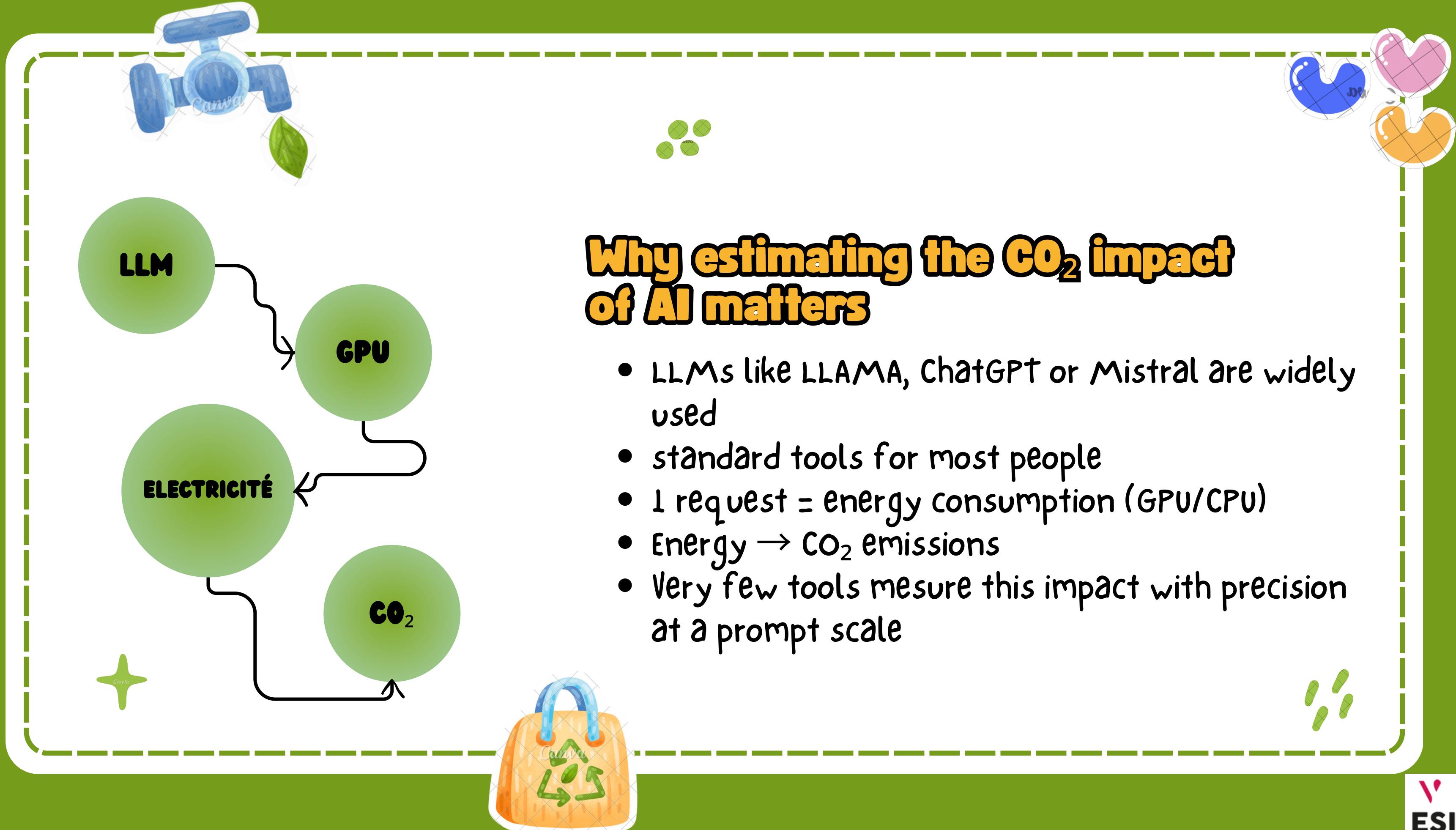
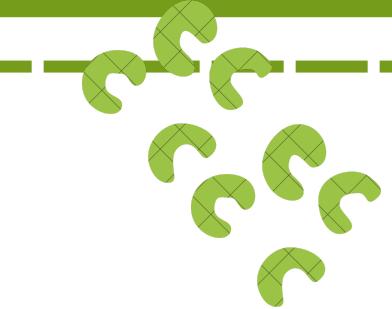


