

# MAGIC

Machine-Assisted Guidance for Integrated Drug Discovery

<b>Programm / Ausschreibung</b>	IWI 24/26, IWI 24/26, Basisprogramm Ausschreibung 2024	<b>Status</b>	laufend
<b>Projektstart</b>	01.09.2024	<b>Projektende</b>	31.12.2025
<b>Zeitraum</b>	2024 - 2025	<b>Projektaufzeit</b>	16 Monate
<b>Keywords</b>			

## Projektbeschreibung

Large Language Models (LLMs) und deren einfache programmatische Ansprechbarkeit via veröffentlichten APIs (OpenAI ChatGPT, u.ä.) eröffnen eine Vielzahl von Möglichkeiten in verschiedenen Bereich. Allerdings sind diese Modelle auf allgemein verfügbaren Informationen trainiert, was sie für den konkreten Einsatz in unseren Datenanalyse und - visualisierungssysteme für die pharmakologische Forschung nicht direkt verwendbar macht.

Wir wollen einen Prototypen für ein Natural Language Interface (NLI) entwickeln, der durch tiefe Einbettung der verfügbaren LLMs in die Applikation und den darunterliegenden biomedizinischen Daten in der Lage ist, auch komplexe biomedizinische Fragestellungen zu beantworten - sofern die Antwort in den Daten enthalten ist. Wir werden evaluieren, in wie fern in der Literatur erwähnte Methoden wie Few-Shot Prompting, Chain of Thought, Tree of Thought, Retrieval-Augmented Generation verwendet und erweitert werden können, damit unsere User:innen inhaltliche Fragen an die 450 GB große biomedizinische Datenbank stellen und die Antworten in unseren Applikationen erhalten können. Wir wollen des weiteren daran forschen, wie wir unsere Forschungsergebnisse aus bisherigen FFG Projekten zum Thema Provenance Tracking dazu einsetzen können, die LLMs noch besser mit Kontext auszustatten.

Unser Ziel ist die Entwicklung und Evaluierung eines NLI-Prototypen, der in beide unsere Produkte integriert werden kann und es den Domänenexpert:innen ermöglicht, Befehle aber auch komplexe Fragestellungen an das System zu richten und dadurch die Applikation zu steuern.

## Projektpartner

- datavisyn GmbH