyping datasets including immunofluorescence and multiplexed mass cytometry[M3 – M22] [Lead - MGU, Participants AU, UK-E, HORG]
T3.5 – Collection and curation of ECM related data and peripheral immune-related metabolic datasets [M3 – M22] [Lead -MOS, Participants AU]
T3.6 – Integration of existing literature data on known molecular, immunologic and broader host-related factors for BC, lung cancer and melanoma [M3 – M22] [Lead – RadboudUMC, Participants AU, UK-E, HORG] [M3 – M22]
T3.7 - Single cell datasets and newly acquired data to supplement fragmented data [M3 – M24] [Lead - UZ, Participants –UK-E, MGU]
T3.8 – Imaging datasets and digital pathology assessment [M3 – M22] [Lead - UVIE, Participants –HORG, INS, NUH, AU, UK-E]

Funded by the European Union

Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

MGU Contribution to T3.1

Collection & Curation of Clinical Datasets



Provides **clinical datasets** from lung cancer, bladder carcinoma, and melanoma patients.

Participates in **clinical data curation,** working with clinical partners to:

- Define **relevant clinical endpoints.**
- Support **age- and sex-based stratification.**

Works with AI teams to align MGU's clinical data with:

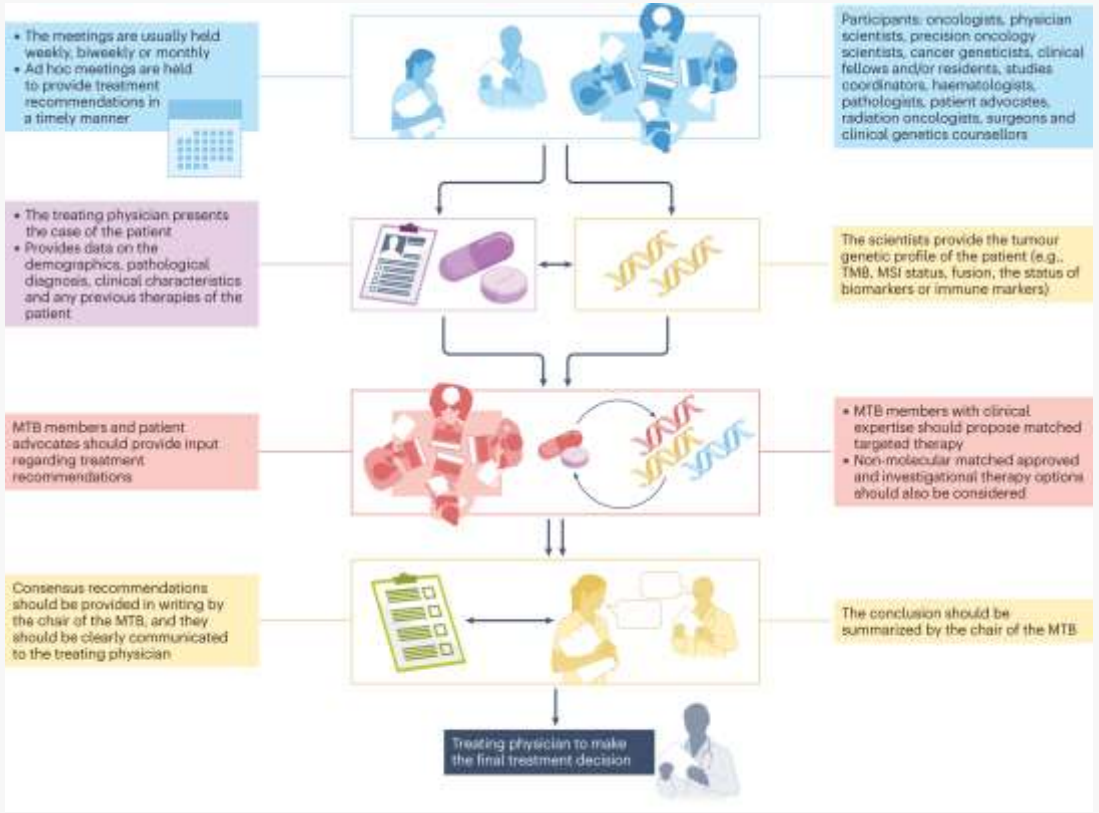
- **Data space integration formats** (ATMC, UCD).
- **Input formats** required for **AI-based and network analyses.**