Estimation & Simulation of the CO₂ Impact of LLM Queries

**Towards a more sustainable use Artificial
Intelligence**

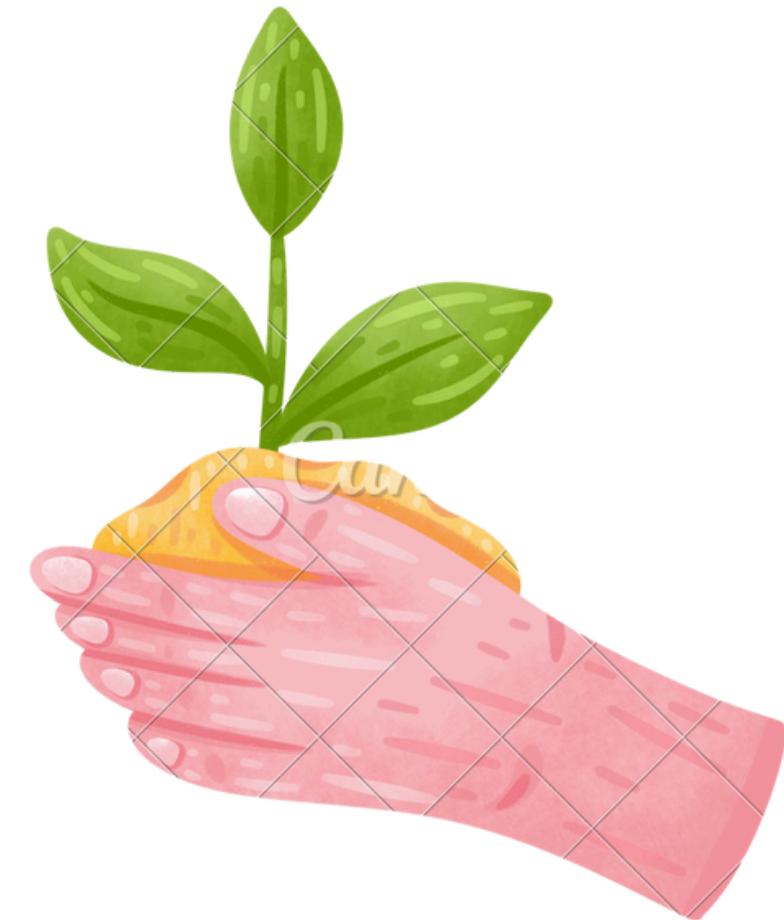
Presented By team 18





Our goal : make AI more transparent and sustainable

- Analyse the energy consumption of open-source models
- Build a predictive model for the energy consumed per request or per token
- Convert this energy into CO₂e (kg) depending on the country's energy mix
- Simulate and visualize its impact through an interactive user interface





```
[+] alpaca_gemma_2b_laptop1.csv  
[+] alpaca_gemma_2b_laptop2.csv  
[+] alpaca_gemma_2b_workstation.csv  
[+] alpaca_gemma_7b_laptop2.csv  
[+] alpaca_gemma_7b_workstation.csv  
[+] alpaca_llama3_70b_server.csv  
[+] alpaca_llama3_8b_laptop2.csv  
[+] codefeedback_codellama_70b_workstation.csv  
[+] codefeedback_codellama_7b_laptop1.csv  
[+] codefeedback_codellama_7b_laptop2.csv  
[+] codefeedback_codellama_7b_workstation.csv  
[+] codefeedback_gemma_2b_laptop2.csv  
[+] codefeedback_gemma_2b_workstation.csv  
[+] codefeedback_gemma_7b_laptop2.csv  
[+] codefeedback_gemma_7b_workstation.csv  
[+] llm_energy_clean_merged.csv
```



Dataset : Real measurements from open LLMs

- Source : HuggingFace – ejhusom/llm-inference-energy-consumption
- Data : 80 000 rows – 79 variables – 15 configurations (model + hardware)
- Models measured : LLAMA, Gemma, CodELLA MA, Falcon
- Key variables : model_name, energy_consumption_llm_total, prompt_token_length, response_token_length, etc...



