LLMs-Assisted Decision-making in Sustainable Contexts

Ana Rodrigues
University of Madeira
ITI/LARSyS, University of Madeira,
Funchal, Portugal
Funchal, Portugal
2066213@student.uma.pt

Diogo Cabral
ITI/LARSyS and ISEL, Polytechnic
University of Lisbon, Portugal
Lisbon, Portugal
dncabral@gmail.com

Pedro Campos
ITI/Larsys, University of Madeira,
Funchal, Portugal
Funchal, Portugal
pedrocampospt@gmail.com

Abstract

Promoting sustainability through creative approaches offers a pathway to achieving sustainability goals, yet the resulting solutions often lack efficiency and effectiveness. Addressing this challenge, we introduce SAIESE, a Creativity Support Tool (CST) leveraging a Large Language Model (LLM) to support the creative process, the generation, evaluation, and selection of sustainable ideas. Our work articulates LLMs role in exploration and selection, focusing on sustainable ideas based on three metrics - Environmental Impact, Economic Benefits, and Social Effects. Our evaluation demonstrates that the prototype outperforms human evaluative tasks in usability, creative thinking, and decision-making while reducing cognitive workload. However, its effectiveness in promoting creativity depends on the clarity of ideas, underlining the importance of human collaboration in ambiguous contexts. This work articulates the potential of AI in creativity, expanding current research on CSTs by highlighting sustainability-oriented creative processes and decision-making support.

CCS Concepts

• **Human-centered computing**; • **Human-computer interaction (HCI)**; • **Empirical studies in HCI**; • **HCI design and evaluation methods**; • **User studies**;

Keywords

Creative Process, Creativity Support Tool, LLMs, Sustainability

ACM Reference Format:
Ana Rodrigues, Diogo Cabral, and Pedro Campos. 2025. LLMs-Assisted Decision-making in Sustainable Contexts. In *36th Annual Conference of the European Association of Cognitive Ergonomics (EACE) (ECCE 2025), October 07–10, 2025, Tallinn, Estonia.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3746175.3746188

1 Introduction

Sustainability, in particular, is a systemic challenge that requires a coordinated interaction between diverse and interconnected systems—environmental, economic, and social—to achieve a lasting balance [1]. As the Paris Agreement [2] highlights, sustainable problems require deep knowledge across different contexts. While specific policies have been addressed (e.g., carbon pricing), there is still a need for a versatile tool to empirically validate how different policy instruments can work together, interacting across multiple sectors. In this context, validation is essential to provide tangible evidence and clear guidance to decision-makers, enabling them to identify and select the most sustainable ideas, maximizing the efficiency and effectiveness of proposed solutions. In this framework, creativity emerges, which is seen [3] as key to addressing sustainability challenges since it allows innovative ideas to transform complex problems into opportunities for positive and lasting change. This ability to bring about meaningful change highlights the critical nature of creativity in shaping sustainable futures and addressing urgent global challenges.

This article is based on the notion that creativity is essential to explore solutions that optimize and accelerate sustainability transitions [3, 4]. As described by Sharma et al. [5], these solutions must be evaluated from multiple perspectives. The evaluation requires appropriate performance metrics. Although several studies (e.g., [6–8]) have proposed specific frameworks and metrics to evaluate ideas, these approaches are highly technical and context-oriented, requiring specialized knowledge. However, according to Rodrigues et al. [9], adopting more general metrics (e.g., Environmental Impact, Economic Benefits, and Social Effects) is imperative. Given this scenario, technology emerges as an essential tool for the evaluation and implementation of these creative solutions, highlighting Creativity Support Tools (CSTs) [10], which expand human capabilities, positively influencing users with different levels of knowledge throughout the creative process. Recent research on CSTs has increasingly explored the use of Artificial Intelligence (AI), which has demonstrated the potential to assist people in creative tasks by offering new perspectives or helping to expand conceptual spaces [11]. However, it remains unclear whether idea selection (i.e., the last step of convergent thinking) receives adequate support from the machine or how idea evaluation influences this step.

We present a CST based on LMM – SAIESE (System Approach to Idea Evaluation and Selection of Ecosystem), which supports sustainable ideas' generation, evaluation, and selection (Figure 1). We evaluated SAIESE through three studies with different participants. We conducted an idea generation process with six participants and the SAIESE prototype in the initial phase. The objective of this phase was to generate two ideas to address key sustainability challenges, such as poverty, climate change, and unsustainable human activities, using the same scenario and task for both humans and the prototype, which resulted in 24 ideas. In the second phase, we conducted a crowdsourcing task through an online study in which the 24 ideas generated were evaluated by 72 participants and by the prototype, both based on three metrics – Environmental Impact,

ECCE 2025, Tallinn, Estonia

ACM ISBN 979-8-4007-2033-8/2025/10
https://doi.org/10.1145/3746175.3746188

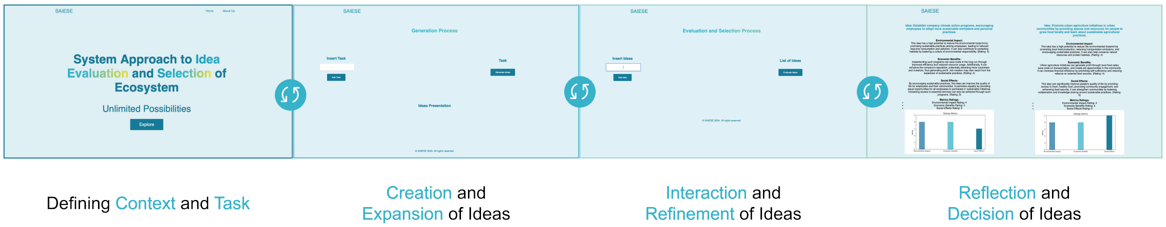

Figure 1: The user interface of SAIESE, a sustainable human-AI idea co-creation environment, was developed using the proposed approach of combining the creative process and sustainability. The approach facilitates human interaction and control over sustainable ideas, with AI facilitating the transformation and synchronization of ideas within and across structures

Economic Benefits, and Social Effects – which allowed us to identify the most efficient and effective ideas as well as a set of ambiguous ones. Finally, we developed two studies with 12 different participants in each study, where we analyzed the usability, creativity support, and decision-making afforded by the SAIESE prototype. The results indicate that the prototype was efficient and effective in supporting the evaluation and selection of ideas. The prototype effectively supports usability and critical thinking, but its ability to foster creativity depends on the clarity of the evaluated ideas. It has also demonstrated its effectiveness in leveraging LLMs to support divergent and convergent thinking, improving decision-making efficiency, and reducing cognitive workload. However, there is a need for greater human collaboration in ambiguous contexts, highlighting the need to balance AI autonomy with human collaboration. The study also revealed that LLM based on a structured and contextualized prompt reduces the homogenization of ideas compared to human ideas. Our contributions are as follows: we articulate how AI can be used to support both divergent thinking (idea generation) and convergent thinking (idea evaluation and selection); we developed an LLM-based CST that utilizes the generative capabilities of AI to support the selection of sustainable ideas through information visualization techniques; and we discussed the challenges in our CST and possible future directions and implications for improving CST. Our findings extend current research on CSTs with the shift toward sustainability by applying informed decision-making through contextualized metrics.

2 Related Work

Our study is positioned in the vision of Oulasvirta et al. [12], which views human-computer interaction (HCI) as a discipline that faces problem-solving by incorporating empirical, conceptual, and constructive aspects. Arguments here show a need to examine the practices of sustainability, creativity, and CSTs. As in HCI, where different types of research and problems are interwoven, sustainable development requires creativity, technology, and social understanding to meet global challenges efficiently and effectively.