Funded by the European Union

Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

MGU Contribution to T3.3

Collection & Curation of Tumour Molecular Datasets



- Provides **tumour molecular datasets** including:
  - **Genomics**: WES/WGS data for mutational landscape, CNVs, epigenetic modifications.
  - **Epigenomics**: DNA methylation profiles (Illumina-based).
  - **Transcriptomics**: Contributes datasets to be **remapped and requantified** using MULTIR-specific pipelines.
- Ensures MGU-generated molecular data adheres to **normalization and quality-control** procedures required for integration with consortium datasets.

Funded by the European Union

Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

Contribution to T3.4

Immunophenotyping Datasets

Immunofluorescence (MGU-led):

- Performs **multistaining and multiplexed immunofluorescence** to characterize immune infiltration:
  - Tumour periphery vs. tumour core.
  - Spatial heterogeneity of immune cell subsets.
- Uses **RNAscope** to identify functional subpopulations of **T cells, B cells, plasma cells, macrophages.**

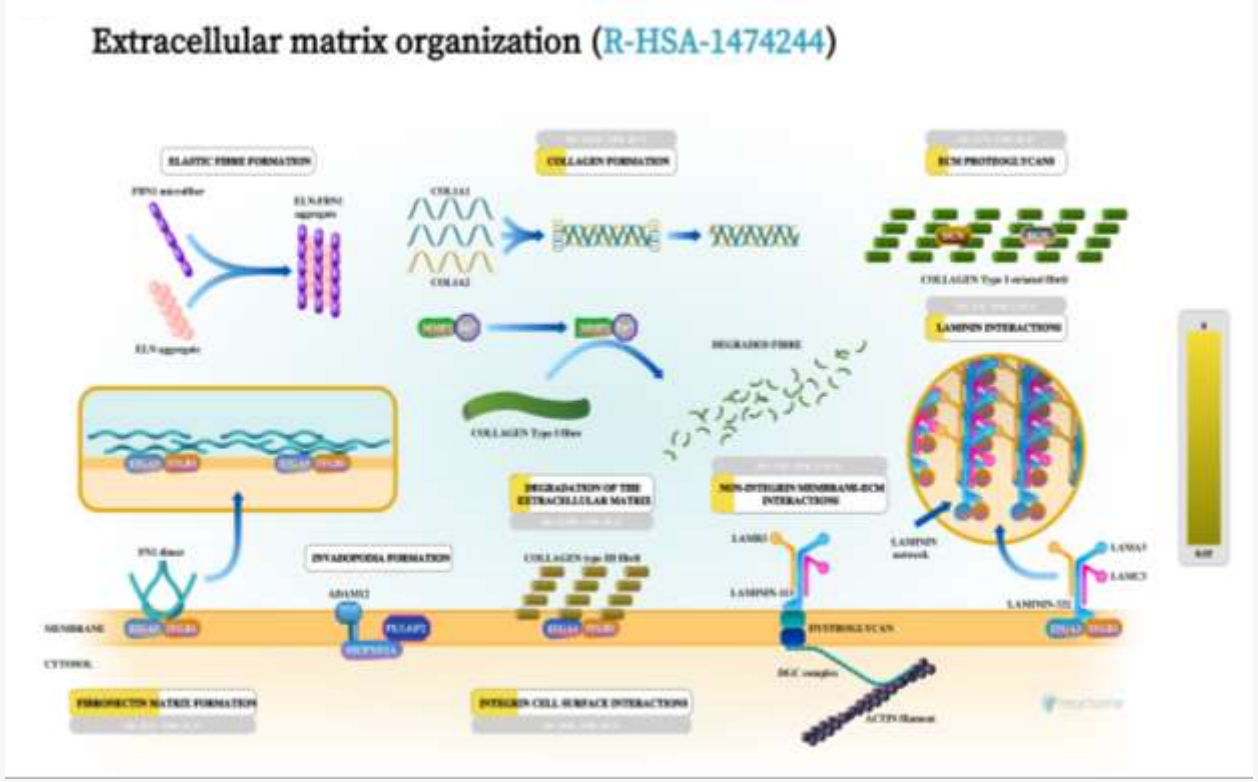
Imaging Mass Cytometry (IMC):

- Develops and applies **innovative analysis pipelines** providing:
  - Fully automated **ML-based cell segmentation.**
  - **Lineage assignment** and phenotype identification.
  - **Cell-cell interaction mapping.**
  - Detection of **cellular neighbourhoods/communities** with spatial resolution.
- Integrates IMC data from partners (AU, UK-E) into a unified spatial immunophenotyping framework.

