

Article

Designing Home Automation Routines Using an LLM-Based Chatbot

Mathyas Giudici ^{*}, Luca Padalino , Giovanni Paolino , Iliaria Paratici , Alexandru Ionut Pascu 
and Franca Garzotto 

Department of Electronics Information and Bioengineering (DEIB), Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy; luca.padalino@mail.polimi.it (L.P.); giovanni.paolino@mail.polimi.it (G.P.); ilaria.paratici@mail.polimi.it (I.P.); alexandruionut.pascu@mail.polimi.it (A.I.P.); franca.garzotto@polimi.it (F.G.)
* Correspondence: mathyas.giudici@polimi.it

Abstract: Without any more delay, individuals are urged to adopt more sustainable behaviors to fight climate change. New digital systems mixed with engaging and gamification mechanisms could play an important role in achieving such an objective. In particular, Conversational Agents, like Smart Home Assistants, are a promising tool that encourage sustainable behaviors within household settings. In recent years, large language models (LLMs) have shown great potential in enhancing the capabilities of such assistants, making them more effective in interacting with users. We present the design and implementation of *GreenIFTTT*, an application empowered by GPT4 to create and control home automation routines. The agent helps users understand which energy consumption optimization routines could be created and applied to make their home appliances more environmentally sustainable. We performed an exploratory study (Italy, December 2023) with N = 13 participants to test our application's usability and UX. The results suggest that *GreenIFTTT* is a usable, engaging, easy, and supportive tool, providing insight into new perspectives and usage of LLMs to create more environmentally sustainable home automation.

Keywords: conversational agents; home automation; domestic sustainability; LLM; TAP



Citation: Giudici, M.; Padalino, L.; Paolino, G.; Paratici, I.; Pascu, A.I.; Garzotto, F. Designing Home Automation Routines Using an LLM-Based Chatbot. *Designs* **2024**, *8*, 43. <https://doi.org/10.3390/designs8030043>

Academic Editors: Saher Javaid and Yuto Lim

Received: 29 March 2024

Revised: 29 April 2024

Accepted: 6 May 2024

Published: 13 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The constant growth in greenhouse gas emissions, primarily caused by human activity, contributes to a concerning escalation in global temperatures [1]. The residential sector's increased electricity use contributes to a considerable amount of global energy consumption [2–4].

Interactive and innovative technologies, as well as Sustainable Human–Computer Interaction [5] research, can play a crucial role in delivering new digital and innovative solutions that can help raise awareness and lead individuals toward more sustainable habits [6]. Digital systems that provide eco-related feedback, such as the real-time visualization of energy consumption, have been found to encourage responsible electricity usage [7]. Especially in home automation contexts [8,9], conversational-based assistants are among the most popular interfaces for interacting with residential IoT networks [10].

The advent of large language models (LLMs) in the field of Artificial Intelligence has revolutionized various applications, including text summarization, story and creative generation, and code development [11]. While LLMs have found applications in diverse domains such as robotics [12] and video games [13], their integration into smart home automation [14], particularly in environmental sustainability, is still largely unexplored. Previous work by Giudici et al. [15] has shown that LLMs are generally good at answering general questions related to environmental sustainability but suffer in accuracy on specific topic-related questions. However, designing and exploring specific applications to advise users on household sustainable practices still needs to be explored.

In this context, our research introduces *GreenIFTTT*, an innovative and novel web-based conversational agent (powered by the GPT4 model) that encourages individuals, particularly homeowners, to adopt environmentally conscious habits within their households. We present the design and the user experience principles followed to create the *GreenIFTTT* application, with features that educate users on sustainable electricity consumption, reducing the overall energy consumption and simplifying routine creation with smart technologies. The system focuses on creating routines: automation sequences within the home environment based on the sequential execution of specific activities triggered by various conditions. The primary interaction paradigm is conversational, leveraging the capabilities of the LLM to assist users in locating and monitoring their smart appliances. The system also integrates data from connected sensors, giving users real-time insights into their daily routines.

In addition, we want to support the system design contribution by reporting an exploratory study conducted to assess *GreenIFTTT* features and answering the following research questions:

- RQ1 What is the *user experience* of a conversational agent powered by a large language model for promoting sustainable household practices?
- RQ2 What is the *engagement and likability* of a conversational agent powered by a large language model for promoting sustainable household practices?
- RQ3 What is the *usability* of a conversational agent powered by a large language model for promoting sustainable household practices?

The results indicate positive user experiences, with high scores on valuable and well-known scales such as the User Experience Questionnaire (UEQ), Parasocial Interaction (PSI), and System Usability Scale (SUS). Participants found the system supportive, easy to use, and engaging, highlighting its potential for promoting sustainable practices in home environments. Our results provide valuable insights into such systems' effectiveness and user acceptance, paving the way for further research using LLM-based conversational agents to promote environmental sustainability in home settings.

This paper is organized as follows. Section 2 frames the current state-of-the-art applications in home automation environments, such as large language models and chatbots for environmental sustainability. Section 3 describes the design, user experience, and high-level implementation of *GreenIFTTT*. Section 4 delves into the methodology of the user evaluation performed, while Section 5 reports and discusses the results, also addressing the limitations of the present work. Finally, Section 6 presents the conclusion and presents future directions for research work.

2. State-of-the-Art Applications

The two pillars on which the presented research relies are home automation environments and conversational technology, focusing on large language models. In order to describe the state-of-the-art applications of such pillars in depth, this section is organized as follows. Section 2.1 delves into the landscape of home automation environments and their primary interaction and programming methods; Section 2.2 presents the landscape of conversational technology—also applied in home automation environments—and the recent literature on large language models.

