Can Energy Vampires make organizations more sustainable? 🧛 The case of LLMs and BPM

Roberto Vergallo*, Andrea Andrenucci†, Simone Casciaro‡, Martina De Gaetanis†,
Andrea Elia†, Gabriele Luigi Gravina† Luca Mainetti‡, Sara Matino§,
*Pegaso Digital University, Dept. of Information Sciences and Technologies, Naples, Italy
roberto.vergallo@unipegaso.it
†University of Salento, Dept. of Innovation Engineering, Via Monteroni 165, 73100 Lecce, Italy
{andrea.andrenucci, martina.degaetanis, andrea.elia3, gabrieleluigi.gravina}@studenti.unisalento.it
‡{luca.mainetti, simone.casciaro}@unisalento.it
§University of Salento, Dept. of Laws, Via Monteroni 165, 73100 Lecce, Italy
sara.matino@unisalento.it

***Abstract*—The environmental impact Generative AI – and in particular, Large Language Models (LLMs) – is far from negligible. Despite this, LLMs can help companies rethink sustainable practices, especially in small organizations lacking technical and methodological expertise in ecological transition. Our work has a twofold aim: (i) to investigate whether LLMs can effectively support the reengineering of business processes toward greener alternatives, and (ii) to assess whether employing such energy-intensive technologies aligns with sustainability objectives in a meaningful and cost-effective way. To address these questions, we have designed a three-step methodological Framework currently under implementation. The Framework involves the use of LLMs "as-as-service" for refactoring textual descriptions of business processes. Preliminary results are promising, with GPT-4o and Claude 3.5 demonstrating strong performance on a benchmark of well-known processes provided by Camunda, while the study on the energetic impact of LLMs in organizational settings is still ongoing.**

***Index Terms*—BPM, sustainability, LLM, benchmark, energy**

I. Introduction

The sudden introduction of Generative AI in the current social context has reopened the debate on technological environmental impact, with its rapid large-scale adoption (evidenced by ChatGPT's 400 million users) raising important concerns about the ecological costs of these solutions.

Notably, a study by Morrison et al. [1] reveals that the entire production cycle (including hardware, model development, and final training), generates approximately 493 metric tons of CO_2, consumes 2.769 million liters of water and 913 MWh of energy, with model development alone representing 50% of the total impact compared to training alone.

Despite its high resource consumption, AI remains an extremely valuable tool with the potential to profoundly transform a wide range of sectors. According to Gao et al. [2], for every 1% increase in AI penetration it can indeed lead to up to a 14.2% increase in overall productivity. Regarding the specific case of generative AI, it has been demonstrated by Noy et al. [3] that the use of tools such as ChatGPT has led not only to an increase in productivity in various tasks, but also to an improvement in the quality of the final result and the users' overall perception of the tasks themselves.

Although sustainability is gaining traction in Business Process Management (BPM), current research still presents several gaps. Most studies focus on theoretical models for sustainable process redesign, overlooking the role of advanced technologies in automating this transformation. To fill these gaps, the study presented here investigates the potential of Large Language Models (LLMs) in sustainability-oriented Business Process Reengineering (BPR), with a focus on reducing carbon impacts. The proposed approach is tested on real processes through practical experiments and comparisons between different GenAI models.

The study therefore has a dual purpose: on the one hand, to demonstrate the effectiveness of LLMs in the sustainable transformation of business processes; on the other, to offer an operational model that can be adapted and reused in different organizational contexts.

II. Framework Architecture

The sustainability resulting from the use of Generative AI for process optimization is assessed through a Framework divided into three main blocks, as shown in the overall architecture illustrated in Figure 1. The first block involves the use of LLM technology for the optimization of processes in terms of sustainability. The second focuses on the application of Generative AI for the graphical conversion of processes into BPMN (Business Process Model and Notation). Finally, the third block completes the system by integrating the Digital Green (DG) Framework [4] for the quantitative evaluation of the coherence in the digital and ecological transformations.

The first block focuses on the Text-to-Text transformation of legacy business process description into sustainable textual versions. This stage is accomplished by using LLM technology, which enables the described process to be analyzed and reformulated to meet sustainability standards. The second block transforms the first block output from text to BPMN 2.0 (XML). This phase involves turning sustainable description into graphical representations of business processes. The workflows can be clearly visualized thanks to the graphical

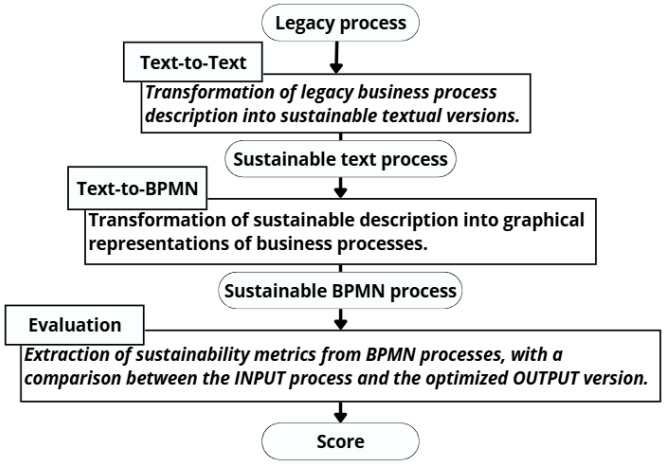

Fig. 1: Architecture of the methodological framework.

representation, which also makes it easier to analyze and evaluate the process, as well as facilitating the implementation of the processes in an operational environment. The "Evaluation" step uses another LLM to extract sustainability metrics.