Funded by the European Union

Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

Contribution to T3.7 Single-Cell & Newly Acquired Datasets

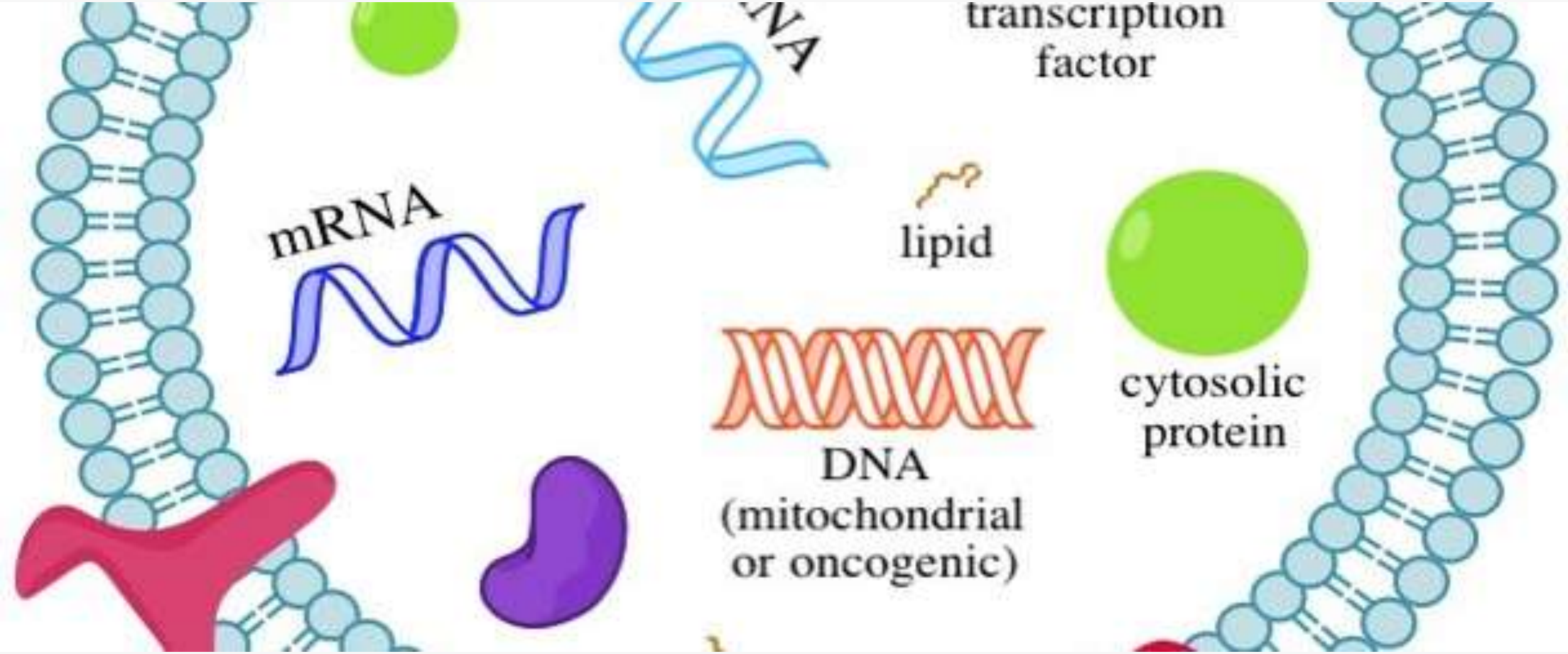


- Contributes to **single-cell immune repertoire studies** in collaboration with UZ and UK-E.
- Supports:
  - Preparation and integration of **TCR/BCR sequencing data.**
  - Quality assessment and downstream analysis of immune-profiling outputs.
- Applies internal expertise to help harmonize **gene expression** and **immune repertoire** single-cell datasets for use in MULTIR.
- Ensures that MGU's single-cell datasets meet consortium-level standards for:
  - Data preprocessing
  - Metadata structure
  - Compatibility with MULTIR analytic endpoints

