Can AI Be Your Sustainability Consultant? Evaluating LLMs for LCA Interpretation and Circular Economy Modelling

Jan Klenner, Alberto Gonzeles, Ivan Salas, Marco Bianchini, Isaac Herraiz

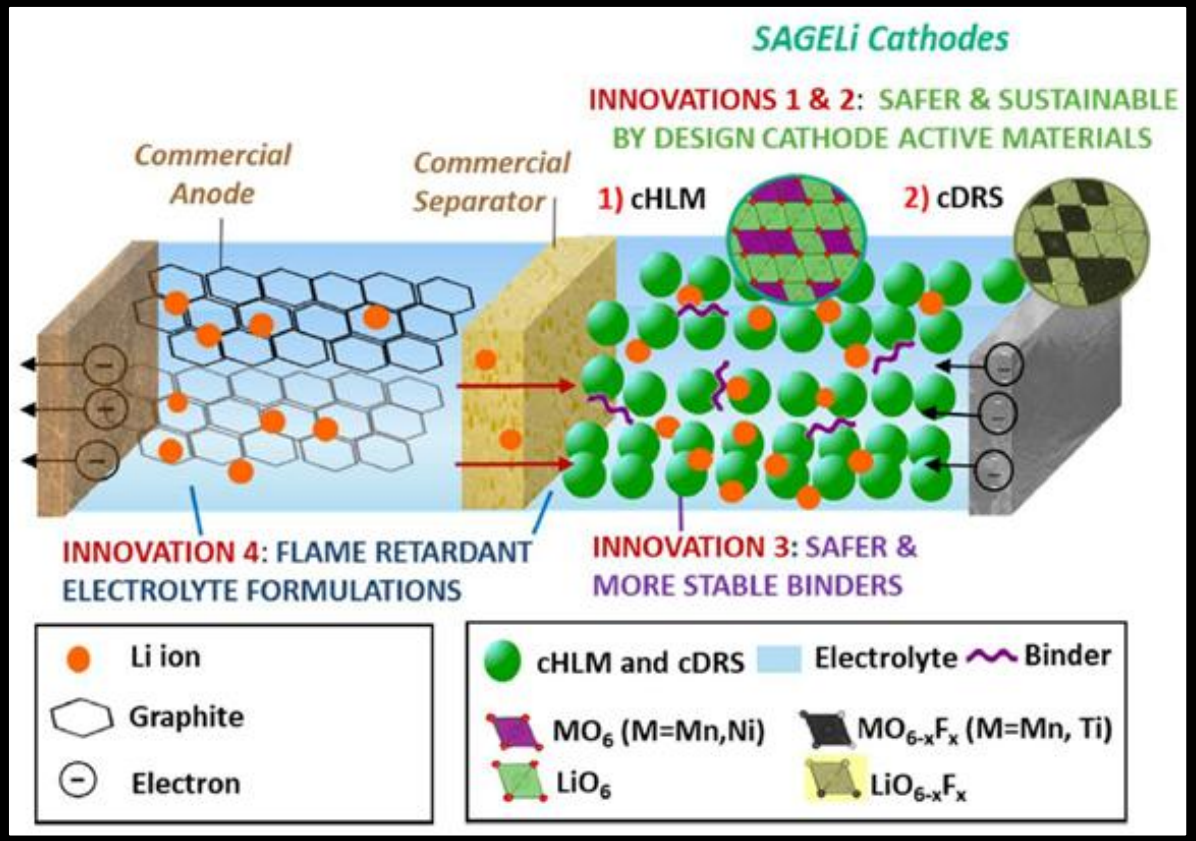
LOMARTOV
[Applied Innovation Engineering]

