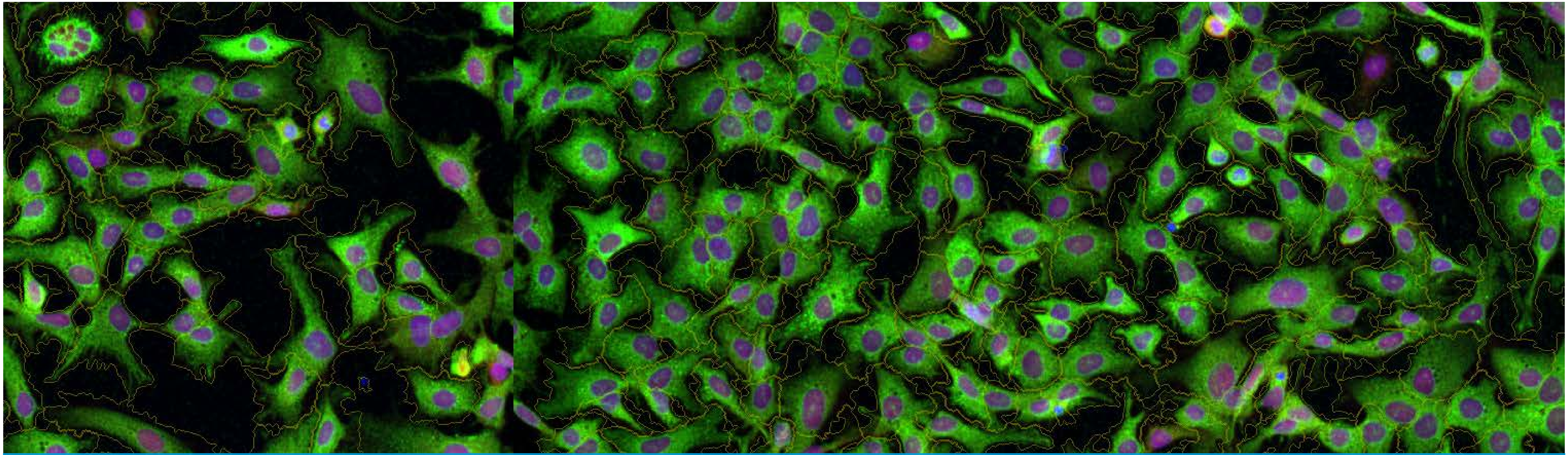


AIDD – kick-off

Janssen Pharmaceutica

26/01/2021

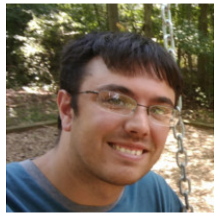
Discovery Sciences

HepG2 hepatocellular carcinoma cells treated with tunicamycin. Nuclei in magenta, and PDI in green.