2.1. Home Automation Environments

Smart environments are physical spaces enhanced with sensing, actuation, communication, and computation capabilities to adapt to users' preferences, requirements, and specific needs [16]. They encompass various applications and settings, from smart homes to smart cities and factories [17]. The development of smart environments faces challenges such as precise activity recognition, effective localization systems, and the need for a smooth transition from traditional to smart environments [18,19]. Emerging wearable devices and wireless communication technologies are expected to further optimize the resource efficiency, comfort, and safety of such environments [17].

In this vast landscape, *home automation* environments refer to systems that enable the management of household activities through computerized control, offering benefits such as comfort, security, and energy efficiency [20]. Advancements in the Internet of Things (IoT) research have enabled the remote control of devices and simplified everyday tasks, making home automation—in general—more accessible and user-friendly. In order to control home appliances, Trigger–Action Programming (TAP) is a simple programming model available to users to create rules automating the behavior of smart homes, devices, and online services [21]. Householders can create TAP rules (also called recipes, applets, or routines) using a simple conditional structure—summarized as if ‘condition’ then ‘action’—using applications designed to be user friendly and accessible without a programming background [22]. For instance, the commercial service IFTTT (<https://ifttt.com/>, accessed on 5 May 2024) (i.e., If This Then That) enables end users to easily use TAP with the most-used commercially available home devices.

However, according to [23,24], the oversimplification of existing TAP systems limits the expressivity of the programs that can be created, leading to inconsistencies in interpreting the behavior of TAP and errors in creating programs with a desired behavior. Another limitation is the lack of standard open protocols; each vendor allows its own product in a closed environment and requires the user to use several applets to manage a fully integrated environment [25]. Still, empirical evidence from Ur et al. [22] showed how users tend to create their TAP recipes rather than using existing ones shared by other users or appliance manufacturers. Finally, Heo et al. [26] discussed how existing IoT frameworks read sensor data periodically, independent of real-time constraints. For example, the IFTTT framework polls sensor data every 15 min, and real-time APIs cause sensors to send unnecessary messages that do not affect any TAP recipe and waste battery power. In their work, the authors introduced the RT-IFTTT language, which allows users to specify real-time sensor constraints. A central manager analyzes the relationships between sensors and the TAP, calculating polling intervals for each trigger condition and turning off the sensor when unnecessary, saving battery and energy.

While TAP has demonstrated great practicality in customizing smart home devices, allowing end users to express a wide range of desired behaviors, and machine learning algorithms have improved the adaptability and functionality of such programs [21], the landscape is still sparse in sustainable living. However, it is worth reporting that in a broader panorama, as in the Sustainable Human–Computer Interaction [5] literature, we can find the evaluations of different visual systems [27] used to reduce the energy consumption of households [28–30] or intelligent technologies that shift people’s thermal comfort consumption [31]. Still, in this field, Alan et al. [32] presented an interactive IoT system designed to help manage energy costs, offering flexible autonomy and detailed information about its operation. Finally, Bang et al. [33] proposed a gamified activity in the context of domestic sustainability, while Beheshtian et al. [34] developed a social robot for sustainable living in a block of flats. These examples from Sustainable HCI can inform the design of new TAP solutions to help people have more sustainable home environments (e.g., designing applets to reduce energy consumption).

2.2. Large Language Models

A *large language model (LLM)* is a subset of Artificial Intelligence algorithms that uses deep learning techniques (e.g., a transformer model) and massively large datasets to understand, summarize, generate, and predict new content [11,35]. Nowadays, LLMs are mainly employed for text summarization [36] and Next Sequence Prediction [37] (i.e., predict the subsequent element or event in a sequence based on patterns and information present in the history of the sequence), the automation and efficiency of repetitive tasks, and language translation [11]. The pioneer model presented in the LLM world was *GPT-2*, released by OpenAI in 2019, which was trained on 1.5 billion parameters, 7 k books, and 8 million web pages [38]. In recent years, other big tech companies have published different versions of LLMs, like Google’s Bard and Microsoft’s Bing AI. Still, OpenAI improved the model, presenting ChatGPT/GPT-3.5 [39] and GPT4. In addition, these models have been

applied in different contexts, for example, to enrich the language capabilities of robots [12], videogame characters [13], and smart home automation [14]. Although LLMs have positive impacts, it is also important to emphasize that there are negative aspects that can impact society [11,40]. For instance, they can generate misinformation and disinformation and amplify biases in training data, raising ethical and data privacy issues [11,41].

Large language models are grounded on long scientific research in natural language processing, aiming to create computers that can interact with the user using natural language [42] (i.e., interpret and generate human language). A *digital tool* that interacts with users using natural language is defined in the scientific literature as a *Conversational Agent (CA)* [9]. CAs have been used to engage users in text-based information-seeking and task-oriented dialogues for many applications [43]. For example, they are integrated into physical devices (such as Alexa and Google Home [10]) and are available in many contexts of everyday life, like in phones (like Siri, the Apple virtual assistant [44]), cars, and online consumer assistance [45].

CAs are also considered promising in the environmental sustainability domain [8,9]. Traditional rule-based chatbots have previously been used to deliver energy feedback [46,47], suggesting sustainable mobility [48] or reducing food waste [49]. Still, Ramasubbu et al. [50] presented a chatbot to optimize the schedule of switching off smart plugs in an office. Furthermore, Gunawardane et al. [51] proposed an example using a data-driven chatbot to suggest recipes with leftover foods. Finally, Giudici et al. [15] presented an evaluation of LLMs to create a hybrid conversational agent that can trigger devices in a home automation environment and address users' open-domain questions.