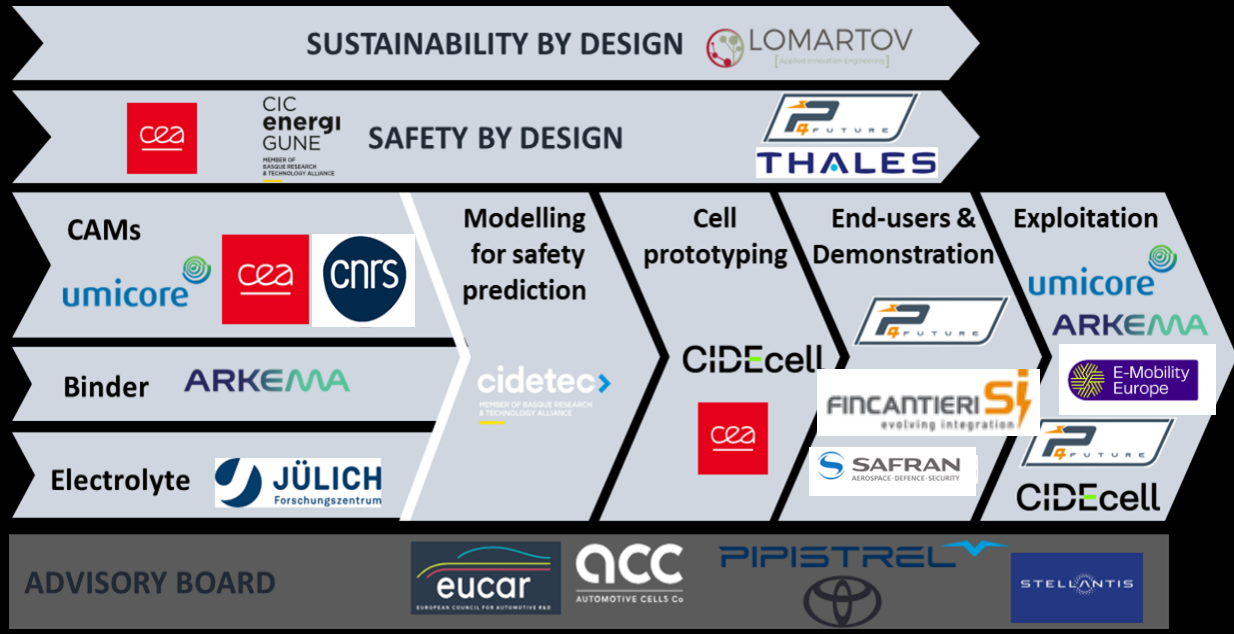
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101147312.

LOMARTOV
[Applied Innovation Engineering]