AI in early drug discovery group in J&J

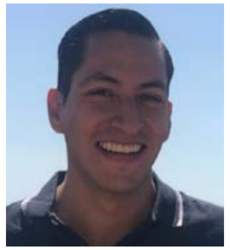
Maciej Kańduła

Natalia Dyubankova

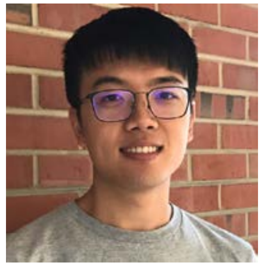


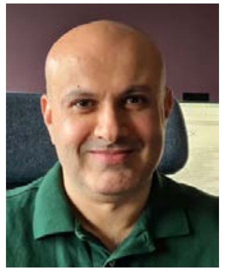
Hugo Ceulemans

Dorota Herman



Oscar Mendez Lucio

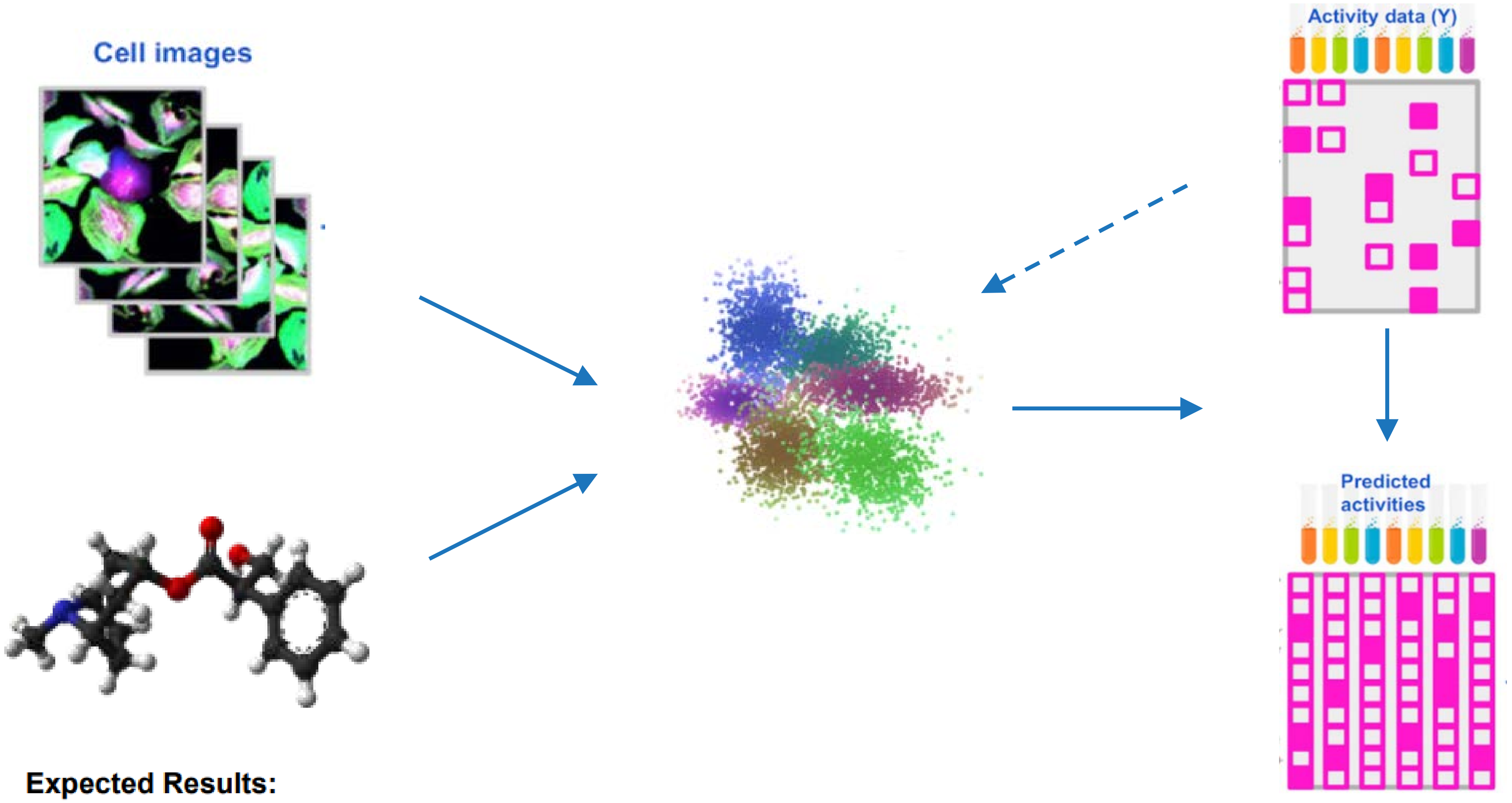




Dries Van Rompaey

Thanh Le Van

Can we learn joint representation for molecules from chemical structures and their image phenotype?

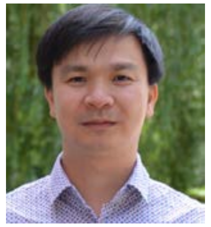


Expected Results:

1. Deep learning models that can learn a single representation for molecules from their low-level chemical structures such as ECFP features, SMILES strings or molecule graphs, and the corresponding microscopy images.
2. Applications demonstrated the advantages of the new representation in a number of domains, including but not limited to a) multi-task learning to predict activity responses of molecules in a number of assays or protein targets; b) tox prediction; c) few-shot learning and domain adaptation to quickly adapt the learned model to a new domain, where only unlabelled data are available and only a small number of images are generated from different cell lines.
3. Pluggable module with the models developed for One-Chemistry model in collaboration with ESR1 and ESR2.
4. Experimental validation of the models within the research project of ESR16.

