ECOCHATTER: YOUR ECO ASSISTANT

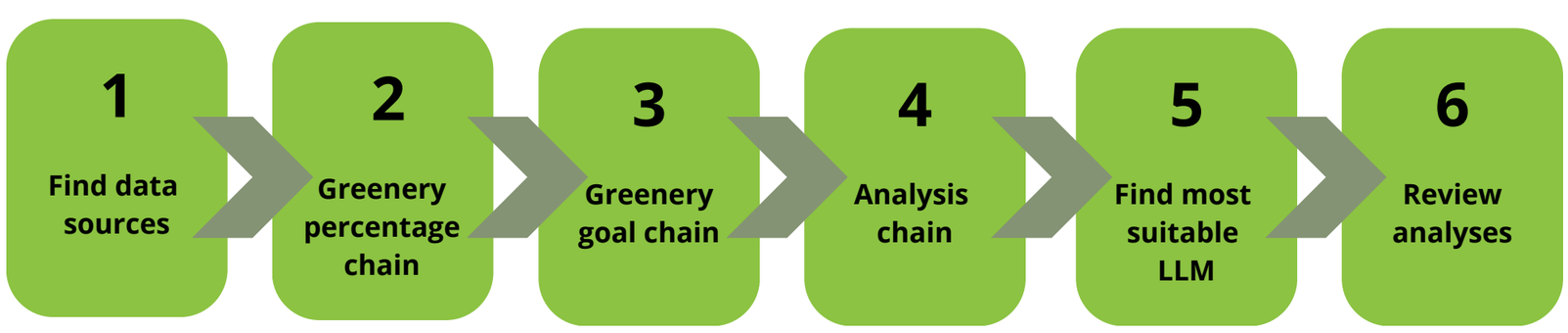
USING LARGE LANGUAGE MODELS (LLMS) TO MAKE NATURE INCLUSIVE DEVELOPMENT MORE ACCESSIBLE