2.1 Creativity as a Tool for Sustainability

Creativity is essential for generating innovative and valuable ideas, defined by its ability to produce "new and useful" solutions [13, 14] to general and specific problems in various domains. Creativity integrates two cyclically interrelated cognitive processes – divergent thinking and convergent thinking [14, 15]. Divergent thinking involves generating various ideas, exploring multiple possibilities, and adopting unconventional approaches [14]. In contrast, convergent thinking focuses on narrowing down these ideas, selecting the most viable options, and refining them into coherent solutions [16]. The creative approach has been seen as one of the paths [14] to follow to achieve sustainable solutions. Sustainability is characterized as being a path for world resources to satisfy not only current needs but also those of future generations, and the basis for a sustainable life must be provided today [1], enhancing the aspects of three interconnected dimensions: environmental impact, economic benefit, and social effect. In recent years, many studies have emerged [3] where researchers try to connect creativity with sustainability to achieve sustainable development. In this context, creativity emerges not only as a driver of innovation but as a bridge between the complexity of sustainable problems and the construction of viable solutions. This ability becomes even more relevant when combined with the possibilities offered by digital technologies. The intersection of individuals, technology, and environmental concerns explores how digital technologies enhance and support the consciousness of societies [17]. Based on this premise, achieving sustainability requires a paradigmatic shift in problem-solving approaches. Instead of simply presenting complex information, we propose that we develop a teaching methodology (e.g., [18, 19]) that minimizes the complexity of sustainability information by leveraging technology. In this sense, we highlight the need to develop educational and methodological approaches that enhance creativity as a fundamental skill in solving real problems (i.e., sustainable problems), emphasizing supporting the practical convergence of ideas—a gap that our approach aims to mitigate directly.

2.2 LLMs-based CSTs

With the advent of digital technologies, new tools have emerged to support and expand the creative capabilities of humans, with CSTs standing out [10]. These tools vary in functionality and scope but aim to expand creative thinking and human capabilities [10]. They are defined as "System that runs on one or more digital systems that encompass one or more resources focused on creativity and are employed to positively influence users of different knowledge in one or more distinct phases of the creative process" [10] and have been revolutionized through AI techniques. Human-AI collaboration is increasingly common, with AI enhancing human creativity with its ability to model data and extract knowledge [20]. This evolution has been driven by the increasing availability of generative AI technologies, such as LLMs, which have significant implications for creative workers [21]. These technologies extend human capabilities to produce innovative and creative content, including images, text, music, videos, and code, with results often similar to human results [22, 23]. In addition to performing tasks autonomously, LLMs can guide users through the conceptual spaces of creative thinking, integrating as and into CSTs [24]. For example, Gonzalez et al. [25] developed the Collaborative Canvas to support brainstorming sessions with Llama 2, while Shaer et al. [21] developed the Brainwriting Group-AI framework that uses GPT-3 to support idea generation and GPT-4 to support evaluation. In the multimedia area, Wang et al. [26] developed a tool to support video creation using GPT-4, and Pedtridis et al. [27] developed AngleKindling, a CST that assists journalists through GPT-3. However, these systems have focused on generating ideas, with the evaluation and selection of ideas being little explored. One of the main obstacles in traditional creativity assessments is the intensive human effort required since many methods require manual assessments, often marked by subjectivity and divergence between evaluators [28]. This aspect is especially critical in complex contexts such as sustainability, where it is necessary to consider the novelty and usefulness of ideas [28] and their environmental, economic, and social impacts [9]. Our approach places the human being at the center of interaction with AI, promoting co-creativity focused on sustainable solutions supported by more consistent, systematic evaluation processes aligned with contemporary challenges—a contribution still little addressed in existing CSTs.

3 SAIESE

SAIESE is a tool that supports creative problem-solving in sustainability contexts using LLMs. The prototype was developed around four design goals addressing key human-AI co-creation challenges. These goals guide the system's functionality and are explored in detail in the following subsections.

3.1 Design Space and Rationale

Based on the analysis of related work, we derive four major goals for the design of an AI-based CST: **Goal 1:** Adopt a Cyclical Problem-Solving Process – inspired by the Double Diamond Model [16]. This objective emphasizes a cyclical approach to problem-solving, which covers generating, evaluating, and selecting ideas (Figure 2). **Goal 2:** Integrates Contextualized Sustainability Metrics – application of the three sustainability metrics from the study by Rodrigues et al. [9] to support the generation and evaluation of ideas during the decision-making process. **Goal 3:** Improves information presentation through Generative AI (GenAI), leveraging the capabilities of GenAI, specifically GPT-3.5-turbo. The tool improves the creative process through visual and textual feedback on each idea. LLMs benefit information visualization in several ways, from generating titles and graphs to pattern recognition [29]. **Goal 4:** Promote co-creativity between humans and AI by defining the LLM agent as an independent and co-creative partner. Transformation of the human creative process into a hybrid process, equally promoting innovation and problem-solving.

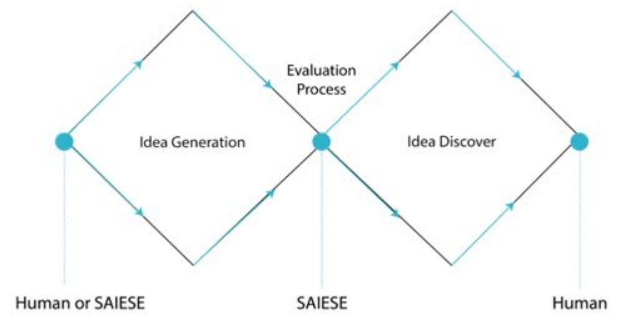

Figure 2: Visual presentation of the Double Diamond model illustrating how the SAIESE prototype supports 'divergence' and 'convergence' during the creative process

3.2 Implementation Details

We implemented SAIESE within the Flask/Python web framework to achieve the objectives (Figure 3). The graphical user interface (GUI) was developed in HTML and JavaScript, allowing tasks or ideas (depending on the stage of the creative process) to be entered through a form that connects to the Python server, which communicates with the model's API GPT-3.5-turbo. The Python server then sends an input prompt to the model with specific instructions for each step in the process, guiding the model's actions. When we defined the context of the model's actions, we formatted its responses to be easily analyzed in the front-end components. So that we can ensure consistency and focus in responses, we set the model's temperature parameter to "0.0", which minimizes randomness and makes responses more concrete and reliable. Although GPT-3.5-turbo is not explicitly trained on sustainability data, it uses an extensive set of web-based data, which allows it to address sustainability-related topics effectively. Setting the temperature aligns with prototype goals without requiring additional specialized training. The server then issues an asynchronous call to the model, which responds over a secure connection encrypted by the OpenAI API secret key. The information is transmitted to the user who requested it after receiving the results (generated ideas, evaluation of ideas, and metric ratings). The model information is then displayed in the user interface (UI), where we develop a loop structure in Jinja2 to iterate over the results. This data is then manipulated by a JavaScript script that dynamically updates the HTML page with the responses received.

3.2.1 Prompt Explanation. The tool is based on a prompt divided into two cognitive processes - idea generation and subsequent evaluation of the same ideas. First, we ask the model to produce 12 creative ideas for a sustainability problem, based on three metrics – Environmental Impact, Economic Benefits, and Social Effects. The prompt includes clear instructions on the number of ideas to be generated and the type of response expected, asking for the presentation of the title and a short description for each idea:

#Generate 12 ideas based on the user-entered task based on the following metrics. Use a maximum of 20 words.