In the home automation context, LLMs unlock a more natural interaction between users and space compared to the interactions that task-specific systems can provide [14]. In addition, they can overcome the limitations of traditional conversational agents integrated with TAP, which include the limited language pattern it must follow and triggering and interfacing with single devices, leading to a low-integrated ecosystem perception [52]. Still, such a feature is possible thanks to a significant exposure of LLM to daily situations and requests inserted into their diverse training data, on which they can be fine-tuned (i.e., using previously available datasets [53]). For instance, Fast et al. [54] presented that an LLM trained only on written works of fiction is able to recognize everyday activities based on semantic relationships between objects and the activities that they are frequently used in. King et al. [14] presented Sasha, an LLM able to realize goal-oriented home automation in domotic environments. They evaluated the ability of such an LLM to create parsable JSON files to be applied in a real-case setting using a data source of smart home action plans. Similarly, Li et al. [55] presented ChatIoT, a zero-code rule generator, to improve the quality of TAP generation while reducing the tokens required for the prompt to the LLM. In addition, a set of confirming rules in the interaction pipeline allows for refining the TAP recipe to be more accurate and safely executed. Finally, Nascimento et al. [56] explored the usage of generative AI versus engineering-crafted coding to create a control structure for an IoT application. The results pointed out how there were cases in which humans outperformed AI algorithms and others where they did not, with the experiment validity being highly impacted by the background experience and skills of the engineers enrolled.

To the best of our knowledge, an aspect that still needs to be addressed is the usage of LLMs to advise users on how to be more sustainable in their houses. In particular, we focused on using GPT4 models to help households create TAP routines that make them more eco-sustainable and save electric energy.

3. The System

We present the design of *GreenIFTTT*, a web-based conversational agent designed to encourage individuals, particularly homeowners, to cultivate environmentally conscious habits within their households, embracing an innovative approach toward emerging technologies and smart devices. The key features of the system are as follows:

1. **Educating Users on Sustainable Electricity Consumption.** The system pushes users to adopt more sustainable behaviors regarding electricity consumption.

2. **Stimulating Cost Optimization in Utility Bills.** The system provides mechanisms and tips to optimize utility bills, ensuring cost-effective energy management.
3. **Energy Consumption Reduction.** The system actively works to induce people to reduce overall energy consumption, promoting an eco-friendly lifestyle through energy consumption shifting.
4. **Simplifying Everyday Life:** The system simplifies users' daily routines, making integrating smart technologies seamless and hassle-free.

The application focuses on creating multiple *routines*. Routines are Trigger–Action Programming automation components within the home environment based on the sequential execution of certain activities. An *activity* is the change in state of an appliance (e.g., turning on the washing machine) when one or more *conditions* occur. Finally, conditions may occur depending on the state of another home device, on some external API (e.g., weather or solar forecast services), or on temporal conditions (timers in a specific span). Figure 1 shows different app pages for creating, viewing, and controlling routines.

Finally, as already described in Section 2, existing smart home assistants lack linguistic patterns to activate devices (or routines). Usually, they cannot create a fully integrated ecosystem enabling the creation of TAP involving multiple devices and trigger conditions. *GreenIFTTT* is a new digital solution to overcome such limitations, using LLMs' new cutting-edge generative power.

This section is organized as follows to organize a complete system description. Section 3.1 describes the design of the user experience of the entire *GreenIFTTT* application, while Section 3.2 illustrates the user workflows, particularly the generation of a home automation routine using the chatbot. Finally, Section 3.3 details the implementation, and an in-depth definition of the integration of the GPT model is provided in Section 3.4.

3.1. User Experience

The primary interaction paradigm employed is conversational. This approach places the core of the user experience in the system's ability to actively engage users, enabling them to leverage the extensive capabilities of an LLM. Users receive assistance locating and monitoring their smart appliances by interacting with a fine-tuned version of *GPT4*. They can also access data collected from connected sensors, streamlining their daily routines. *No particular on-top machine learning algorithm is used to inform the routine creation; this relies only on the prediction of the LLM, driven by the energy consumption values included in the prompt (see Section 3.4).*

To analyze and define existing user experience with the application, we employed the *5W+H heuristic* framework proposed by Jia et al. [57]. The model addresses the well-known WH questions, *who, what, when, where, why, and how (5W+H)*, to explore the various elements with which the user interacts.

Who Homeowners seek to develop new eco-friendly routines in their homes with an open-minded approach toward embracing new technologies and smart devices. The primary goal is to reduce power consumption and cultivate a more eco-sustainable lifestyle.

What The issue the system aims to resolve is to simplify users' everyday interactions with smart devices [58]. The system was designed to introduce ease and efficiency into managing these devices.

When The system is most suitable for house owners who are open-minded toward innovative technologies and willing to leverage the significant advancements in the field of large language models.

Where The system was designed for implementation in residential settings where homeowners seek to integrate smart technologies seamlessly into their daily routines. It is adaptable to various housing environments, from urban apartments to suburban homes, fostering a sustainable and energy-efficient lifestyle. In addition, we created a web-based application to allow people to interact with the *GreenIFTTT* agent from desktops and mobile devices.

Why The motivation behind the system lies in addressing the growing need for simplicity and efficiency in managing smart devices within households.

How The system aims to streamline daily routines by incorporating new AI technologies, enhance energy efficiency, and contribute to more sustainable electric consumption. The system offers homeowners a convenient way to embrace eco-friendly practices while enjoying the benefits of innovative solutions.