III. Validation and Results

This section reports the early results for the first block, whose validation includes an experimentation based on a selection of case studies taken from the Camunda repository [5] which contains the textual descriptions BPMN representation of four real, complex and legacy processes. The selected processes include "Dispatch of Goods," which describes the entire shipping process, including documentation, delivery, and issue resolution; "Recourse," which covers the handling of claims or appeals, from request evaluation to final decision; "Credit Scoring," which involves analyzing customer data to calculate their credit score; and "Self-Service Restaurant," which includes customer flow, meal selection, payment, and resource management. The analysis employs five different LLMs, each provided with a use case and a tailored prompt specifying sustainability criteria to transform BPMN process descriptions by enhancing sustainability, digitalization, and green software practices while explicitly mapping BPMN elements in a text-based output format. The aim of the validation is to generate optimized versions of the four processes: each process is addressed by all five models, and each model produces five output generations per process, resulting in a total of 100 distinct outputs used for the quantitative analysis. For the testing phase, LLMs offered "*as-a-service*" are used, i.e. tools based on advanced linguistic models accessible via Cloud infrastructures, chosen for their scalability, flexibility and ease of access. This choice enables the immediate use of pre-trained models, eliminating the time, costs, resources and skills required for training models from scratch. The models considered include GPT-4.0, Claude 3.5 Haiku, Copilot, DeepAI and Perplexity AI.

The final step involves developing and deploying Python code for the G-Eval framework [6], which evaluates and examines the output quality generated by different AI models, with the text-to-text transformation revealing a high level of optimization and systematic integration of the specified sustainability criteria. The code calculates the average score (ranging from 0 to 1) for each output, based on a comparison with a textual description of the reengineered processes, created by a domain expert and evaluating how well it meets the specified constraints relative to the expected output. The results show that the Claude 3.5 Haiku model achieves the highest score (0.77), whereas GPT-4 Turbo performs the worst (0.66), as shown in Figure 2.

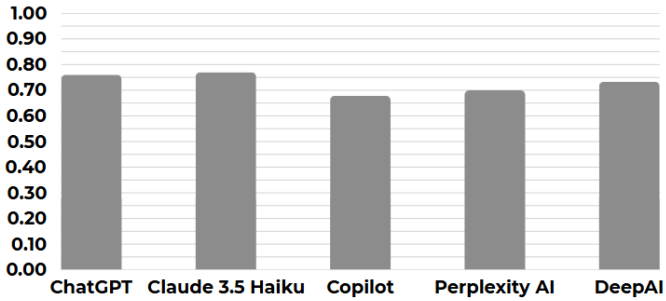

Fig. 2: Performances of the considered AI models. Considered metric: G-Eval.

The complete results of the experiments are accessible for verifiability via the GitHub link: https://github.com/softengunisalento/LLM-Process-Transformations#.

IV. Future Works

A key issue that is still ongoing concerns the balance between the benefits achieved and the energy required to operate these models. Currently, there is no clear estimate of the environmental impact associated with the use of LLMs in the context of BPR, which limits the ability to assert with certainty that their adoption is truly sustainable. For this reason, before developing the remaining two blocks of the proposed Framework, it will be essential to complete the evaluation phase by assessing the actual coherence of the approach. Despite this, LLMs already prove to be promising tools for supporting process designers in rethinking more sustainable practices, especially within small organizations that lack advanced technical expertise in ecological transition.

References

[1] Morrison, J., Na, C., Fernandez, J., Dettmers, T., Strubell, E., & Dodge, J. (2025). Holistically evaluating the environmental impact of creating language models. arXiv preprint arXiv:2503.05804.

[2] Gao, X., & Feng, H. (2023). AI-driven productivity gains: Artificial intelligence and firm productivity. Sustainability, 15(11), 8934.

[3] Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. Science, 381(6654), 187-192.

[4] Vergallo, R., D'Alò, T., Mainetti, L., Paiano, R., & Matino, S. (2024). Evaluating Sustainable Digitalization: A Carbon-Aware Framework for Enhancing Eco-Friendly Business Process Reengineering. Sustainability, 16, 1-18.

[5] "Camunda BPMN repository", url: https://github.com/camunda/bpmn-for-research (accessed June 2025)

[6] "G-Eval", url: https://docs.confident-ai.com/docs/metrics-llm-evals (accessed June 2025)