Metrics:

Figure 3: SAIESE Design Interface. Design space and rationale of follow-up action: A & B - Human-AI Co-Creation Process (collaborative interface), and C - Presentation of each idea's textual and graphic assessment

#Environmental Impact:

- The idea's potential to reduce the environmental footprint, conserve natural resources, mitigate pollution, and protect habitats.

#Economic Benefits:

- The idea's potential to generate profit, save costs, increase financial efficiency, and create jobs.

#Social Effects:

#- The potential of the idea to improve people's quality of life, promote equality, strengthen communities, and increase access to essential services.

Provide the title for each idea and a small explanation.

#"""

In the idea evaluation phase, the model receives the 12 ideas individually. We asked the prototype to evaluate each idea based on three metrics – Environmental Impact, Economic Benefits, and Social Effects – using a 5-point Likert Scale, with justification for each score given to each idea and the respective bar graph showing each metric's score for each idea:

#Evaluate each idea based on the following metrics for each idea. Use a maximum of 20 words.

Metrics:

#Environmental Impact:

- Evaluate the idea's potential to reduce the environmental footprint, conserve natural resources, mitigate pollution, and protect habitats.

#Economic Benefits:

- Evaluate the idea's potential to generate profit, save costs, increase financial efficiency, and create jobs.

#Social Effects:

#- Evaluate the potential of the idea to improve people's quality of life, promote equality, strengthen communities, and increase access to essential services.

#Classification:

#Rate each metric using a Likert scale from 1 to 5, where:

#- 1: Not important

#- 2: Moderately

#- 3: Important

#- 4: Very important

#- 5: Extremely important

#Provide the rating for each metric and feedback presentation.

4 Ideas Generation and Evaluation Process

We analyzed the human-to-human generation and evaluation process of the SAIESE prototype. To conduct this analysis, we recognized that in real-world settings (such as private or public departments), workers' expertise varies, including different levels of familiarity with sustainability issues. The participants in this study have different work backgrounds and levels of knowledge about the topic of sustainability, which reflects a realistic workplace. In the following sections, we describe the entire experiment on the creative process.

4.1 Idea Generation

To address idea generation (i.e., divergent thinking), we recruited 6 human participants (4 women and 2 men) with experience in the topic of sustainability across different areas of work – Environmental Sciences (2), Biological Invasions (1), Research Marine (1), Biology (1), Marine Biology (1), through the word-of-mouth method. At the beginning of the study, we explained the purpose and task through a scenario that highlighted how poverty, human activities, and climate change threaten the planet and current and future generations. With the definition of the Sustainable Development Goals (SDGs) [4], areas of critical importance for human well-being and protection of the planet (poverty, threats from human activities, and climate change) were identified. After obtaining informed consent, each participant was invited to propose two innovative ideas to inspire social action and behavior change for 12 human ideas. We then applied the same method and scenario to

Table 1: Ideas generated by humans and the SAIESE prototype

Human Ideas	Prototype Ideas
1. The creation of a virtual before and after (through virtual reality) of same places was affected so that society could explore and better understand the impact of their actions.	13. Create viral social media campaigns encouraging people to share information about the SDGs and how they can contribute to achieving them.
2. Create groups that consider predictions for the future (for example, temperature or pollution) and show changes along a timeline to society.	14. Organize outdoor screenings of films and documentaries related to the SDGs, followed by group discussions on issues covered.
3. Facilitate volunteering opportunities where people can get involved in environmental conservation and community development projects to combat poverty and the effects of climate change.	15. Develop virtual reality experiences that transport people to environments impacted by poverty and climate change, enabling a deeper understanding of the challenges faced by communities worldwide.
4. Promote sustainable ideas competitions in schools and universities, challenging participants to develop innovative projects to tackle the problems of poverty and climate change.	16. Develop mobile applications that provide personalized tips and incentives to help people adopt more sustainable behaviors in their daily lives, such as reducing energy consumption and using public transport.
5. Develop applications or games that teach about poverty, climate change, and the impacts of human activities interactively and engagingly.	17. Organize immersive events that simulate the living conditions faced by individuals in poverty or show the effects of climate change first-hand, creating empathy and inspiring future action.
6. Establish an online platform where communities can share and collaborate on local solutions to problems related to poverty and climate change.	18. Launch reforestation programs in underserved communities to mitigate climate change's effects, provide economic opportunities, and improve quality of life.
7. Virtual reality experiences in the world and, over time, integrate images from different countries received over the decades, showing development, destroyed parts, and living standards.	19. Implement awareness campaigns in public transport systems, informing passengers about the benefits of sustainable transport and encouraging a reduction in the use of private vehicles.
8. Virtual reality or Augmented Reality is where "players" choose items to buy and return in time, visualizing the environmental costs.	20. Establish company climate action programs, encouraging employees to adopt more sustainable workplace and personal practices.
9. Collaborate with local artists to create murals and public art installations that depict the challenges communities face about climate change.	21. Promote urban agriculture initiatives in urban communities by providing spaces and resources for people to grow food locally and learn about sustainable agricultural practices.
10. Establish mentoring programs where more experienced individuals can guide younger people on sustainable practices and how to get involved in social activism.	22. Organize swap fairs where people exchange clothes, books, toys, and other items, promoting reuse and reducing waste.
11. Hold themed hackathons that bring together designers, developers, and policy experts to collaborate on creating innovative solutions to poverty and climate change.	23. Organize interactive exhibitions in cultural centers and museums highlighting the challenges communities face about poverty and climate change, providing opportunities for the public to learn, engage, and be inspired.
12. Create interactive digital media campaigns that engage the public through quizzes, games, and challenges related to the SDGs, encouraging learning and action.	24. Provide training programs for young community leaders, equipping them with the skills and knowledge necessary to lead sustainability initiatives in their communities.

our prototype for comparison purposes, totaling 12 machine ideas. We present the 24 obtained from this study in Table 1.

4.2 Idea Evaluation

We evaluated the 24 ideas obtained in the previous stage through crowdsourcing and controlled evaluative tasks (prototype). The decision to use crowdworkers instead of experts was motivated by the need to collect large-scale evaluations with diverse perspectives and to reflect the general perception of the perceived value of ideas by ordinary users – a relevant audience for the acceptance of sustainable solutions. First, we developed a questionnaire on the XM Edit Survey | Qualtrics Experience Management, which initially included informed consent. Then, we presented the 24 ideas with the respective metrics and rating scale to evaluate each idea based on the three metrics (Environmental Impact, Economic Benefits, and Social Effects) - 5-point Likert Scale, where 1 – Not Important and 5 Extremely Important. Subsequently, we published the questionnaire on the Prolific platform, recruiting 72 crowd workers who evaluated the ideas, with an average duration of 8:03 min. At the end of the task, each participant received a reward of $1.50, according to Prolific payment principles.

