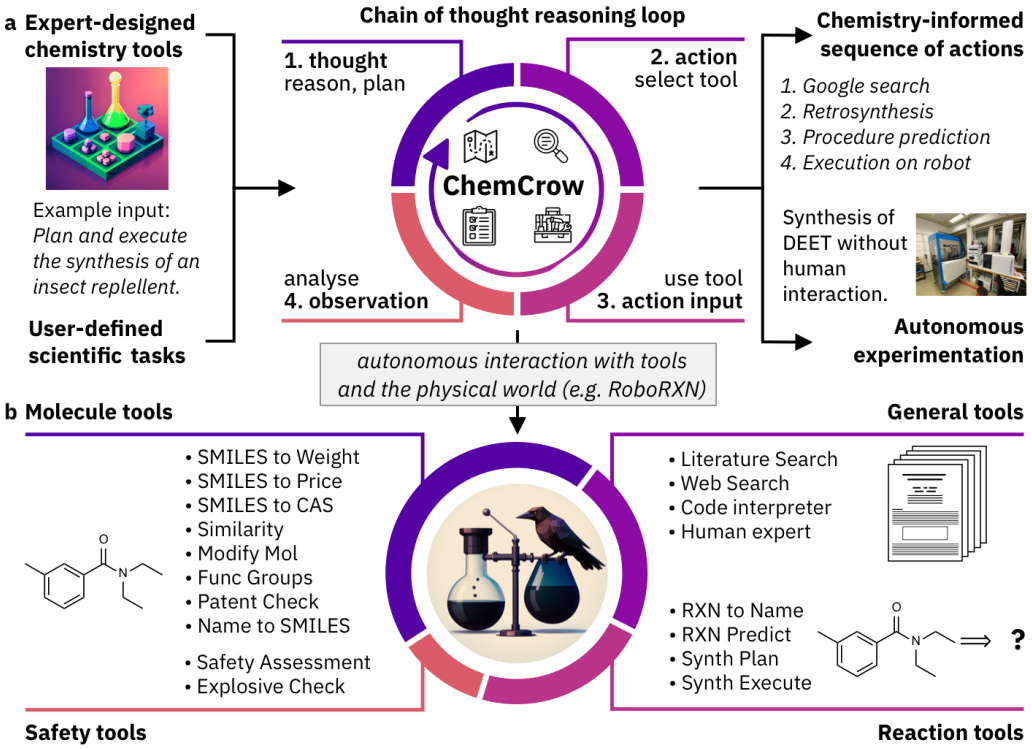
Augmenting large language models with chemistry tools

P. Schwaller[1,4], A. M. Bran[1,4], S. Cox[2], O. Schilter[3,4], C. Baldassari[3], A. D. White[2]*

[1]Laboratory of Artificial Chemical Intelligence (LIAC), Institut des Sciences et Ingénierie Chimiques, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland., [2]Department of Chemical Engineering, University of Rochester, USA. , [3]Accelerated Discovery, IBM Research - Europe, [4]National Centre of Competence in Research (NCCR) Catalysis, Switzerland.

Over the last decades, excellent computational chemistry tools have been developed. Their full potential has not yet been reached as most are challenging to learn and exist in isolation. Recently, large-language models (LLMs) have shown strong performance in tasks across domains, but struggle with chemistry-related problems. Moreover, these models lack access to external knowledge sources, limiting their usefulness in scientific applications. In this study, we introduce ChemCrow, an LLM chemistry agent designed to accomplish tasks across organic synthesis, drug discovery, and materials design. By integrating 18 expert-designed tools, ChemCrow augments the LLM performance in chemistry, and new capabilities emerge. ChemCrow autonomously planned and executed the syntheses of an insect repellent, three organocatalysts, and guided the discovery of a novel chromophore. Our evaluation, including both LLM and expert assessments, demonstrates ChemCrow's effectiveness in automating a diverse set of chemical tasks. Surprisingly, we find that GPT-4 as an evaluator cannot distinguish between clearly wrong GPT-4 completions and Chemcrow's performance. There is a significant risk of misuse of tools like ChemCrow and we discuss their potential harms. Employed responsibly, ChemCrow not only aids expert chemists and lowers barriers for non-experts, but also fosters scientific advancement by bridging the gap between experimental and computational chemistry.



[1] AM Bran, S Cox, O Schilter, C Baldassari, AD White, P Schwaller - arXiv preprint arXiv:2304.05376, 2023