LABORATORY FOR CHEMICAL TECHNOLOGY
Technologiepark 125 9052 Gent, Belgium

Coach	**Supervisor(s)**	**Funding**
Maarten Dobbelaere	Dr. Maarten Dobbelaere Prof. Kevin Van Geem	FWO

Science for AI: Physics-Informed Agents for Predictive Reaction Kinetics

Aim: The aim of this master thesis is to develop a "Science for AI" chemistry agent that integrates a physics-based kinetic estimation method and a solvent selection tool. By embedding these mechanistic constraints into an LLM chemistry tool, the agent will enable experimental researchers to identify realistic reaction conditions *a priori*, minimizing trial-and-error in the laboratory.

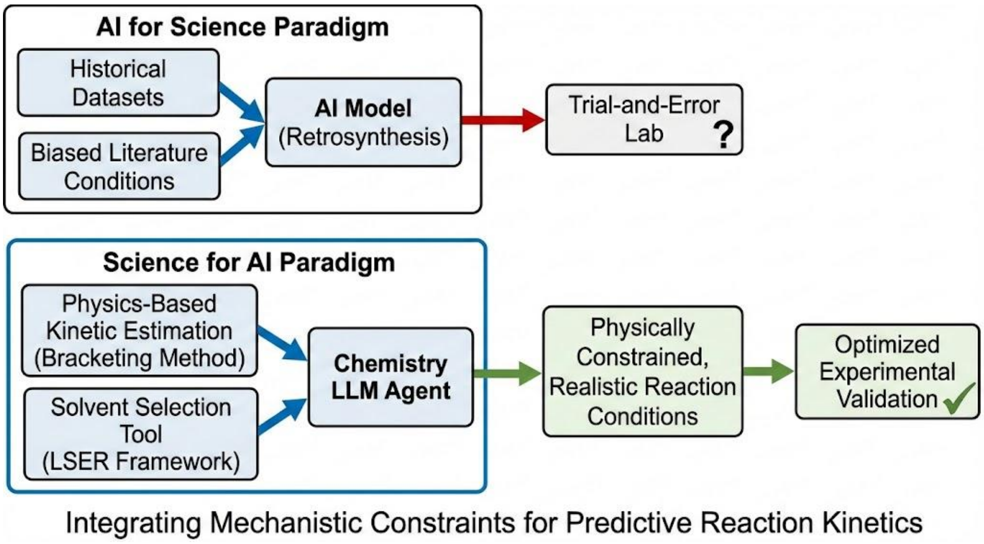

Figure 1: Proposed workflow for introducing physical knowledge into LLMs for a priori reaction conditions.

Justification: Self-driving laboratories (SDLs) are transforming chemical research by integrating Artificial Intelligence (AI) with robotic experimentation. However, a critical disconnect remains in the workflow: while current AI tools excel at structural retrosynthesis (determining *what* to make), they frequently fail to predict optimal process conditions (determining *how* to make it). This failure stems from the "AI for Science" paradigm, which relies on vast historical datasets to predict new reaction conditions. These datasets are inherently biased, with literature conditions often reflecting researcher experience or popularity trends rather than scientific optimization. Consequently, models trained on this data merely mimic historical averages rather than discovering optimal kinetics.

To overcome this, we must adopt a "Science for AI" approach, in which fundamental physical laws are embedded directly into the AI's decision-making process. Previous work at the LCT has established the "Bracketing Method" to define physical bounds for reaction rate coefficients and a linear solvent energy relationship (LSER) framework for calculating solvent effects. This thesis will integrate these mechanistic models into the LCT's existing chemical LLM agent architecture. By constraining the agent's predictions with rigorous physical bounds on activation energy and collision frequency, the resulting tool will generate conditions based on thermodynamic feasibility rather than statistical popularity.

Program

- Literature survey on AI approaches for reaction condition estimation
- Generation and curation of a reaction condition database
- Development of a physics-informed ML algorithm for *a priori* estimation of reaction kinetics
- Integration of physics-informed ML algorithm with LLM agent