Funded by the European Union

Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

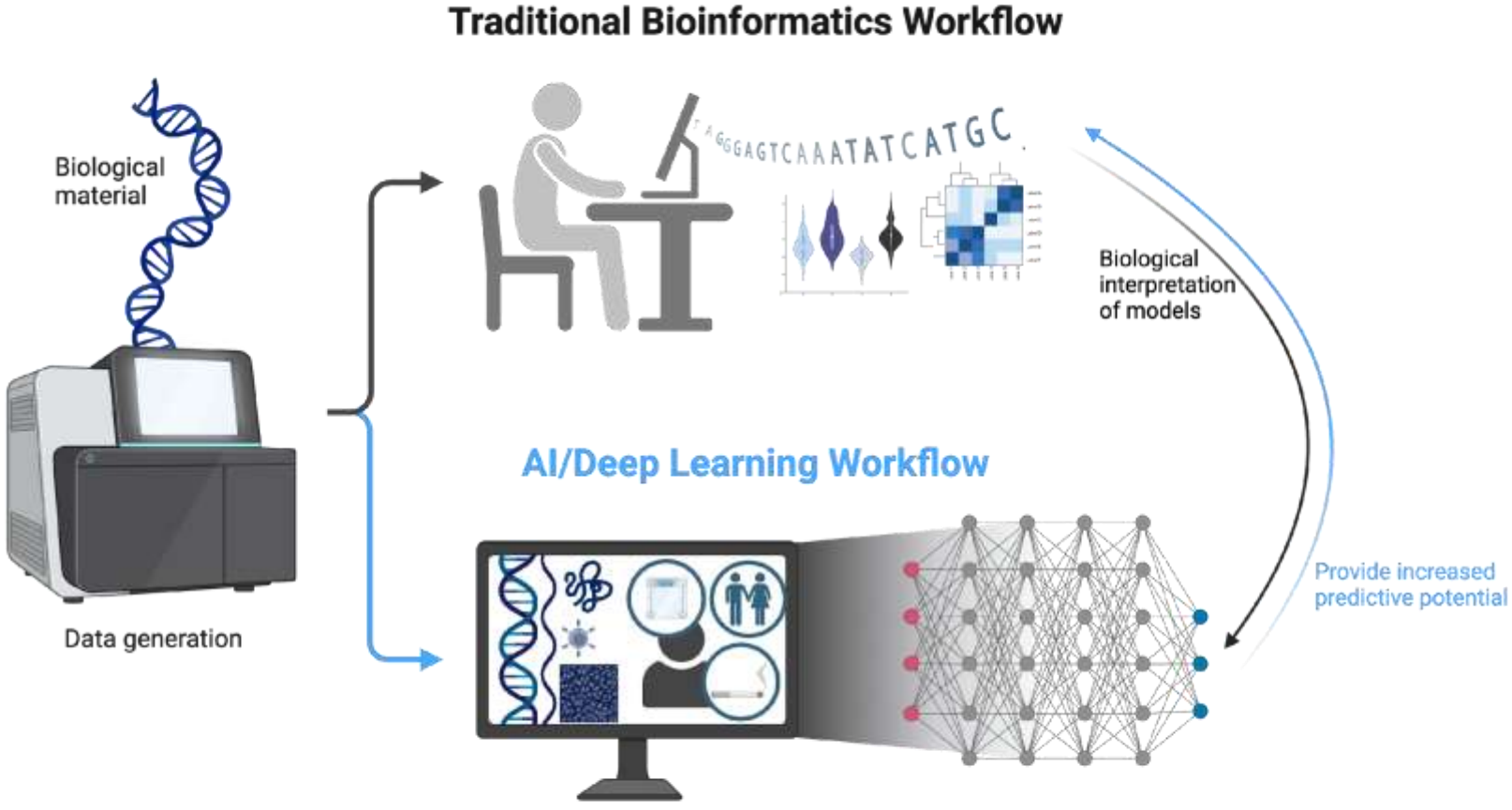
WP4 – Modelling tumour host interactions and response



Funded by the European Union

Funded by the European Union under GA No. 101136926. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

Close interplay between T4.2 and tasks regarding AI-based modelling (T4.3-T4.7)



Funded by the European Union

Objective: Identify and analyze statistical correlations between various factors at both single and multiple tumor levels.