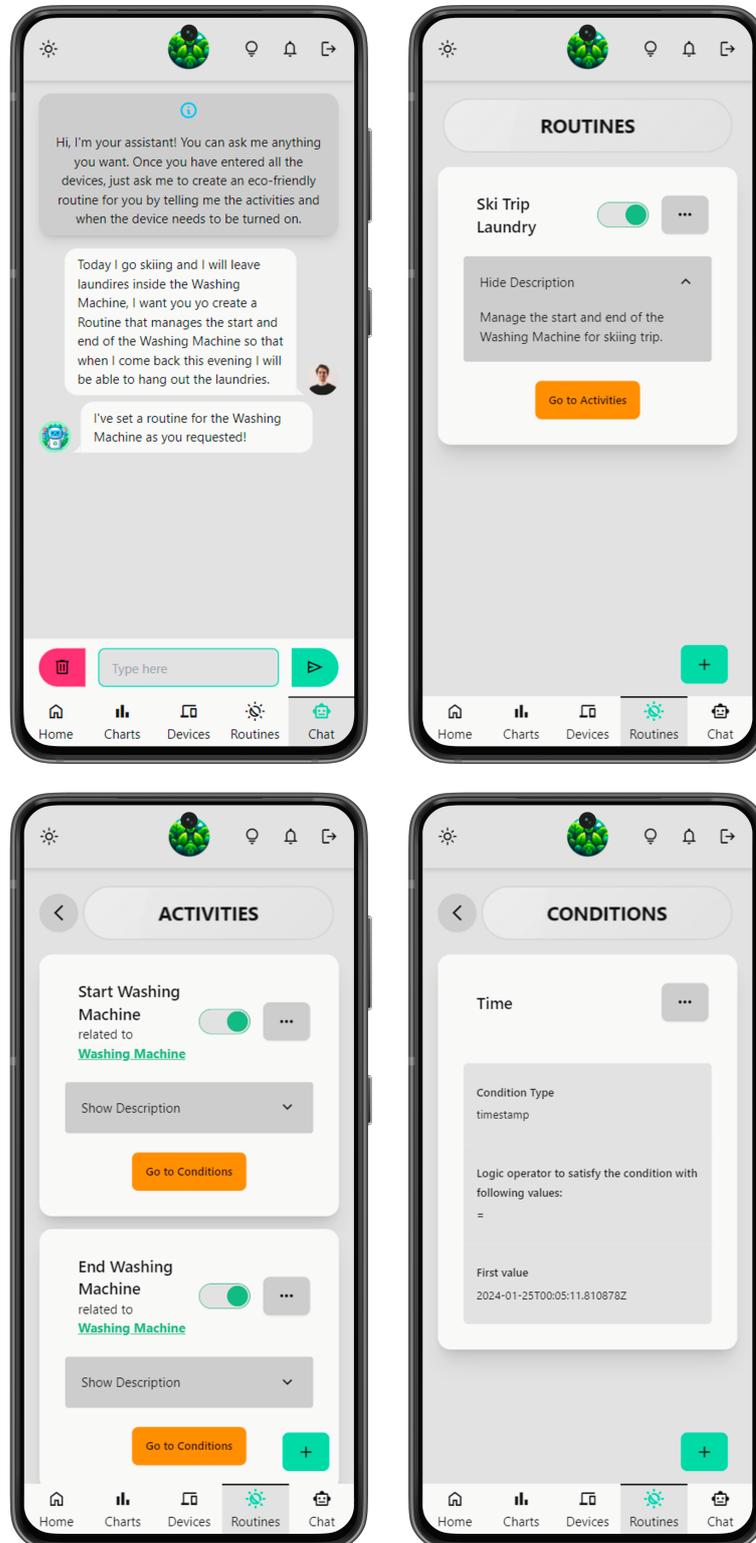


Figure 1. Overview of GreenIFTTT application.

Finally, following the principles presented by Liao and Vaughan [59], the application design aimed to keep users well informed about ongoing processes through essential and straightforward visuals. For example, we included appropriate feedback mechanisms (e.g., toasters while the model is computing the response), delivering response messages within a reasonable time frame and ensuring that users feel guided during the conversational process. Actions automatically performed by the system are stated via text in a short, clear, and comprehensible manner, fostering a sense of transparency.

The system also offers users a high level of freedom, allowing them to customize every *routine* component and empowering users to tailor the system to their specific needs and preferences.

Lastly, the entire application wants to be the following:

- **Simple.** The interface was designed to be straightforward, minimizing complexity and facilitating easy navigation.
- **Intuitive.** The system’s features and functions are accessible and usable without a steep learning curve.
- **Easy to use.** The overall user experience was designed for accessibility and user friendliness, ensuring a smooth and efficient interaction.

The above features collectively contribute to fostering a positive user experience, aligning with the overarching goal of creating a user-friendly and highly adaptable system to individual preferences.

3.2. User Workflows

User workflows represent users’ steps to navigate the system and complete desired actions [60]. Such workflows are crucial for designing an intuitive and efficient user experience. The system offers various functionalities accessible through a designated tab in the application. Such functionalities are represented in Figure 2, and they include the following:

- **Dashboard.** Provides an overview of user energy information and graphical trends.
- **Charts.** Allow the exploration of detailed power consumption and billing data.
- **Devices.** Enable connection with and the monitoring of smart home appliances.
- **Routines.** Facilitate the creation and management of automated routines.
- **Chat.** Integrates with a pre-trained ChatGPT model for routine creation through text prompts.