Dr Gunter Klambauer
ULinz

Maciej Kańduła

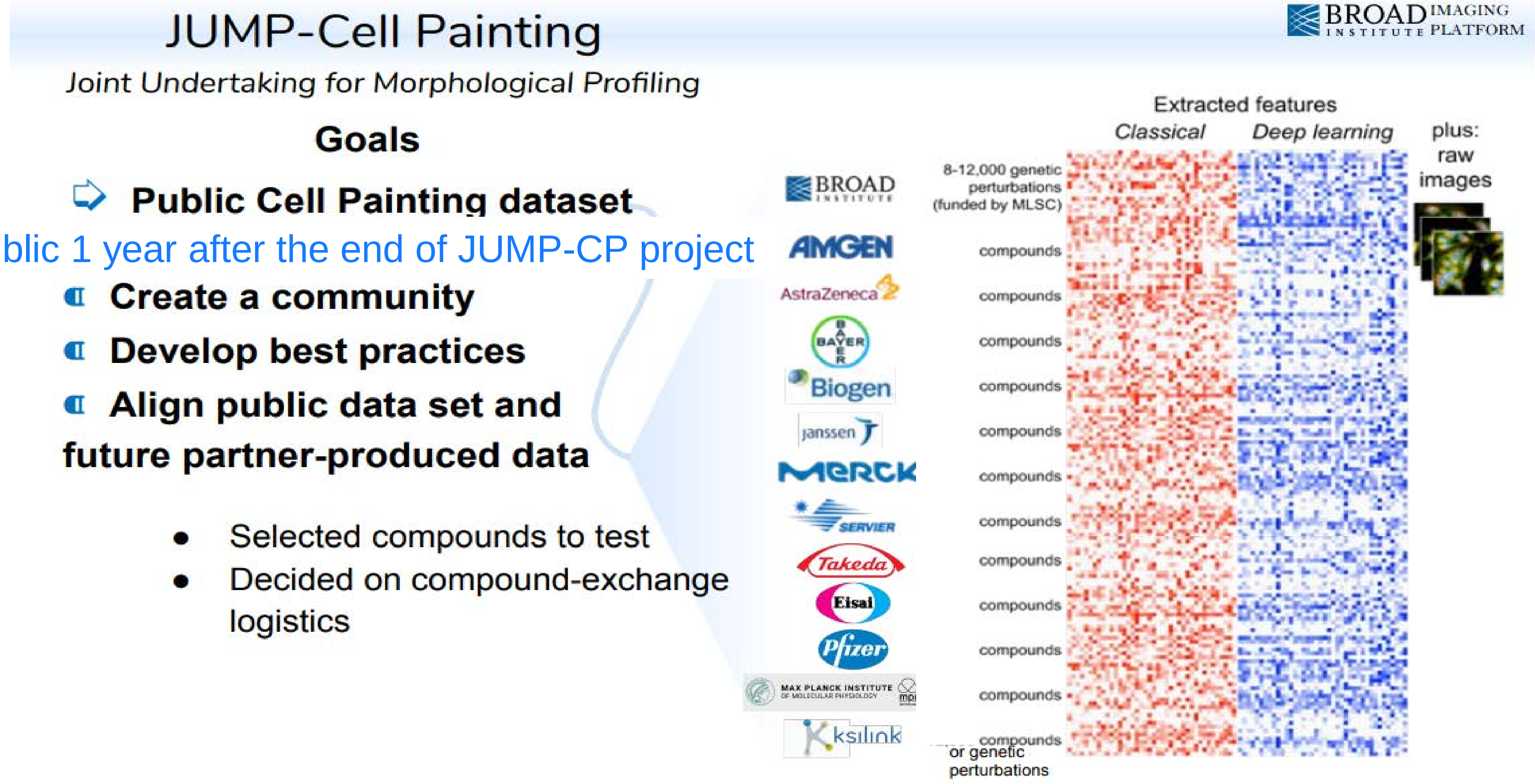


Thanh Le Van

Discovery Sciences (DS)

confidential

Can we integrate microscopy images from different sources to inform compounds design?



Data will be public 1 year after the end of JUMP-CP project

Expected Results:

1. Developed methodology to integrate various microscopy datasets from different experimental settings.
2. Publicly available source code for the pipeline for feature extraction from microscopy images using deep learning.
3. Publicly available predictive models and interpretability framework that associates observed morphological changes with predicted biological activity and structural motifs of the small molecules.
4. Pluggable module for One-Chemistry model in collaboration with ESR1 and ESR2.