Methodology:

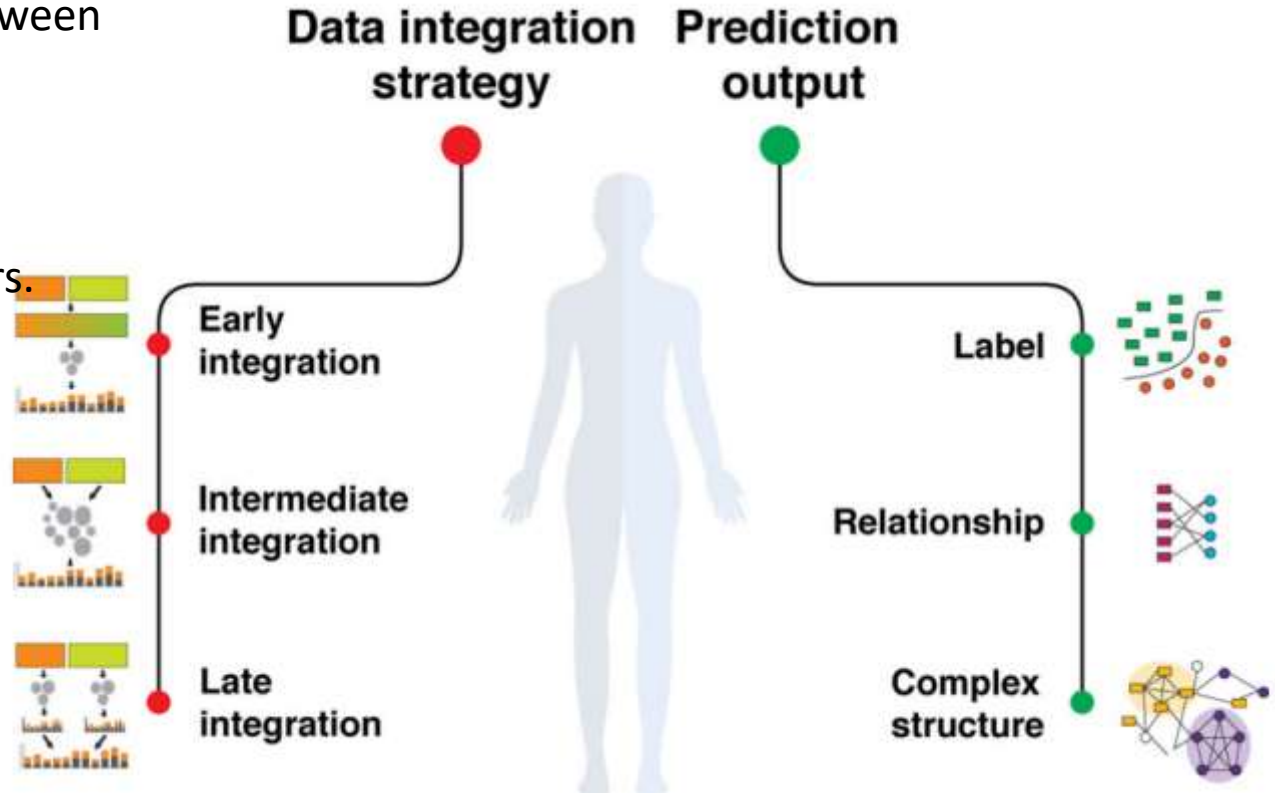
- Collection and analysis of data from individual tumors.
- Comparative analysis across multiple tumor types.

