**Supervisory team:** Ke Li, Krasimira Tsaneva-Atanasova, Anelia Horvath (external collaborator and co-supervisor at the George Washington University)

**Title:** Causal Perturbation Prediction Across Diverse Cellular Contexts via Diffusion Language Models

**Abstract:**

+++

Cancer is not a uniform disease: even within one tumour, cells can differ in identity, state and drug sensitivity. This diversity helps explain why treatment may eliminate some cells while sparing others, leading to resistance and relapse. Yet it is impossible to test every gene target or drug in every cancer context in the laboratory. This PhD project will develop AI methods to predict how cancer cells respond to perturbation, using single-cell gene-expression data to track changes in cell state.

The project will focus on diffusion language models, a new type of AI model that learns biological patterns by reconstructing them step by step. We will adapt these models to predict how gene expression changes after perturbing genes or pathways across diverse cellular contexts. In collaboration with Professor Anelia Horvath at the George Washington University (USA), we will combine public single-cell perturbation resources with cancer datasets from collaborators to train and evaluate models that can generalise to unseen tumour backgrounds, cell states and interventions. We will integrate prior biological knowledge, including gene networks, pathway annotations and protein-level information, to improve accuracy and make predictions easier to interpret.

A central aim is to move beyond pattern recognition towards causal prediction: estimating which perturbations are most likely to push cells into resistant, dormant or therapeutically vulnerable states. The resulting models will be tested on held-out perturbations and contexts and used to prioritise follow-up experiments. By enabling in silico screening before wet-lab studies, this project aims to reduce experimental burden, uncover context-specific therapeutic opportunities, and support precise cancer treatment.

+++