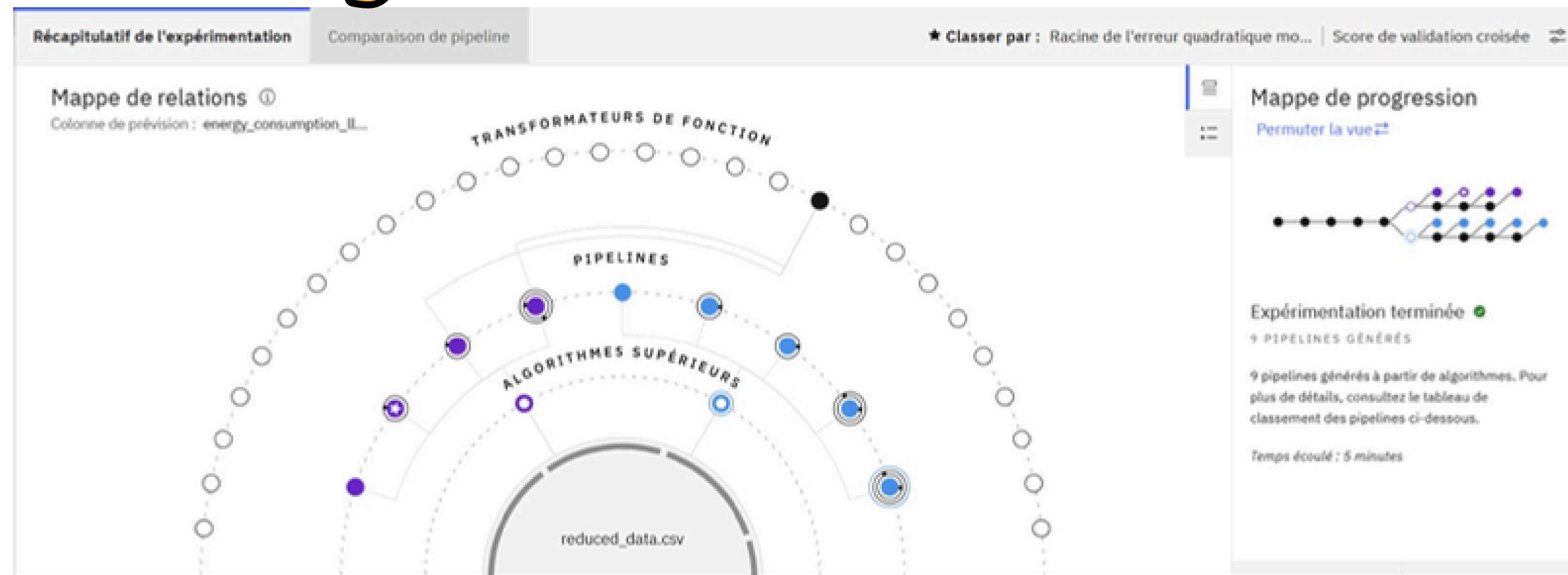
From raw data to clean features

RAW CSV → CLEANING → FEATURE SELECTION → MODELING

- Deleting null values and duplicates
- Standardization of durations (nano-seconds → seconds)
- Harmonization of energy units (kwh)
- Variables selection

model_name	total_duration	prompt_token_length	unique_word_count	response_token_length	avg_word_length	sentence_count	energy_consumption_kwh_total
codellama	4.6284090323.0	60.0	18.0	2043.0	4.909090909090909	1.0	0.0009817730934404
codellama	3.662480170.0	65.0	28.0	203.0	3.9705882362941178	1.0	9.633786433033927e-05
codellama	9.344900499.0	35.0	21.0	493.0	0.1923076923076923	2.0	0.0002148261593305
codellama	4.695640763.0	97.0	47.0	242.0	3.4	3.0	0.0001140618194974

Model training with AutoAI



- Data preprocessing
- Automated model selection
- Automated feature engineering and hyper-parameter optimization

“Everything is automated.”



Model evaluation with cross-validation

MODELS :
DECISION TREE REGRESSORS

Classement de pipeline					
Rang	Nom	Algorithm	RMSE (Optimisé) Validation Croisée	Améliorations	Heure de création
2	Pipeline 1	● Régresseur d'arbre de décisions	1.859e-5	Aucun	00:00:04
★ 1	Pipeline 2	● Régresseur d'arbre de décisions	1.859e-5	HPO-1	00:00:09
4	Pipeline 3	● Régresseur d'arbre de décisions	2.008e-5	HPO-1 FE	00:00:54
3	Pipeline 4	● Régresseur d'arbre de décisions	2.008e-5	HPO-1 FE HPO-2	00:01:04

EVALUATION :
ROOT MEAN SQUARE ERROR
+ CROSS VALIDATION

**HYPER-PARAMETER
OPTIMIZATION**

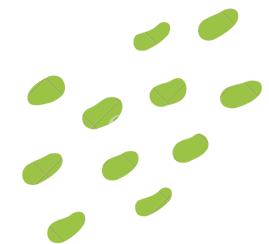


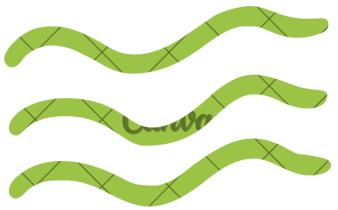


From energy to CO₂e

FORMULA : CO₂e (kg) = Energy (kWh) × $\frac{\text{Intensity (gCO}_2/\text{kWh)}}{1000}$

Mix énergétique	Intensity	Source
France	60 gCO ₂ /kWh	RTE
Europe	300 gCO ₂ /kWh	EEA
Monde	450 gCO ₂ /kWh	IEA





Application : Simulate your LLM's carbon footprint

1. Enter the chosen model and the prompt
2. Predict energy consumed and carbon footprint (CO_2e)
3. Compare models and devices
4. Visualization : web dashboard

Our visualization :

Dashboard template (Streamlit / Gradio / web UI)

+

Cursor “prompt length → CO_2 impact”



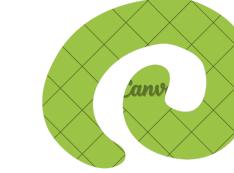
Towards greener AI



- Extend to proprietary models (GPT-4, Claude, etc.) by extrapolation
- Add training energy
- Integrate other parameters: temperature, GPU efficiency
- Raise awareness among users via a Chrome extension



Business impact and added value



A new tool for ESG reporting (the environmental aspect of CSR)

Helping businesses and individuals set clear and quantifiable goals in terms of environmental impact

Give individuals an overview of the environmental impact of each request and workflow