Integration with Other Tasks:

- Utilizes data from WP2 and WP3 for analysis.
- Contributes findings to WP5 for model validation

Expected Outcome:

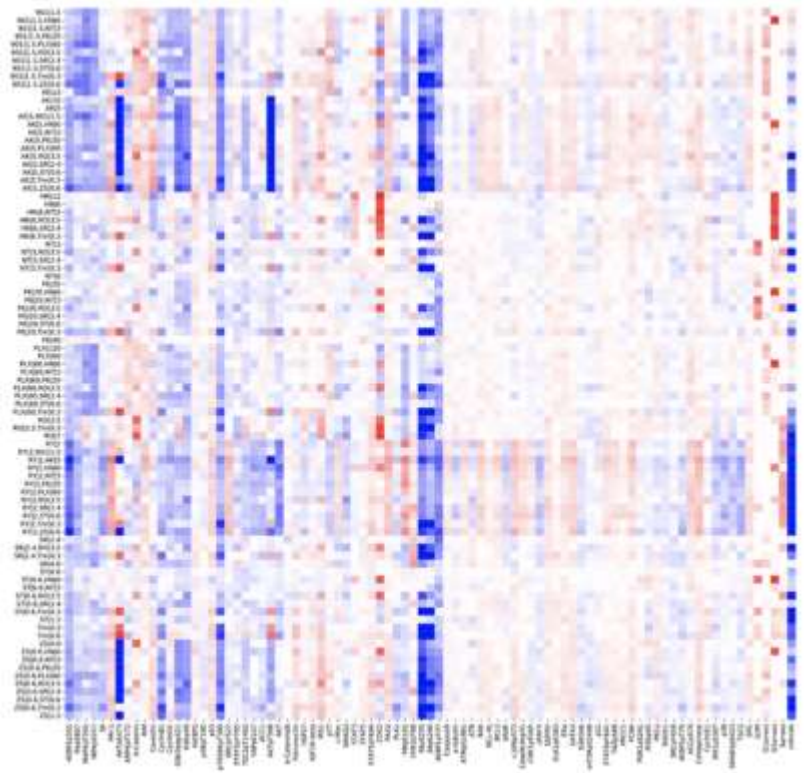
- Comprehensive understanding of tumor behavior at different levels.
- Insightful data for model development and treatment strategies.

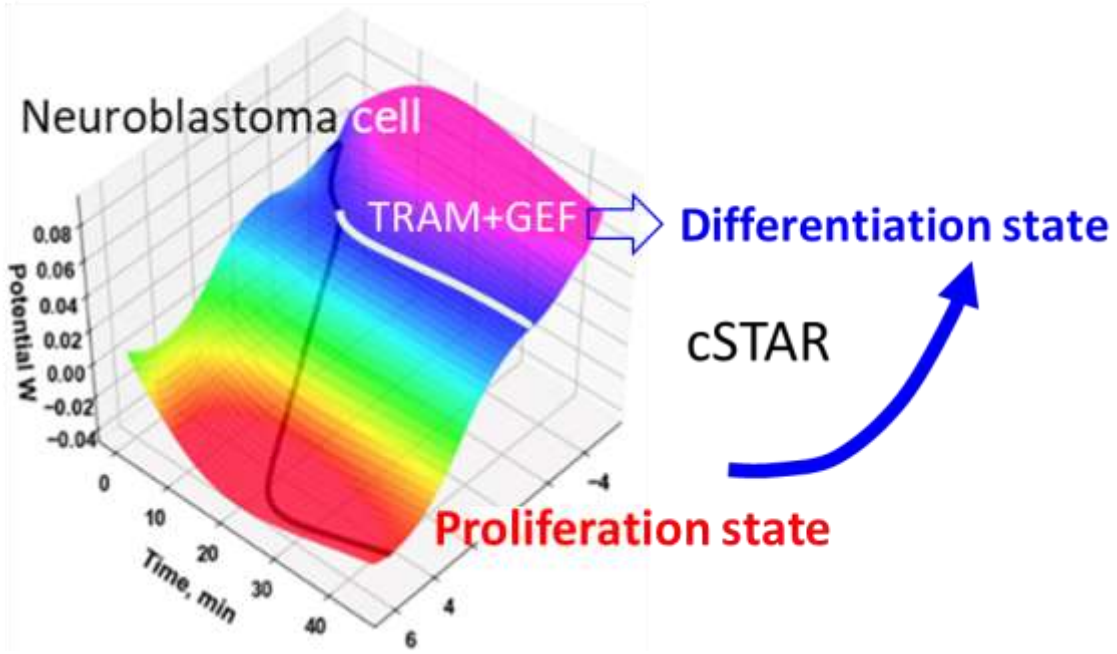


Funded by the European Union

cSTAR - cell State Transition Assessment and Regulation - Enables Us to Reverse Pathological States

cSTAR allows us to push cancer cells back to normal (differentiation)