We finished this stage with the control evaluative task, introduced the ideas into our prototype, and asked to evaluate the 24 ideas like humans. In Table 2, we present the exact evaluations

Table 2: Scores from evaluations made by crowdworkers and the SAIESE prototype. Idea 1 - Idea 12 are human-generated, and Idea 13 - Idea 24 are prototype-generated ideas

Generative Source	Ideas	Human Evaluation (N=72)				Prototype Evaluation			
		EI	EB	SE	Global Score	EI	EB	SE	Global Score
Human	Idea 1	4.00	3.00	4.00	11	5.00	3.00	4.00	12
	Idea 2	4.00	4.00	4.00	12	4.00	3.00	3.00	10
	Idea 3	4.00	4.00	4.00	12	5.00	4.00	5.00	14
	Idea 4	4.00	4.00	4.00	12	4.00	4.00	4.00	12
	Idea 5	3.00	3.00	4.00	10	4.00	4.00	4.00	12
	Idea 6	4.00	3.00	4.00	11	4.00	4.00	4.00	12
	Idea 7	4.00	3.00	4.00	11	5.00	3.00	4.00	12
	Idea 8	4.00	3.00	3.00	10	5.00	3.00	4.00	12
	Idea 9	3.00	3.00	3.00	9	3.00	3.00	4.00	10
	Idea 10	4.00	3.00	4.00	11	4.00	3.00	4.00	11
	Idea 11	4.00	4.00	4.00	12	4.00	4.00	4.00	12
	Idea 12	4.00	3.00	4.00	11	4.00	3.00	4.00	11
Prototype	Idea 13	4.00	4.00	4.00	12	4.00	5.00	4.00	13
	Idea 14	4.00	3.00	4.00	11	3.00	3.00	4.00	10
	Idea 15	4.00	4.00	4.00	12	5.00	3.00	4.00	12
	Idea 16	4.00	3.00	4.00	11	5.00	5.00	5.00	15
	Idea 17	4.00	4.00	4.00	12	4.00	3.00	5.00	12
	Idea 18	4.00	4.00	4.00	12	5.00	5.00	5.00	15
	Idea 19	4.00	3.00	4.00	11	5.00	4.00	4.00	13
	Idea 20	4.00	4.00	4.00	12	5.00	4.00	5.00	14
	Idea 21	4.00	4.00	4.00	12	5.00	4.00	5.00	14
	Idea 22	4.00	4.00	4.00	12	4.00	4.00	4.00	12
	Idea 23	4.00	4.00	4.00	12	4.00	4.00	4.00	12
	Idea 24	4.00	4.00	4.00	12	4.00	4.00	4.00	12
Mdn ($\boldsymbol{N} = 24$)	4.00	4.00	4.00	4.00	12	4.00	4.00	4.00	12

of each idea and the overall score of each idea, where we add the scores of the three metrics. As all metrics are evaluated with the same scale (5-point Likert Scale), the sum reflects a direct aggregation without the need to perform normalization, and each criterion contributes equally to the total score - fairness in the data. We considered a minimum score of 3 if all metrics received a score of 1 and a maximum score of 15 when all metrics had a score of 5. We boldly highlighted the two ideas considered most efficient and effectively generated and evaluated by humans. Two were considered the most efficient and effectively generated and evaluated by the prototype to constitute our ideas.

4.2.1 Data Analysis and Results. The analysis focused on the human and prototype assessments' consistency, quality, and comparison. We applied nonparametric statistical tests since the data did not follow a normal distribution (Shapiro-Wilk normality test). Using Cronbach's alpha, the human assessments showed high internal consistency ($\alpha = 0.81$), with scores concentrated around 4. In contrast, those of the prototype showed slightly lower internal consistency ($\alpha = 0.75$), reflecting greater diversity in the scores, possibly due to the formulation of the prompts used. The ideas presented different levels of originality and applicability. It was observed that human-generated ideas were more practical and focused on environmental and social impact, while prototype-generated ideas tended towards technological and systematic approaches. The Mann-Whitney U test indicated significant differences in the assessments of Environmental Impact ($U = 397,\ p = 0.005$) and Social Effects ($U = 387.5,\ p = 0.005$) but not in Economic Benefits ($U = 304.5,\ p = 0.716$), suggesting different perceptions about the value of ideas depending on their origin. The comparison between the two types of ideas and their respective evaluation (Table 3) indicated statistically significant differences in the overall scores ($U = 178,\ p = 0.014$), with an advantage for the ideas generated by the prototype, both in the human evaluation ($U = 37.5,\ p = 0.029$) and in the self-evaluation ($U = 36.5,\ p = 0.030$). On the other hand, when comparing human and prototype evaluations on human-generated ideas, there was no significant difference ($U = 47,\ p = 0.131$). In the ideas generated by the prototype itself, there was a significant difference between the human evaluations and the prototype ($U = 34.5,\ p = 0.017$), and the prototype again attributed higher values. These results indicate a tendency for the prototype to present higher scores for both its ideas and human ideas, emphasizing the environmental impact of its suggestions.

4.2.2 Concluding Remarks. After analyzing the generative and evaluative process, we selected the four most efficient and effective ideas and the four most ambiguous ideas. We present only four ideas to simplify the evaluation process, focusing on the in-depth

Table 3: : Presentation of the medians of human evaluation and the prototype of each metric and overall median for each condition for every 12 ideas generatedMetrics

	Human Evaluation (N = 72)		Prototype Evaluation	
	Human GeneratedIdeas (N = 12)	Prototype GeneratedIdeas (N = 12)	Human Generated Ideas (N = 12)	Prototype Generated Ideas (N = 12)
Environmental Impact	Mdn = 4.00	**Mdn** = 4.00	**Mdn** = 4.00	**Mdn** = 4.50
Economic Benefits	Mdn = 3.00	**Mdn** = 4.00	**Mdn** = 4.00	**Mdn** = 4.00
Social Effects	Mdn = 4.00	**Mdn** = 4.00	**Mdn** = 4.00	**Mdn** = 4.00
Global Scores	Mdn = 11.00	**Mdn** = 12.00	**Mdn** = 12.00	**Mdn** = 12.50

analysis of the most representative ideas. The selection of the most efficient and effective ideas followed an evaluation criterion based on three sustainability metrics: Environmental Impact, Economic Benefits, and Social Effects. The ideas with the highest scores in each of the metrics (Environmental Impact – 4, Economic Benefits – 4, and Social Effects - 4) that presented feasible practices in terms of costs, available technology, and community support were ideas 2 and 3.

Similarly, we selected the ideas generated and evaluated by the prototype, which obtained the maximum scores (Environmental Impact – 5; Economic Benefits – 5, and Social Effects - 5), ideas 16 and 18. In addition, four ambiguous ideas were selected (ideas 1, 6, 20, and 21), characterized by potentially promising impacts but with contradictions or sustainability issues. If an idea generates negative environmental or social impacts, the economic benefits may not be sustainable in the long term [30, 31]. Therefore, the final ideas highlight the Environmental Impact and Social Effects metrics, as these two dimensions are fundamental to ensuring more balanced solutions in sustainability.

5 SAIESE

To evaluate the usability, effectiveness, and creative support of SAIESE as a creativity support tool, we followed a mixed-methods approach, combining quantitative insights with qualitative results through two within-subject designs. The objective was to evaluate the effectiveness of the prototype in supporting the decision-making of sustainable ideas in different contexts of complexity. In the first study, we focused on the process involving the most efficient and effective ideas, while in the second study, we investigated the process with ambiguous ideas. In the former study, we aimed to study a more straightforward selection, whereas in the latter, we targeted more complex decisions.

5.1 Study 1 – Selection of the most efficient and effective ideas

5.1.1 Methodology. 5.1.1.1 *Participants.* We recruited 12 participants using the word-of-mouth method (denoted as P_1-P_{12}). Of the 12 participants recruited, 5 were women and 7 men, aged between 27 and 41 years ($M = 31.58,\ SD = 4.68$). Participants reported their familiarity with creative processes when solving tasks due to their work areas. They presented good knowledge about sustainability problems and solutions: "What it is, what it covers, and practical examples." Among the twelve participants, 60% had good knowledge of the topic, 25% had some knowledge, and only 15% had little knowledge. Four of the twelve participants were from the Graphic Design area, four were from Design, three were from the Industrial Design area, and one was from the Service Design area.