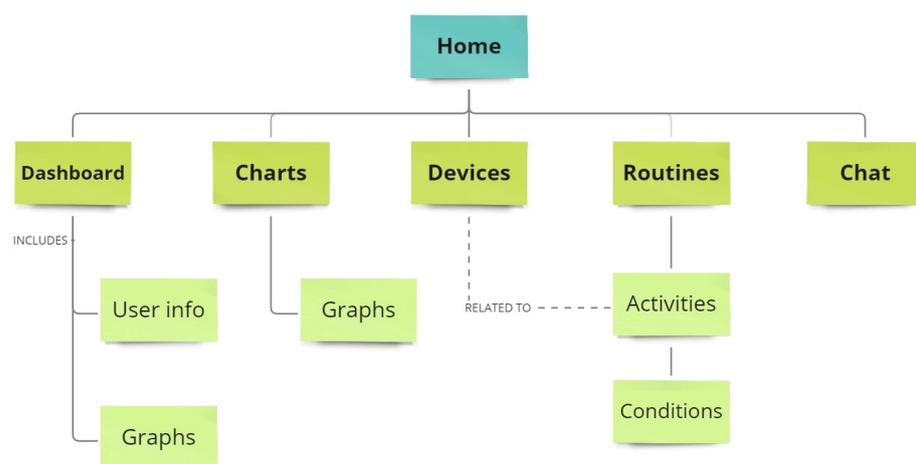


Figure 2. Overview of the main components of *GreenIFTTT*.

One remarkable user workflow involves creating a home automation routine using the chatbot (see Figure 3). This workflow starts with user initiation through the *Chat* tab. The user submits a text prompt describing the desired routine. The GPT model processes

the prompt and generates a JSON response. The backend system parses this response to implement the custom routine and provides a feedback message to the user (in the chat). Following generation, the user can review the created routine. They have the option to manually edit or delete the routine entirely. Alternatively, the user can instruct the chatbot to perform these actions on their behalf. Thus, this user workflow demonstrates the system’s ability to guide users through task completion, offering manual and chatbot-assisted options for creating custom routines.

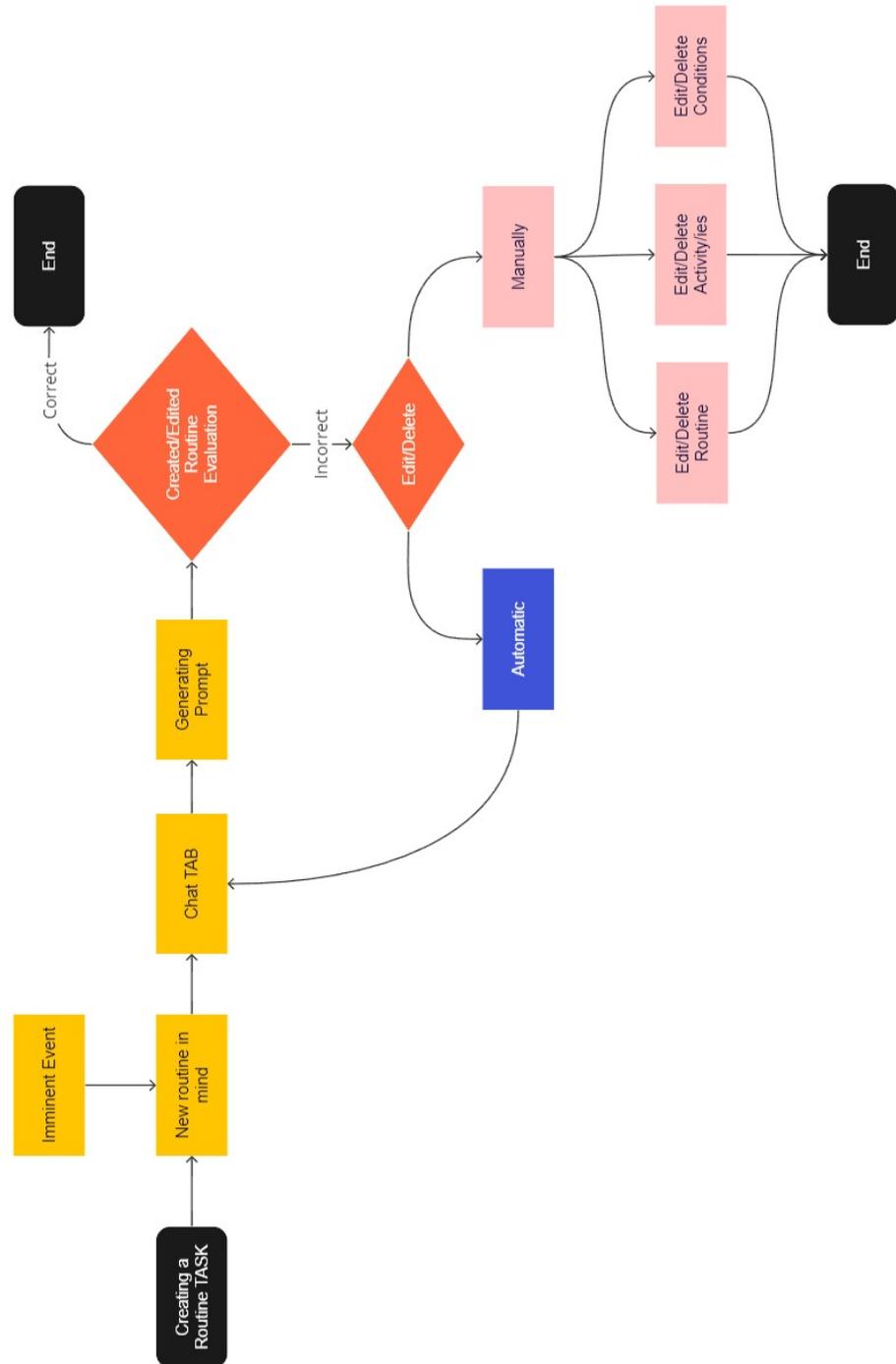


Figure 3. Schema representing the user workflow followed to create a home automation routine using the chatbot.