Prof. Paul Czodrowski

TUDO

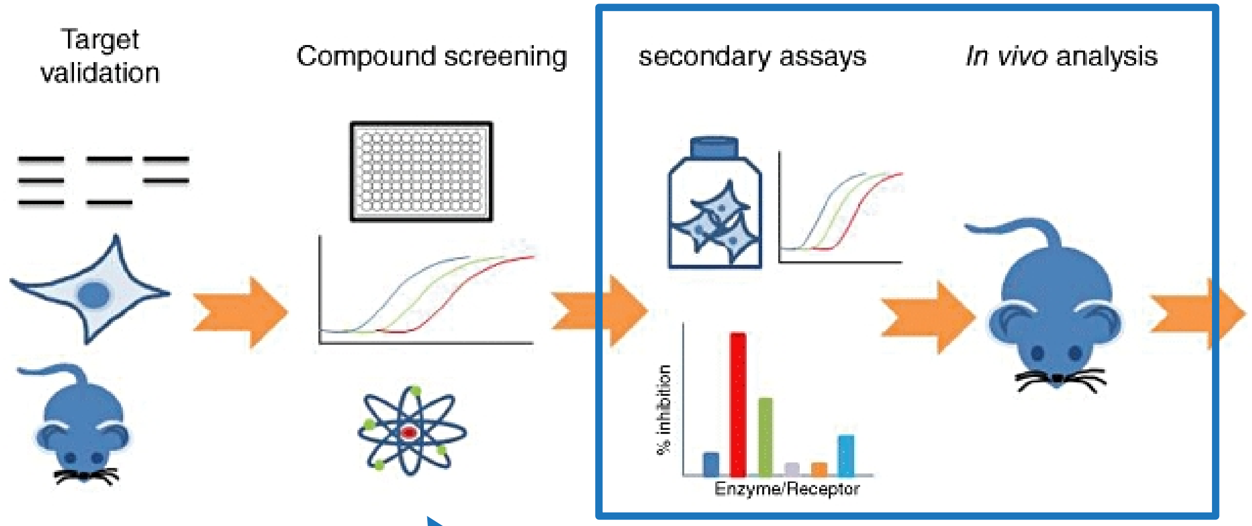
Discovery Sciences (DS)

Oscar Mendez Lucio

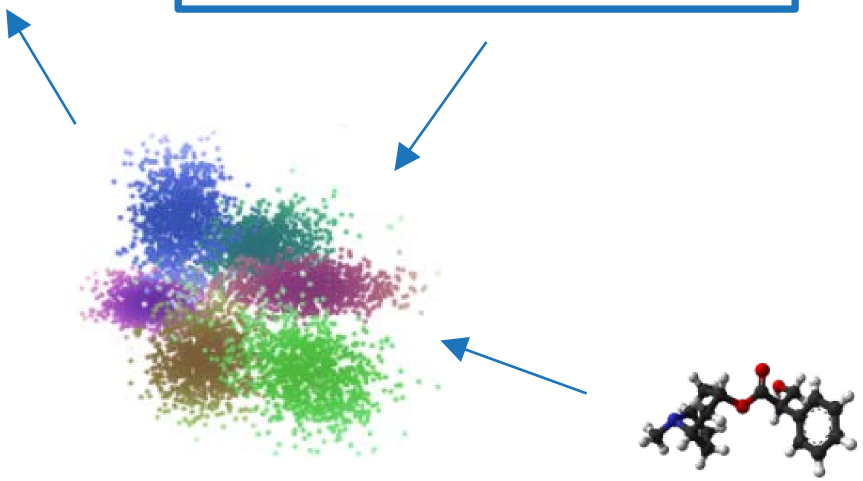
Steffen Jaensch

confidential

Can we bring information from preclinical studies to early drug development?



Br J Pharmacol. 2011 Mar; 162(6): 1239–1249.



Prof. Samuel Kaski

UAalto

Discovery Sciences (DS)

Dorota Herman

Expected Results:

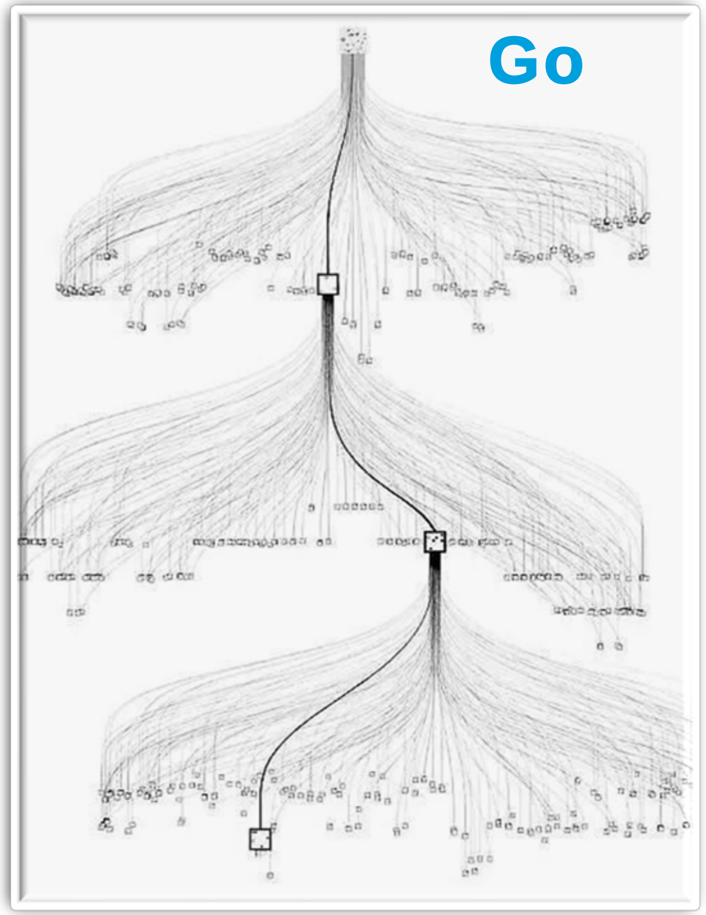
1. Developed methodology for interpretability of the molecular space through decomposable representations coupled with intuitive explanations.
2. Incorporation of the methodology into One-Chemistry model in collaboration with ESR1 and ESR2.
3. Integration of multiple assays with improved in vivo prediction performance.
4. Experimental validation of the technique in ESR16's research project.

Can we predict single-step synthesis?

product

retrosynthesis

building block

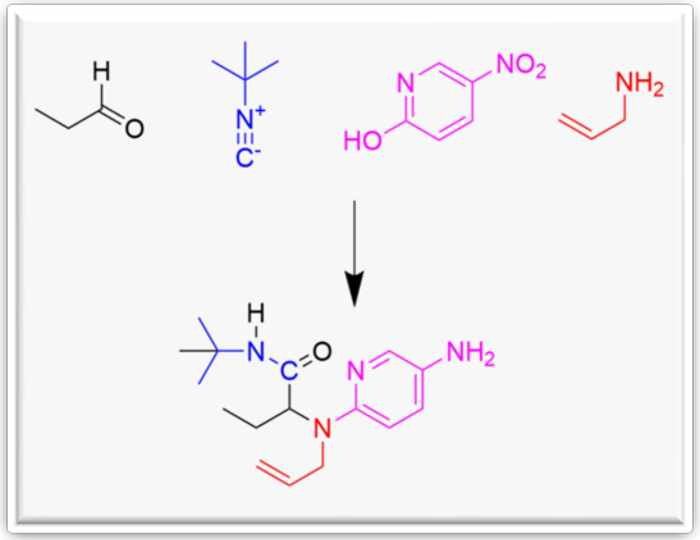


Nature **550**, 354–359 (19 October 2017)

scaffold

forward synthesis

active compound



Expected Results:

1. Publicly available software for predicting the most likely reactions for synthesis of a target compound including reactants, reagents, conditions of a suggested reaction.
2. Interpretation techniques to clarify why the model chooses a particular path.
3. Pluggable module into One-chemistry model in collaboration with ESR1 and ESR2.

Dr. Igor Tetko

HMGU

Natalia Dyubankova

Discovery Sciences (DS)

Can we learn representation for reactivity?

Reference QM data → **Deep Tensor Neural Network** → **Output**
- Energetics
- Transition state geometry
- QM descriptor set

→ **Evaluation**

Prof. Alexandre Tkatchenko

ULUX

Dries Van Rompaey

Discovery Sciences (DS)

Expected Results:

1. Publicly available robust an extensive quantum chemistry dataset of transition state geometries and energies, generated in a standardized and reproducible manner, suitable for machine learning purposes.
2. Publicly available model trained on this dataset that is capable of predicting transition state geometries and energies, enabling such predictions at a fraction of the current cost, both in terms of computational expense and time required.
3. Pluggable model for One-Chemistry model in collaborations with ESR1 and ESR2.

confidential

janssen
PHARMACEUTICAL COMPANIES OF Johnson & Johnson