5.1.1.2 *Task.* Participants were asked to select one idea from the four ideas obtained in the idea in the previous study. They were informed that they had to find an innovative awareness solution that would inspire social action and behavior change in the face of the following scenario: "Almost a billion people still live below the international poverty line, and human activities and climate change threaten our planet and the livelihoods of current and future generations. The Sustainable Development Goals (SDGs) have identified areas of critical importance for human well-being and the protection of the planet. Researchers and investigators are reframing and questioning the effectiveness of leveling citizen behavior that impacts the environmental crisis".

5.1.1.3 *Procedure.* The studies were carried out in person, lasting approximately one hour. Initially, participants received an overview of the study and signed an informed consent form. Below, we present the scenario with the problem and the four ideas. The experiment lasted an average of 30 minutes and was recorded using a GoPro camera. At the end, participants answered post-study questionnaires and participated in an interview about their experience with the prototype.

5.1.1.4 *Measures.* The post-study evaluation included the System Usability Scale (SUS) questionnaire [32], which has a dual purpose. On the one hand, the usability of the SAIESE prototype is assessed as a specific tool to evaluate and select ideas focused on sustainability. On the other hand, it contributes to validating a broader proof of concept, demonstrating how a certain level of usability can increase engagement by supporting creative processes in sustainability-oriented decision-making environments. This dual perspective highlights the importance of usability not only as a technical attribute but also as a facilitator of meaningful user experiences and innovation, and the Creativity Support Index (CSI) questionnaire [33] focuses on four factors (excluding Collaboration, as it does not apply to the individual prototype environment) to measure the level of creative support. Three sustainability metrics were integrated through textual and graphical visualizations to support decision-making, aiming at accessibility and clarity for non-technical users. The evaluation was completed by analyzing the selection of ideas and the results of the semi-structured interviews, focusing on the experience with the evaluation process, impact on workload, clarity of information, and perception of the accuracy and usefulness of the system.

5.2 Analysis and Results

In this section, we respectively report the responses to the SUS and CSI questionnaires on the usability and creativity of supporting SAIESE. We finalized the analysis with the results of the semi-structured interviews.

5.1.2.1 *Usability and Support Creativity.*

The evaluation results indicate that the SAIESE prototype presented high usability (SUS: $M = 80.21$, $Mdn = 81.25$, $SD = 5.27$) and good support for creativity (CSI: $M = 63.33$, $Mdn = 63.33$, $SD = 5.41$). The data suggest that aspects related to usability positively influenced participants' creative experience with the tool. Most participants (P_1, P_3, P_4, P_5, P_7, P_8, P_9, P_{10}, P_{11}, P_{12}) reported ease of use, efficiency, and overall satisfaction, which is reflected in the positive evaluations of the CSI factors, namely, Enjoyment ($M = 10.36$, $Mdn = 9.34$, $SD = 5.07$), Exploration ($M = 10.14$, $Mdn = 11.34$, $SD = 4.38$), and Immersion ($M = 13.22$, $Mdn = 12.67$, $SD = 6.86$). These results suggest that the intuitive navigation and reduced learning effort highlighted by some participants (P_1, P_4, P_{12}) may have fostered an environment more conducive to creative engagement. On the other hand, some participants (P_2, P_3, P_6, P_9, P_{11}) pointed out difficulties related to the complexity of the textual and graphic presentation of the evaluation of ideas. These perceptions coincide with greater variability in the Expressiveness factor ($M = 11.64$, $Mdn = 8.83$, $SD = 7.16$), indicating that the clarity of the interface can influence the perception of the ability to express and develop ideas effectively. The Results Worth Effort (RWE) factor ($M = 17.42$, $Mdn = 19.50$, $SD = 6.20$) was the best evaluated, reinforcing the relationship between usability and creativity. The perception that significant results compensated for the effort expended appears to be associated with the system's ease of use, particularly the reliability and integration of resources. The results suggest high usability contributes to system acceptance and enhances users' creative involvement. The difficulties identified, although limited to a few participants, reinforce the importance of optimizing the clarity and consistency of the interface to maximize support for creativity.

5.1.2.2 Ideas Selection Analysis.

The analysis revealed that ideas 16 (four participants) and 18 (five participants) were the most chosen, highlighting the role of the CST in the decision-making process and demonstrating the prototype's competence in generating and evaluating ideas aligned with users' expectations. Although human-generated ideas were also relevant, they were selected by fewer participants (idea 2 - one participant, idea 3 - two participants), indicating that the prototype significantly supported the creative process by offering solutions preferred by users.

5.1.2.3 Interview Results.

During the interviews, participants explained in detail how the prototype supports the evaluation and selection of ideas by influencing convergent thinking. Table 4 summarizes the interview results and their implications. When reporting the strengths and weaknesses indicated by participants, we use P_x to denote participant numbers x in the user study.

5.3 Study 2 – Selection of Ambiguous Ideas

5.2.1 Methodology.

5.2.1.1 Participants. We recruited 12 professional participants (8 women and 4 men) through word of mouth (denoted as P_{13} - P_{24}), aged between 30 and 53 years ($M = 39.08$, $SD = 8.34$). Participants reported familiarity with creative processes due to their training and work areas. Three of the twelve participants were from the Interior Design area, two were from Product Design, five were from Design, and two were from the Graphic Design area. We asked about knowledge of sustainability problems and solutions, such as – "What it is, what it covers, and practical examples". 40% of participants had good knowledge, 36% had basic knowledge, and 25% had little knowledge.

5.2.1.2 Task. The task was the same as the task described in Study 1, but modifying the ideas used and with a new organization of ideas in the SAIESE tool, displaying the chart's idea scores simultaneously. This change aimed to facilitate comparisons among ideas with different scores, reducing the cognitive load. For this study, according to their scores, we selected the four most ambiguous ideas - Ideas 1, 6, 20, and 21.

5.2.1.3 Procedure & Measures. This study replicated the procedure and measures used in Study 1, maintaining consistency in methodological aspects and only changing the ideas presented to participants.

5.4 *Analysis and Results.*

This section first reports the analyses performed on usability and creativity support. Finally, we report the analysis of the selection and the results of the semi-structured interviews.

5.2.2.1 Usability and Support Creativity. Similar to the previous experiment and with little variation, the results indicate that the SAIESE prototype presented high usability (SUS: M = 80.42, Mdn = 82.50, SD = 11.67), but with moderate support for creativity (CSI: = 50.72, $Mdn = 50.67$, $SD = 3.95$). Some participants (P_{13}, P_{15}, P_{16}, P_{17}, P_{18}, P_{19}, P_{21}, P_{22}, P_{23}, P_{24}) highlighted the usefulness of the system and the interest in its frequent use, which indicates good acceptance. This positive experience is reflected in the dimensions: Results Worth the Effort (RWE) ($M = 12.06$, $Mdn = 12.67$, $SD = 4.87$) and Immersion (= 10.33, $Mdn = 10.67$, $SD = 3.98$), where participants recognized the value of the results obtained during the experiment and demonstrated their satisfaction in their involvement during the tasks. However, other participants (P_{14}, P_{18}, P_{21}, P_{23}) identified some difficulties, such as the complexity of the interface and during the

Table 4: Strengths and weaknesses of SAIESE output for the most efficient and effective ideas. Strengths highlight the efficiency and objectivity of data analysis. Weaknesses highlight accuracy, explainability, and interpretability issues. Possible solutions through human involvement and initiatives were proposed to address the weaknesses