Controlling the cell journey through Waddington's landscape

Article

Control of cell state transitions

Nature, *609, 975–985 (2022)*

https://doi.org/10.1038/s41586-022-05194-y

Received: 11 January 2021

Oleksii S. Rukhlenko[1], Melinda Halasz[1,2,5], Nora Rauch[1,5], Vadim Zhernovkov[1], Thomas Prince[1], Kieran Wynne[1], Stephanie Maher[1], Eugene Kashdan[1], Kenneth MacLeod[3], Neil O. Carragher[3], Walter Kolch[1,2] & Boris N. Kholodenko[1,2,4]

Trends in Cell Biology

Reversing pathological cell states: the road less travelled can extend the therapeutic horizon

Boris N. Kholodenko, [1,2,3,*] Walter Kolch, [1,2] and Oleksii S. Rukhlenko [1]

CellPress

T4.6: *AI validation based on pre-defined AI parameters*

Objective: To validate AI models against established parameters for accuracy and reliability.

Methodology:

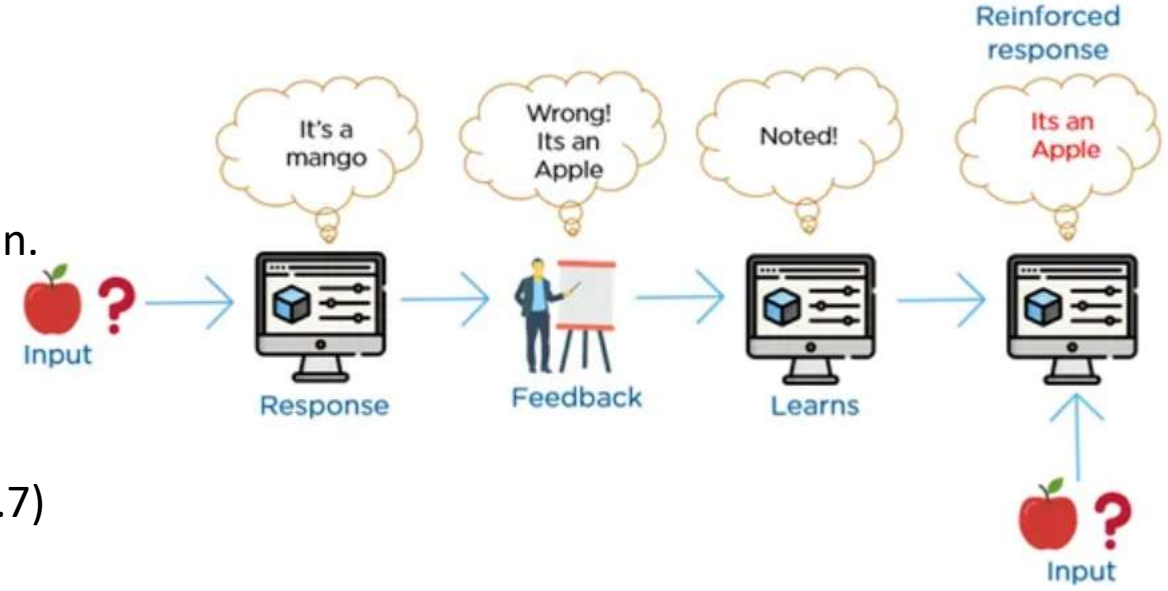
- Testing AI models using predefined parameters.
- Ensuring models meet set standards for clinical application.

Integration with Other Tasks:

- Collaboration with WP3 for data integration.
- Utilizing outputs from WP4 for modeling (T4.2, T4.4 & T4.7)

Expected Outcome:

- Reliable and clinically applicable AI models.
- Enhanced model credibility and trustworthiness.



Objective: Develop AI models to predict immunotherapy outcomes using transfer learning.

Approach:

- Utilize transfer learning techniques for model development.