SAGELi

SAGELi project / Overview

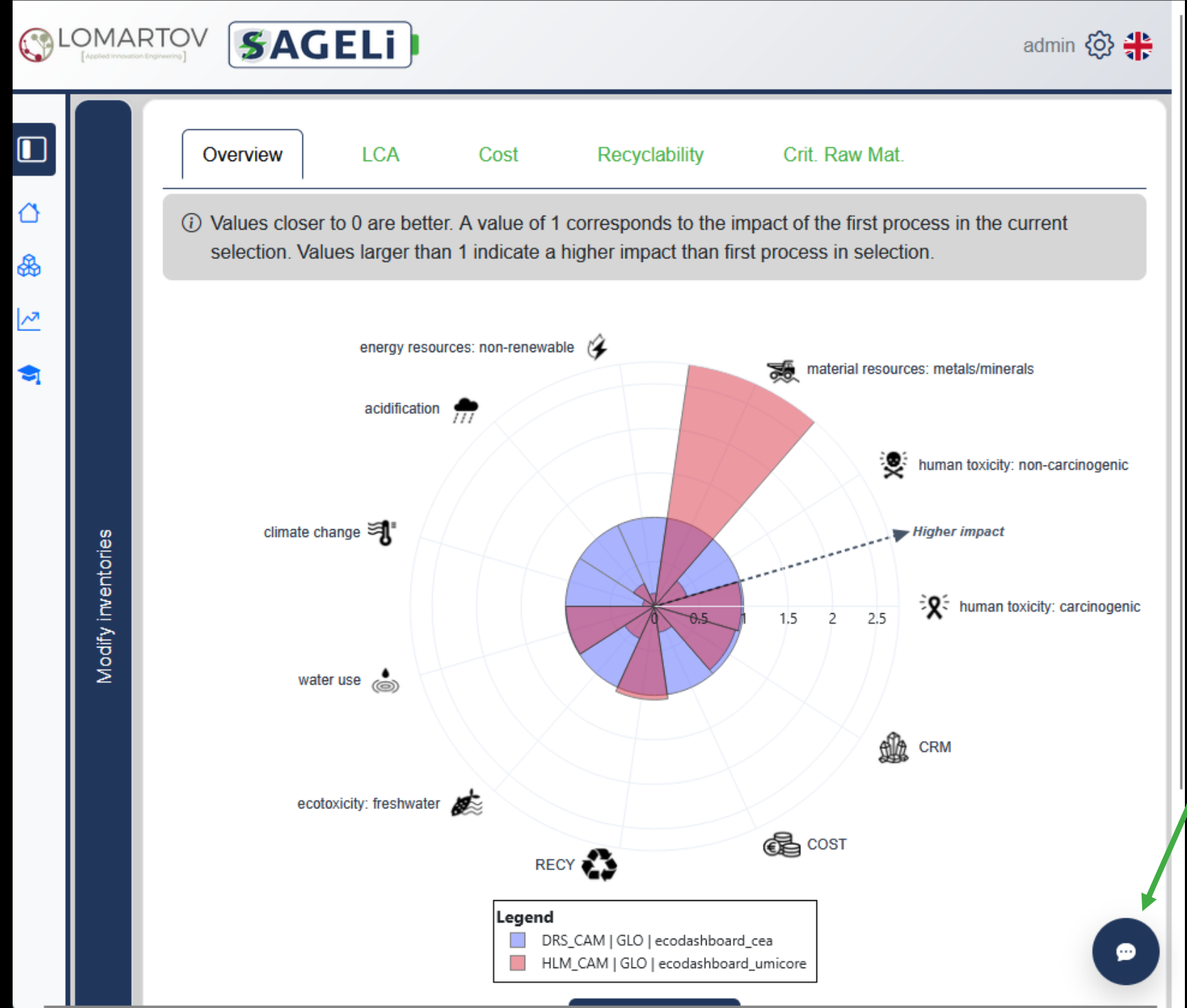




This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101147312.

LOMARTOV [Applied Innovation Engineering]

SAGELi



SAGELi´s Sustainability Dashboard

LLM chat bot

LOMARTOV [Applied Innovation Engineering] | SAGELi

Goal: Compare different LLMs for interpretation of LCA results

- Four LLMs (no extra training) and one colleague
- Standardised prompts
- Variation of input, prompts, etc.
- Scorecard to evaluate results
- Criteria: Correctness, Relevance, Consistency, Clarity

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101147312.

LOMARTOV [Applied Innovation Engineering]

SAGELi

Results: Hotspot identification and mitigation recommendations (case 1) and CFF (case 2)

Case 1: Hotspot	Groq compound	openAI OSS 120B	Claude Sonnet 4.6	Allam 2.7	Human
Correctness					
Relevance					
Consistency					
Clarity					
Overall score					

* Rankings based on qualitative expert judgement

Case 2: CFF	Groq Compound	openAI OSS 120B	Claude Sonnet 4.6	Allam 2.7	Human
Correctness					
Relevance					
Consistency					
Clarity					
Overall score					

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101147312.

Good | Average | Poor

LOMARTOV [Applied Innovation Engineering]

SAGELi

Conclusions

- Relevant differences between LLMs
- LLMs with default settings and simple prompts do not perform great ➔ Effort needed to design prompts
- For now, I would not trust the LLMs for these tasks, human verification needed

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101147312.

LOMARTOV [Applied Innovation Engineering]

SAGELi