Convergent-Thinking	Strengths of Prototype	Weaknesses of Prototype	Possible Solutions
Evaluation	Visual and textual assessments were seen as clear, useful and complementary ($\mathbf{P}_3$, $\mathbf{P}_4$, $\mathbf{P}_7$)	Textual assessments sometimes did not capture the necessary depth of some ideas ($\mathbf{P}_2$)	Refining text generation with a focus on context and medium/long-term impact
	Visualizations helped identify patterns and trends quickly ($\mathbf{P}_1$, $\mathbf{P}_5$)	Some ideas were perceived as superficial in more complex ideas $\mathbf{P}_5$, $\mathbf{P}_{10}$)	Include extra layers of analysis or targeted questions for further assessment
	More informed analysis due to the mix of quantitative and qualitative data ($\mathbf{P}_4$, $\mathbf{P}_5$, $\mathbf{P}_9$)	Lack of collaborative elements prevented discussion of evaluations between prototype and participants ($\mathbf{P}_6$)	Add co-assessment functionality
Selection	Idea selection has become more objective and consistent ($\mathbf{P}_8$, $\mathbf{P}_9$)	Reliance on final human review to ensure fairness and accuracy in selection ($\mathbf{P}_3$, $\mathbf{P}_7$, $\mathbf{P}_8$)	Include field for manual notes or custom justifications in the selection
	It made it easier to compare ideas, highlighting aspects that could go unnoticed manually ($\mathbf{P}_8$, $\mathbf{P}_9$)	Some specific and subjective nuances of the ideas were not fully understood ($\mathbf{P}_2$)	Incorporate a shared decision stage or synchronous discussion between reviewers
	Helped reduce human bias, promoting more impartial decisions ($\mathbf{P}_{11}$, $\mathbf{P}_{12}$)	The absence of a collaborative dimension prevented the exchange of perceptions between users in the final choice ($\mathbf{P}_{12}$)	

presentation of the content generated by the prototype. We observed that these limitations negatively influenced the factors Expressiveness ($M = 8.70,\ Mdn = 6.67,\ SD = 4.90$) and Enjoyment ($M = 9.53,\ Mdn = 9.33,\ SD = 4.83$). Finally, the Exploration factor ($M = 10.11,\ Mdn = 9.33,\ SD = 1.53$) also showed good consistency, which indicates that the prototype provided reasonable support for creativity.

5.2.2.2 Ideas Selection Analysis.

The results suggest a tendency for ideas that directly impact the community and are easily understood and applicable. The selected ideas were: idea 1 was selected by two participants ($P_{16},\ P_{21}$), idea 6 was selected by three participants ($P_{13},\ P_{17},\ P_{19}$), idea 20 was selected by two participants ($P_{18},\ P_{22}$), and idea 21 was selected by five participants ($P_{14},\ P_{15},\ P_{20},\ P_{23},\ P_{24}$)

This selection indicates that participants prefer solutions with direct, practical, and community impact, such as idea 21 – “Promote urban agriculture initiatives in urban communities by providing Spaces and resources for people to grow food locally and learn about sustainable agricultural practices”. Although the human-generated ideas were well accepted, the solutions proposed by the prototype demonstrate direct actions and clear community impact

5.2.2.3 Interview Results. Following the analysis method applied in Study 1, we detail how the prototype supported evaluating and selecting ambiguous ideas (Table 5). We use P_x to denote participant number x in the user study.

6 Discussion

In this section, we summarize the findings of the study and discuss its implications. We also report the limitations of the study and our plans for future work.

6.1 CST for Sustainability Support

Our user study's quantitative and qualitative results confirmed the possibility of externalizing cognitive operations during the creative process, supporting divergent and convergent thinking through an LLM-based CST. SAIESE significantly contributed to reducing the cognitive load associated with idea evaluation and reducing the workload during the process, allowing participants to focus on the process more efficiently and effectively, indicating to us that the prototype is well structured. These findings align with the research presented by Oh et al. [34], where they argue that well-designed CSTs can support creative flow without replacing it, supporting users in exploration and making informed decisions. Our approach supports the approach to real and complex problems, such as challenges associated with sustainability. It demonstrates the ability to support critical evaluation and the facilitation of idea selection.

6.2 Interaction between Humans and AI

The prototype adopts the GPT-3.5-turbo LLM in the creative process workflow, facilitating divergent and convergent thinking as described by Oh et al. [34]. AI assistance in the prototype covers

Table 5: Strengths and weaknesses of SAIESE production for the most ambiguous ideas. Strengths highlight the efficiency and objectivity of data analysis. Weaknesses highlight issues of accuracy, explainability, and interpretability. Possible solutions through human involvement and initiatives were proposed to address the weaknesses

ConvergentThinking	Strengths of Prototype	Weaknesses of Prototype	Possible Solutions
Evaluation	Metrics provide an objective assessment and minimize human bias (e.g., preference criteria) (**P**$_{14}$, **P**$_{15}$)	Impersonal evaluation process, without detailed context for some ideas (**P**$_{16}$,**P**$_{17}$,**P**$_{18}$)	Add more context and detailed explanations to complex idea reviews
	Significant reduction in time and mental effort when viewing information (**P**$_{17}$, **P**$_{24}$)	Lack of confidence in some aspects of automated assessment and some superficial assessments (**P**$_{22}$)	Integrating more human insights and customization into evaluative analysis
	Information Visualization makes it easier to compare ideas (**P**$_{13}$,**P**$_{19}$,**P**$_{20}$)	Lack of human collaboration in assessments (**P**$_{15}$,**P**$_{23}$,**P**$_{24}$)	Deeply explore the long-term impacts and complexity of ideas
Selection	Effectiveness in highlighting more relevant and accurate ideas (**P**$_{23}$)	Lack of flexibility in the selection process for adjustments after the initial analysis (**P**$_{14}$, **P**$_{17}$)	Offer manual adjustment options for greater customization in the selection process
	Good organization of ideas in a logical and structured way, facilitating decision-making (**P**$_{13}$)	The selection process should be more dynamic and flexible, with more human interaction (**P**$_{22}$)	
	The speed and clarity of the selections helped to prioritize promising ideas (**P**$_{13}$, **P**$_{19}$)		

three aspects: idea generation, idea evaluation through process automation, and data transformation into text and graphics to support decision-making. Analysis of participant feedback and study results revealed distinct interactions between humans and AI, particularly in the idea evaluation process.

In Study 1, we found a higher degree of trust in the prototype, where participants demonstrated reduced cognitive workload and confidence in decision-making in convergent thinking tasks, in contrast to participants in Study 2, who preferred greater human collaboration in evaluating ambiguous ideas. These results highlight the challenges that arise when AI systems are faced with less structured and objective information.

Another observation is that although idea 21 (idea generated and evaluated by the prototype) was the most selected, the attention given to idea 6 (idea generated and evaluated by humans), both ambiguous ideas, suggests that the prototype managed to attract participants' interest even in contexts of greater uncertainty. These findings align with the study presented by Amershi et al. [35], which emphasizes the need for AI systems to balance autonomy with mechanisms that encourage collaboration between humans and AI. According to the investigations [34, 36], the prototype increased the efficiency and effectiveness of participants' decision-making by presenting superior and more complete evaluations than human evaluations. These positive empirical results indicate that the concern with highly centralized and data-driven AI systems (such as GPT), which present a diverse range of results [37], can be minimized by using more contextualized models, such as the ones we developed in our prototype (incorporating task-oriented metrics). Thus, the prototype can recognize and value aspects of ideas that may not be as evident to humans.