3.3. Implementation

We implemented the solution considering the principles of *scalability*, *modularity*, and *ease of implementation*, and possible future extensions. The project was configured to integrate with existing IoT devices, enabling real-time control. However, for the empirical evaluation of the system, we simulated device behavior by emulating their connections within a local network.

Our chosen technological solution is a single-page *web application*, adopting the *three-tier client–server architecture*, segmented into a *frontend*, *backend* (application logic), and *data layer* or external connections. Figure 4 shows a high-level system overview.

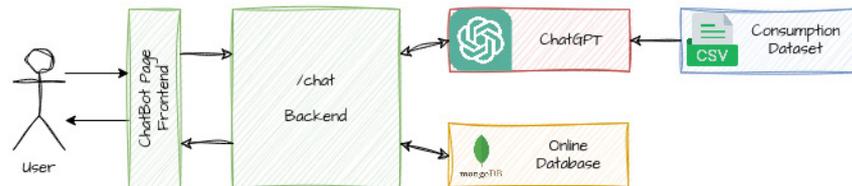


Figure 4. High-level system overview.

The application’s *backend* plays a crucial role in providing the main logic of the application (routine handling and authentication), integration with external components, data storage in the database, and interfaces with the frontend. The backend is structured by separating the data and application logic layers for efficiency and code clarity.

In particular, as shown in Figure 5, the backend was developed using the *Node.js* framework and *Prisma ORM* to facilitate communication with the *MongoDB* database. The frontend was structured as a single-page web application developed with the *Vue.js* framework.

The software (version v1.0.0) (<https://gitlab.com/i3lab/GreenIFTTT>, accessed on 6 May 2024) is being released as open source to allow study repeatability, improvement, or any other suggestion from the scientific community.

3.4. Integration with GPT4

In *GreenIFTTT*, GPT4 plays a central role in processing user input. The prompt sent to GPT4 is fed with data previously entered by users, their home automation (in a JSON format), and live consumption of appliances. In addition, a publicly available dataset (<https://www.kaggle.com/datasets/ecoco2/household-appliances-power-consumption>, accessed on 6 May 2024) containing appliance-monitored consumption data is used to fine-tune the model, enhancing its responsiveness and ability to create more eco-related advice and interpret appliance consumption data.

When a user sends a message, the system checks for a *thread* previously opened by the user with GPT4; if no thread exists, one is created. A thread, which is considered an OpenAI API beta feature, represents and contains an ongoing conversation between the system and a specific user.

Creating a thread involves sending instructions to GPT4 to guide the model generation and shaping responses in a parsable format for the client (i.e., the backend with the app logic). A prompt engineering technique [61,62] was applied. Generation instructions (zero-shot prompting) are sent to the fine-tuned GPT4 model in the initial *system* prompt using the text provided in Appendix A.

According to such an initial prompt, GPT4 usually responds with a specific *JSON structure*, reported in Appendix B, consisting of key–value attributes. The *message* attribute is delivered directly to the user and stored in the database. Other operations are addressed depending on the message’s type, which can be either *chat* or *routines*. In the former, no further operations are performed on the database. In the latter case, additional JSON

attributes are analyzed to execute database operations, such as creating a *routine* involving *activities* and related *trigger conditions* to be satisfied.