Timo Sevarts, t.sevarts@student.fontys.nl

Introduction

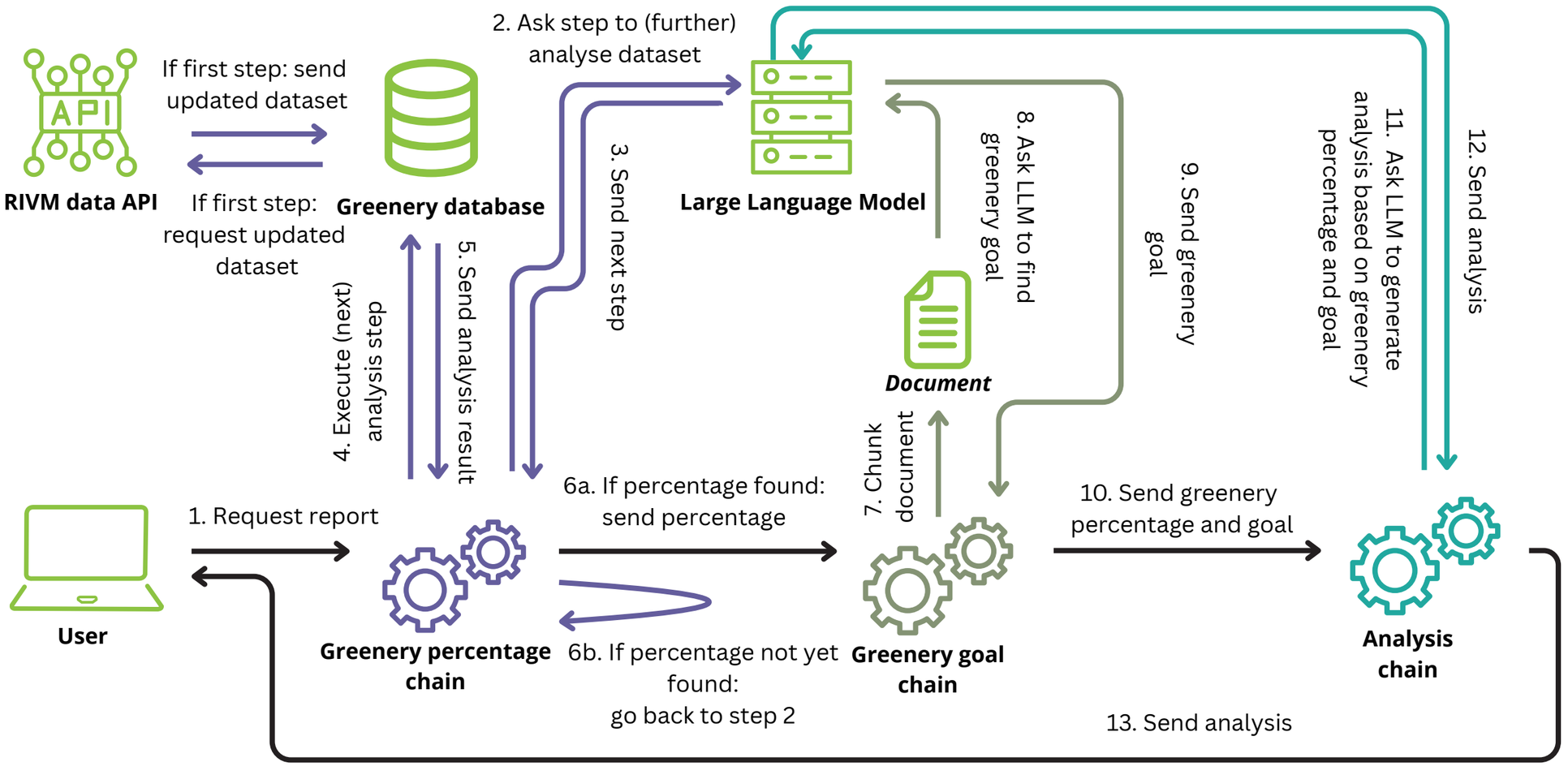
While nature-inclusive urban development (NIUD) has gained a lot of attention in recent years, it has challenges such as uncertainty (because of limited mainstream adoption), unclear regulations and a lack of urban ecology expertise. Background research showed that there is a research gap when it comes to using LLMs for analyzing ecological geographic datasets and what LLM would be suited best for these analyses. Biodiversity reports give insights into such geographic datasets, making it an instrument for informing policy-officers. This could help them understand the context of people requesting permits better and because of this, the officers could give better advice to the people constructing nature-inclusive projects. This is why this study investigated if large language models can be used to generate geographic analyses for biodiversity reports.

Methodology



Results

The figure on the right shows a suitable architecture for generating biodiversity reports using LLMs. The architecture consists of three chains: a chain for retrieving a greenery percentage and goal and for creating an analysis based on this data.



The table below shows five LLMs compared on time and token usage, failed runs, costs and data privacy. The domain expert that reviewed the analyses said that the analyses often assume steps have been taken to improve the greenery percentage, while that is not necessarily true.

	Time (seconds)	Tokens	Failed runs	Costs per 100 analyses	Hosting costs?	Data used for training?
L: Gemma 3 (4B)	19,8	6103,1	0	-	Yes, but can also run locally	No
L: Qwen 3 (8B)	133,3	8206,7	2	-	Yes, but can also run locally	No
C: o3-mini	40	5880,8	0	€2,52	No, pay-per-use model	Opt-out when using free or individual plans
C: gpt-4o-mini	8,1	5469,3	0	€0,32	No, pay-per-use model	Opt-out when using free or individual plans
C: Gemini 2.5 Flash	16,3	6825,3	0	€2,08	No, pay-per-use model	When using free tier

Legend: L = running locally, C = running in cloud

Conclusion

The results demonstrate the potential of LLMs for analyzing ecological geographic data. With adjusted datasets and prompts, this architecture serves as framework for such analyses. Among the tested models, gpt-4o-mini and Gemma 3 perform best: Gemma 3 is suitable for on-device or self-hosted use, while gpt-4o-mini is more cost-effective when only a small number of generations are needed.

Future work

The architecture needs to be developed further, and more validation methods should be used to ensure the generated analyses are reliable and usable for the domain.

nest natuurinclusief

Fontys