Our results represent the center of a unidirectional spectrum of human–LLM/AI collaboration presented by Faggioli et al. [36], where LLMs support humans in varying degrees of interdependence, supporting humans in their judgments. Another observation was that participants did not realize that the prototype generated some of the ideas presented. This demonstrates that generative AI, when guided within the context and application domain, can effectively reproduce the nuances of human-generated text, i.e., Natural Language Processing techniques. Our experimental results suggest that our AI-based CST can simplify human decisions during the creative process, from the easiest to the most complex.

6.3 Mitigating the Generative and Evaluative Stage of the Creative Process

We analyzed the results against concerns raised by Anderson et al. [38] regarding the risk of homogenization when using LLM-based CSTs. Specific thematic patterns emerged among the ideas generated by the prototype, such as raising public awareness of SDGs (ideas 13, 14, 19, 23, 24), immersive experiences (ideas 15, 17, 23), and community engagement (ideas 18, 21, 22). However, these ideas included a variety of tools and platforms, with a combination of technology and physical actions in ideas 16, 20, 21, and the incorporation of cultural/artistic elements in ideas 14 and 23.

These findings indicate that, despite some common themes, the prototype did not exhibit the level of homogenization suggested in [38]. Instead, the diversity of methods and formats indicates the model's ability to generate diverse and contextually rich results when prompted appropriately. In contrast, human-generated ideas tended to be more thematically and technologically repetitive, focusing on VR/AR, as in ideas 1, 2, 7, 8, addressing educational strategies, as in ideas 3, 4, 5, 10, and including social media, as in ideas 6 and 12. Although idea 11 presented a different approach (hackathons), the overall pattern suggests that manual idea generation tends to be related to familiar or dominant concepts.

Regarding the evaluative stage, human evaluations showed high internal consistency but limited variability, which aligns with previous findings of uniformity in human judgment. In contrast, evaluations performed by the prototype showed greater variation, likely due to the contextual metrics incorporated into the prompt.

In summary, our findings highlight the potential of LLMs, focusing on the goal to be achieved to help mitigate homogenization in both the generative and evaluative stages of the creative process.

6.4 Challenge and Future Work in Developing AI-based CSTs

Although SAIESE demonstrates significant potential in supporting sustainability-oriented creative processes, several areas for improvement were identified. Semi-structured interviews revealed mixed perceptions about the complexity of the prototype, such as positive assessments of usability and creative support, until the prototype consideration was somewhat complex. he complexity of the evaluation presented on ambiguous ideas was also a limitation that emerged from our analysis, thus addressing the lack of trust of humans in AI technology. We also observed that some participants had difficulty interpreting the results, using manual methods, such as Likert scales, to independently assess whether the evaluations presented by the prototype were aligned with their judgments. The feedback obtained from the interviews highlighted the interest of participants in hybrid interactions of this kind, interactions during the creative process that the user can guide through feedback loops. These mechanisms are believed to increase usability and support for creativity, user confidence, and greater control over the process, especially when dealing with ambiguous ideas. We also observed that information visualization techniques present an area for improvement since both Study 1 and Study 2 participants reported cognitive overload and difficulties interpreting the information the prototype returned. A possible solution would be incorporating interactive panels for simultaneous visualization, which would correlate with textual and graphical data.

These challenges indicate the diverse preferences and cognitive abilities of users.

7 Conclusion

This paper aims to investigate how AI can support creativity and reveal important aspects that need to be considered in AI-based CSTs in the context of sustainability. Creativity is a subjective concept applied differently depending on people and environments. In this work, we address the creative process aimed at sustainability by developing a CST based on an LLM, SAIESE. The work contributes to advances in how users can interact with LLMs in creative tasks, while also offering opportunities to improve the usability, interpretability, and flexibility of CSTs to support the creative process.

Acknowledgments

This research was funded by FCT – Fundação para a Ciência e Tecnologia by project reference 2021.06734.BD and DOI identifier 10.54499/2021.06734.BD, by Portuguese Recovery and Resilience Program (PRR), IAPMEI/ANI/FCT under Agenda C645022399-00000057 (eGamesLab), FCT LARSyS (DOI:10.54499/UIDB/50009/2020), FCT projects: 10.54499/LA/P/0083/2020; 10.54499/UIDP/50009/2020 & 10.54499/UIDB/50009/2020.

References

[1] Atul Adya, Paramvir Bahl, Jitendra Padhye, Alec Wolman, and Lidong Zhou. 2004. A multi-radio unification protocol for IEEE 802.11 wireless networks. In Proceedings of the IEEE 1st International Conference on Broadnets Networks (BroadNets'04). IEEE, Los Alamitos, CA, 210–217.
[2] https://doi.org/10.1109/BROADNETS.2004.8
[3] Carl DiSalvo, Phoebe Sengers, and Hrönn Brynjarsdóttir. 2010. Mapping the landscape of sustainable HCI. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10), 1975–1984. https://doi.org/10.1145/1753326.1753625
[4] Annika Stechemesser, Nicolas Koch, Ebba Mark, Elina Dilger, Patrick Klösel, Laura Menicacci, Daniel Nachtigall, Felix Pretis, Nolan Ritter, Moritz Schwarz, Helena Vossen and Anna Wenzel. 2024. Climate policies that achieved major emission reductions: Global evidence from two decades. Journal of Science, 385, 6711 (Aug 2024), 884-892. DOI: 10.1126/science.adl6547
[5] Rabab Saleh and Alexander Brem. 2023. Creativity for sustainability: An integrative literature review. Journal of Cleaner Production, 388 (February 2023). https://doi.org/10.1016/j.jclepro.2023.135848
[6] Swagata Sharma, Venkat Duvasulu, Berk Clik, Siddharth Suryanarayanan, Timothy Hausen, Antony Maciejewski, Howard Jay Siegel.Metrics-Based Assessment of Sustainability Demand Response. In IEEE 19th International Conference on HPCC/SmartCity/DSS, (2017), 130-137. doi: 10.1109/HPCC-SmartCity-DSS.2017.17.
[7] Sudhendu Rai, and Marc Daniels. 2015. An event-log analysis and simulation-based approach for quantifying sustainability metrics in production facilities," 2015 Winter Simulation Conference (WSC), pp. 1033-1043, doi: 10.1109/WSC.2015.7408231
[8] Ecochain - Life Cycle Assessment (LCA). https://ecochain.com
[9] Jason Ord, Ellen Chappell, Scott Canonico, and Tim Strecker. Product environmental metrics for printers. ACM SIGMETRICS Performance Evaluation Review, 37, 4, (March 2010), 79-83.
[10] https://doi.org/10.1145/1773394.1773408
[11] Ana Rodrigues, Diogo Cabral, and Pedro Campos. 2024. Fostering Informed Decision-Making: A Visual and Textual Approach to Sustainability Metrics for Idea Selection. International Conference on Graphics and Interaction (ICGI' 24), 1-8. 10.1109/ICGI64003.2024.10923803.
[12] Ben Shneiderman. 2000. Creating creativity: user interfaces for supporting innovation. ACM Transactions on Computer-Human Interaction 7, 1 (March 2000), 114–138. https://doi.org/10.1145/344949.345077
[13] Jonas Frich, Lindsay MacDonald Vermeulen, Christian Remy, Michael Mose Biskjaer, and Peter Dalsgaard. 2019. Mapping the Landscape of Creativity Support Tools in HCI. In CHI Conference on Human Factors in Computing Systems Proceedings (DIS '19), 1-18.
[14] https://doi.org/10.1145/3290605.3300619
[15] Anca Serbanescu and Frank Nack. 2023. Human-AI system co-creativity for building narrative worlds. In De Sainz Molestina, D., Galluzzo, L., Rizzo, F., Spallazzo, D., (eds.), IASDR 2023: Life-Changing Design (October 2023), 9-13, October. https://doi.org/10.21606/iasdr.2023.293
[16] Antti Oulasvirta and Kasper Hornbæk. 2016. HCI Research as Problem-Solving. In Proceedings of the Conference on Human Factors in Computing Systems (CHI '16), 4956–4967. https://doi.org/10.1145/2858036.2858283
[17] Joy Paul Guilford. 1950. Creativity. American Psychologist, 5, 9, 444- 454. https://doi.org/10.1037/h0063487
[18] Arthur Cropley. 2010. In Praise of Convergent Thinking. Creativity Research Journal 18, 3 (June 2010), 391-404. https://doi.org/10.1207/s15326934crj1803_13

[19] Bela H. Banathy. 1996. Designing Social Systems in a Changing World. Springer New York, NY.