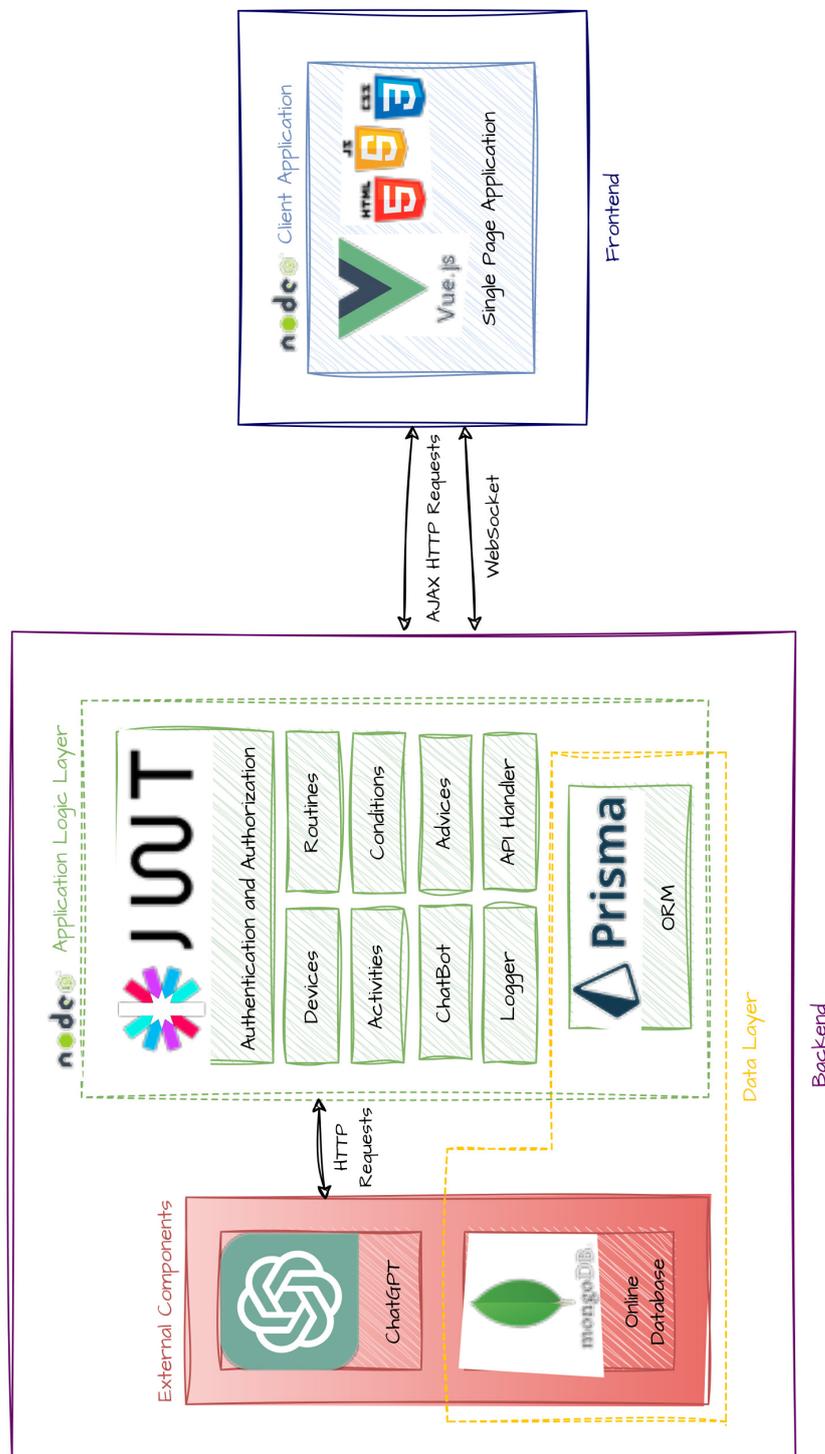


Figure 5. System overview with employed technologies.

Every exchange of messages between the user and GPT4, and vice versa, is stored in the database, ensuring that the conversation history is readily available as soon as the user loads the chat page, providing a comprehensive and continuous user experience.

Finally, discussing and highlighting some challenges for future work integrating GPT4 in domain-specific applications is crucial. As described above, the *thread* function of Ope-

nAI, which was in a beta version (preview) at the time of the study, was used, which resulted in slightly longer response generation times compared to traditional APIs. Furthermore, as described, the fine-tuning contributed (only in the early stages of interaction) to lengthening these times. In the future, it is necessary to evaluate the time generation of the *thread* function in a production environment (not in beta). Still, in a real-case environment, the creation of edge or mixed computing systems (with part of the LLM generation carried out in users' systems) is worth investigating.

4. Empirical Study

The exploratory evaluation of our application touches on three areas: *engagement*, as the ability to keep the subject engaged in the activity for a prolonged period; *likability*, which is defined as the degree of appreciation and the ease of use of the tool; and *usability*, which is the degree to which something is able or fit to be used. This also included an evaluation of the interaction paradigms.

4.1. Research Variables

Data gathering was performed using a web-based questionnaire. Participants answered questions referring to different quantitative scales, which were as follows:

- *The User Experience Questionnaire (UEQ)* [63] ($\alpha = 0.87$) was used—in its short version—to extract feedback about the user experience (UX) of an interactive digital tool;
- *The Parasocial Interaction (PSI) scale* [64,65] ($\alpha = 0.91$) measures the degree to which participants feel connected to and attached to the system;
- *The System Usability Scale (SUS)* [66,67] ($\alpha = 0.74$) was used to determine participants' perceptions of the interface's usability.

The UEQ scale presents a set of items that contains two different adjectives; participants selected, on a seven-stage scale, which of the opposing terms for each item better described the system (e.g., the first item of the UEQ is *obstructive vs. supportive*; 1 is linked to “obstructive” and 7 to “supportive”). All the items in the PSI and SUS scales were evaluated using a seven-point Likert scale ranging from 1 (“Completely Disagree”) to 5 (“Completely Agree”).

4.2. Participants

The study involved 13 participants (3 females and 10 males) with a mean age of 27 years ($M = 27.85$, $SD = 11.58$). All study participants were recruited voluntarily using snowball sampling (started by our community, colleagues, and university students). The participants were mainly undergraduate or postgraduate students; from the qualitative results, they reported varying expertise in the home automation field (from those interested in home automation to those who did not consider it). All participants signed a consent form informing them about procedures, goals, and data treatment. The investigation was conducted using the same laptop in our research laboratory in Milan (Italy) in December 2023.

4.3. Procedure

The experimental protocol consisted of three phases (Figure 6) with a total session duration of about 15 min. Participants were asked to fill out general biographical information in the first phase. In addition, a researcher presented the scenario and the tasks of the study using a supplemental paper sheet. The scenario and tasks used in the study are detailed in Appendix C. During the second phase, participants were invited to interact with the system and complete the tasks presented in the previous phase. Finally, in the last phase, participants filled out a questionnaire with all the inquiries needed to assess the research variables presented in Section 4.1, which, respectively, were proposed by Laugwitz et al. [63], Tsai et al. [65], and Bangor et al. [67].

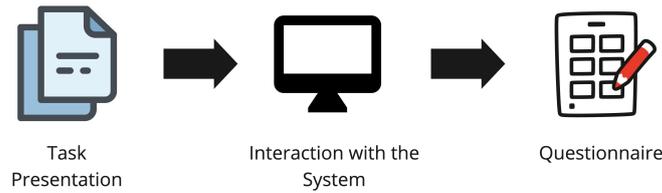


Figure 6. Empirical study procedure.