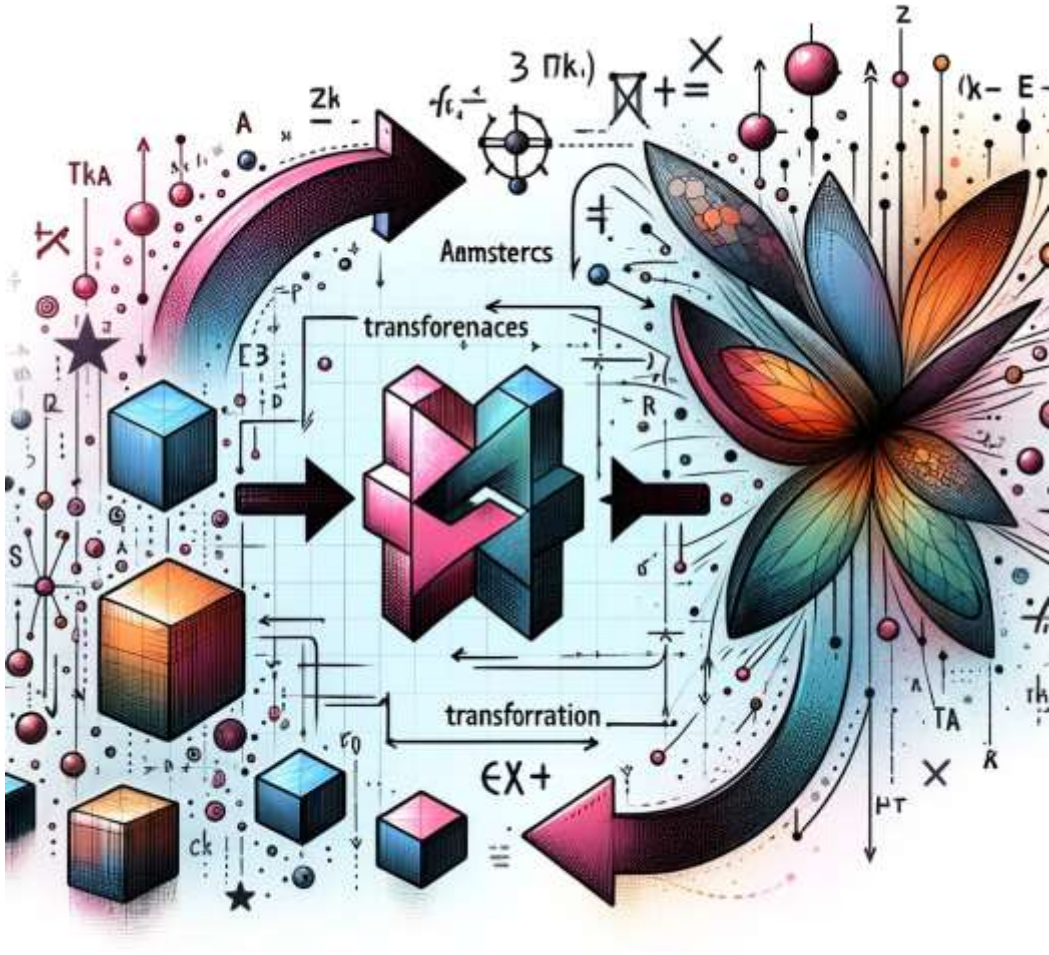
  Immunotherapy Outcomes=BMRA+GCN+Transfer Learning

  T4.4. -> fine tune -> outcome prediction

- Incorporate diverse datasets to enhance model robustness.

Expected Impact:

- Improved prediction accuracy of patient responses to immunotherapy.
- Tailored treatment plans based on model predictions.



Funded by the European Union

WP6 – Stakeholder engagement

Task 6.6: Development of Web based application for predictive and guidance tool (M39-M44) [Lead – ATMC]

- Develop Web based applications to predict outcome and response to therapy based on multiple host (and tumour) features fed to AI models through a Web user interface
- App will be available for every project partner to deploy it on own IT landscape
- Web application to be created based on input from patients (partner PU) and regulators (EMA)
- App to be validated *in silico* using data available to the consortium and collected by consortium members over course of project duration
- Post-project prospective study planned for validation conform to EU regulation 2017/746

Funded by the European Union