[20] Special Eurobarometer 501. (2020). Attitudes of European citizens towards the Environment Report. https://europa.eu/eurobarometer/surveys/detail/2257

[21] Filipe Quintal, Clinton Jorge, Valentina Nisi, and Nuno Nunes. 2016. Watt-I-See: A Tangible Visualization of Energy. In Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI '16), 120–127. https://doi.org/10.1145/2909132.2909270

[22] Marta Ferreira, Valentina Nisi, and Nuno Nunes. 2023. Interactions with Climate Change: a Data Humanism Design Approach. In Proceedings of the Designing Interactive Systems Conference (DIS '23), 1325–1338. https://doi.org/10.1145/3563657.3596003

[23] Tim Coughlan and Peter Johnson. 2008. Idea management in creative lives. In Extended Abstracts on Human Factors in Computing Systems (CHI EA '08). Association for Computing Machinery, pp 3081–3086. https://doi.org/10.1145/1358628.1358811

[24] Angel Hsing-Chi Hwang. 2022. Too Late to be Creative? AI-Empowered Tools in Creative Processes. In Extended Abstracts of the 2022 Conference on Human Factors in Computing Systems (CHI EA '22), 1–9. https://doi.org/10.1145/3491101.3503549

[25] Orit Shaer, Angelora Cooper, Osnat Mokryn, Andrew L Kun, and Hagit Ben Shoshan. 2024. AI-Augmented Brainwriting: Investigating the use of LLMs in group ideation. In Proceedings of the Conference on Human Factors in Computing Systems (CHI '24), 1–17. https://doi.org/10.1145/3613904.3642414

[26] Michael Muller, Lydia B Chilton, Anna Kantosalo, Q. Vera Liao, Mary Lou Maher, Charles Patrick Martin, and Greg Walsh. 2023. GenAICHI 2023: Generative AI and HCI at CHI 2023. In Extended Abstracts of the Conference on Human Factors in Computing Systems (CHI EA '23), 1–7. https://doi.org/10.1145/3544549.3573794

[27] Selcuk Acar. 2023. Creativity assessment, research, and practice in the age of artificial intelligence. Creativity Research Journal (2023), 1–7.

[28] Gabriel Enrique Gonzalez, Dario Andres Silva Moran, Stephanie Houde, Jessica He, Steven I. Ross, Michael Muller, Siya Kunde, and Justin D. Weisz. 2024. Collaborative Canvas: A Tool for Exploring LLM Use in Group Ideation Tasks. In Proceedings of the Conference on Intelligent User Interfaces (IUI 2024). https://ceur-ws.org/Vol-3660/paper4.pdf

[29] Bryan Wang, Yuliang Li, Zhaoyang Lv, Haijun Xia, Yan Xu, and Raj Sodhi. 2024. LAVE: LLM-Powered Agent Assistance and Language Augmentation for Video Editing. In Proceedings of the Conference on Intelligent User Interfaces (IUI '24), 699–714. https://doi.org/10.1145/3640543.3645143

[30] Savvas Petridis, Nicholas Diakopoulos, Kevin Crowston, Mark Hansen, Keren Henderson, Stan Jastrzebski, Jeffrey V Nickerson, and Lydia B Chilton. 2023. AngleKindling: Supporting Journalistic Angle Ideation with Large Language Models. In Proceedings of the Conference on Human Factors in Computing Systems (CHI '23), 1–16. https://doi.org/10.1145/3544548.3580907

[31] Janet Rafner, Michael Mose Biskjær, Blanka Zana, Steven Langsford, Carsten Bergenholtz, Seyedahmad Rahimi, Andrea Carugati, Lior Noy, and Jacob Sherson. 2022. Digital Games for Creativity Assessment: Strengths, Weaknesses and Opportunities. Creativity Research Journal 34, 1 (2022), 28–54. https://doi.org/10.1080/10400419.2021.1971447

[32] Roni Reiter-Palmon, Boris Forthmann, and Baptiste Barbot. 2019. Scoring divergent thinking tests: A review and systematic framework. Psychology of Aesthetics, Creativity, and the Arts 13, 2 (2019), 144.

[33] Weikai Yang, Mengchen Liu, Zheng Wang, and Shixia Liu. 2024. Foundation Models Meet Visualizations: Challenges and Opportunities. Journal Computational Visual Media, 10, 3 (June 2024), 399-424. https://doi.org/10.48550/arXiv.2310.05771

[34] Sabrina Scuri, Marta Ferreira, Nuno Jardim Nunes, Valentina Nisi, and Cathy Mulligan. 2022. Hitting the Triple Bottom Line: Widening the HCI Approach to Sustainability. In Proceedings of CHI Conference on Human Factors in Computing Systems (CHI '22), 1–19. https://doi.org/10.1145/3491102.3517518

[35] Aaron Bangor, Philip T Kortum, and James T Miller. 2008. An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction 24, 6 (July 2008), 574–594. https://doi.org/10.1080/10447310802205776

[36] Erin Cherry and Celine Latulipe. 2014. Quantifying the Creativity Support of Digital Tools through the Creativity Support Index. ACM Trans. Comput.- Hum. Interact. 21, 4, Article 21 (June 2014), 25 pages. https://doi.org/10.1145/2617588

[37] Jeongseok Oh, Seungju Kim, and Seungjun Kim. 2024. LumiMood: A Creativity Support Tool for Designing the Mood of a 3D Scene. In Proceedings of the Conference on Human Factors in Computing Systems (CHI '24), 1–21. https://doi.org/10.1145/3613904.3642440

[38] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In Proceedings of the Conference on Human Factors in Computing Systems (CHI '19), 1–13. https://doi.org/10.1145/3290605.3300233

[39] Guglielmo Faggioli, Laura Dietz, Charles L. A. Clarke, Gianluca Demartini, Matthias Hagen, Claudia Hauff, Noriko Kando, Evangelos Kanoulas, Martin Potthast, Benno Stein, and Henning Wachsmuth. 2024. Who Determines What Is Relevant? Humans or AI? Why Not Both? Commun. ACM 67, 4 (April 2024), 31–34. https://doi.org/10.1145/3624730

[40] Daniel Buschek, Lukas Mecke, Florian Lehmann, Hai Dang. 2021. Nine Potential Pitfalls when Designing Human-AI Co Creative Systems. In Proceedings of the Conference on Intelligent User Interfaces. https://doi.org/10.48550/arXiv.2104.00358

[41] Barrett R Anderson, Jash Hemant Shah, and Max Kreminski. 2024. Homogenization Effects of Large Language Models on Human Creative Ideation.In Proceedings of the Conference on Creativity & Cognition (C&C '24), 413–425. https://doi.org/10.1145/3635636.3656204