

Advancements in Automated OMOP Concept Code Selection: Leveraging **GPT-4** for Efficient and Accurate **Mapping** of Drug description and **ATC to RxNorm (Extension)**.



Problem statement

Drugs at the General Hospital AZ Maria Middelaes (Belgium) are assigned a **non-structured** textual **description** and an **ATC** code. The OMOP CDM favours **RxNorm** (Extension) codes for drugs.

Solution

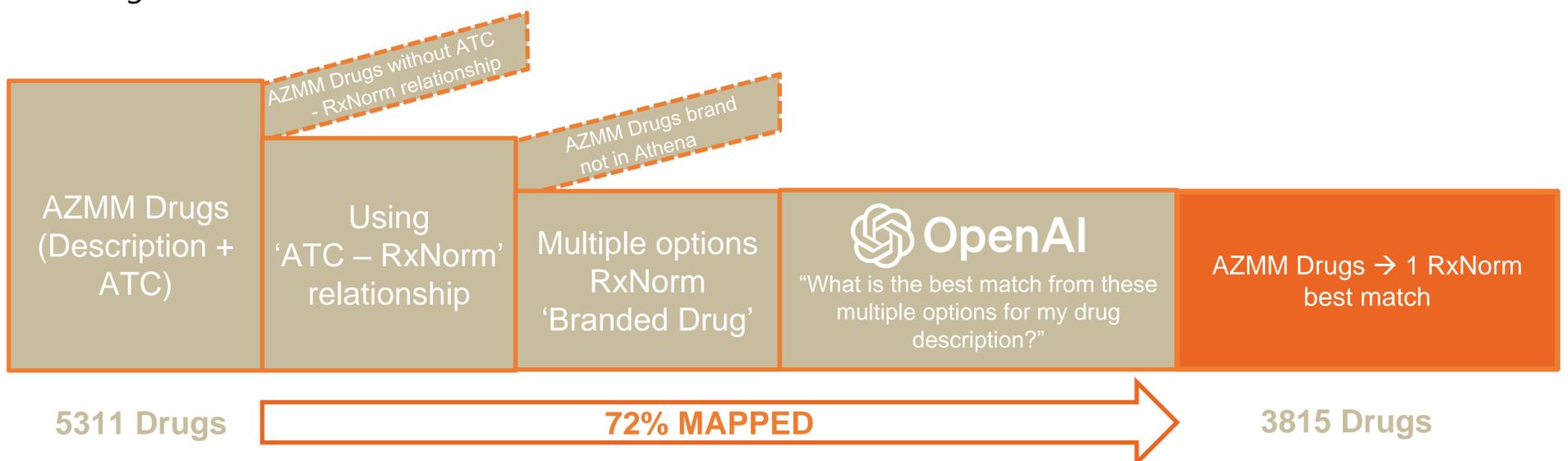


The workflow **transforms** the drug description and ATC code to an RxNorm (Extension) code. To streamline the process and **minimise manual search**, the usage of the Large Language Model **GPT-4** is suggested.



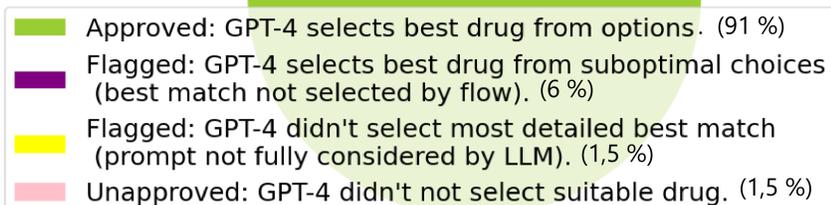
Automated workflow

ATC codes are automatically mapped to RxNorm (Extension) codes through Athena relationships. As this leads to multiple RxNorm options, the LLM GPT-4 selects the best matching option based on drug brand, drug form and drug dose.



Results

20% RxNorm codes selected by GPT-4 are validated



Limitations



- High dependency on completeness and level of detail of drug description.
- Validation performed on 20% of the data.
- Non-deterministic behaviour of LLM and parts of prompt might be ignored.

Conclusion



Drug descriptions were **correctly mapped in 91%** of the validated cases. **Using GPT-4** enables conversion from ATC to RxNorm. Prompt engineering is of high importance to leverage the use of the LLM. Validating the outcomes is still necessary.