5. Empirical Results

This section reports the empirical evidence from the exploratory study described above. This section is organized as follows. Section 5.1 reports the descriptive results, while Section 5.2 discusses this evidence and attempts to answer the research questions, focusing on previous scientific results. Finally, Section 5.3 reports this study’s limitations.

5.1. Results

Participants reported (as also reported in Table 1 and shown in Figure 7) the application as more *supportive* rather than *obstructive* (M = 5.92, SD = 0.76), as *easy* (M = 6.08, SD = 1.04), *efficient* (M = 5.62, SD = 1.12), *clear* (M = 5.92, SD = 1.04), *exciting* (M = 5.23, SD = 1.24), *interesting* (M = 5.92, SD = 0.86), *inventive* (M = 5.85, SD = 1.14), and *leading edge* (M = 5.92, SD = 0.86).

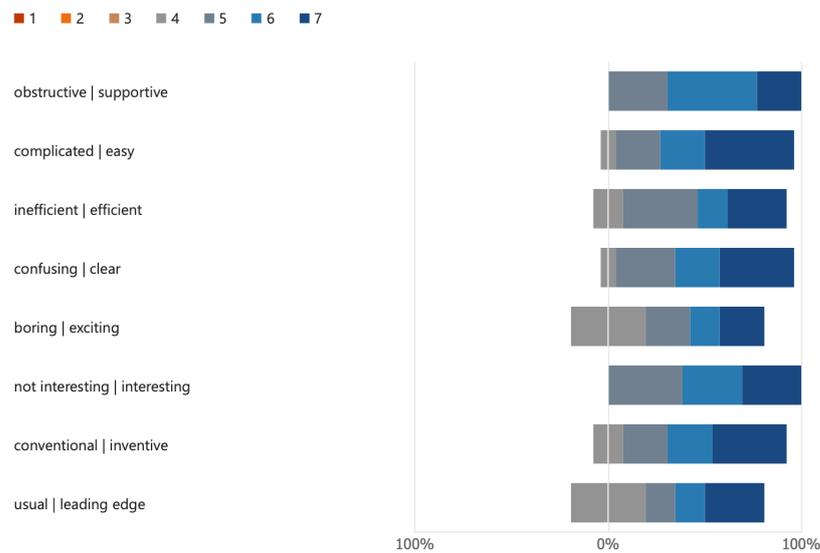


Figure 7. UEQ scale result representation.

Table 1. Descriptive results of the research variables.

	Variable	AVG	SD
UEQ scale	obstructive supportive	5.92	0.76
	complicated easy	6.08	1.04
	inefficient efficient	5.62	1.12
	confusing clear	5.92	1.04
	boring exciting	5.23	1.24
	not interesting interesting	5.92	0.86
	conventional inventive	5.85	1.14
	usual leading edge	5.38	1.33
PSI scale	PSI	5.18	1.42
	PD	4.85	1.38
	UE	4.69	0.98
	IS	6.35	0.49
SUS scale	SUS	83.79	5.07

The different items in the PSI scale indicated a *Perceived Dialogue (PD)* with a mean equal to 4.85 (SD = 1.38), *User Engagement (UE)* with an average value of 4.69 (SD = 0.98), *Interaction Satisfaction* with an average of 6.35 (SD = 0.49), and *Perceived Parasocial Interaction (PSI)* with a mean of 5.18 and a standard deviation of 0.49.

Finally, the SUS score had an average value of 83.79 (SD = 5.07).

5.2. Discussion

As stated in the introduction (Section 1), this preliminary evaluation aimed at understanding more about the engagement, likability, and usability of a large language model-powered conversational agent for promoting sustainable household practices. Regarding usability, the average SUS score was 83.79. According to Bangor et al. [67], the usability of our system is more than acceptable, with a grade of B and an adjective rating between “good” and “excellent”.

Our results on the Parasocial Interaction scale are greater than those previously determined by Tsai et al. [65], suggesting a positive and effective interaction likeability between participants and *GreenIFTTT*. In particular, users reported a higher *Interaction Satisfaction* and *Perceived Parasocial Interaction*. Such higher results could be related to using an LLM instead of a traditional rule-based chatbot. LLMs can engage users with their advanced language capabilities [11], almost keeping the main functionalities of traditional agents. Notably, in our context, we imposed on the LLM that “If you don’t know how to respond to me, you can ask me to repeat the question or to show you the available commands” (Appendix A, line 3). Still, from the scientific literature, we have examples of recent studies reporting the perceived high consciousness [68] and human likeness [69] that users attribute to such agents. In addition, Ross et al. [70] reported the high quality of the generated responses and the agent’s ability to assist users in specific domain tasks (e.g., produce or create code). Finally, our results are also better than those of the previous work by Giudici et al. [71], in which the authors evaluated the usage of a traditional rule-based chatbot to promote environmental sustainability in home environments. The UEQ output also confirmed such results and pointed out that participants found the application easy, clear, supportive, and interesting. In particular, the highest average score was obtained on the *easy* adjective. Considering that the experimental task was executed by participants who were not experts in the specific environmental sustainability field and generating good and reliable home automation represents a quite trivial task [23,24], we can argue that the approach presented in this paper of using an LLM to create home automation tasks can enable more people (even non-experts) to set up routines to make their home consumption more optimized and sustainable. Still, such results are aligned with the instructions in the LLM prompt, in particular, *You are a helpful assistant and should answer me clearly and it must be human-like, friendly, in a way that I see you as my best friend* (Appendix A, line 1 and 13). Finally, previous research by Zhang et al. [62] has indicated a connection between a user’s level of engagement with a digital application and their environmentally sustainable attitudes, corroborating the results obtained by our application.

5.3. Limitations

Our work presents limitations. First of all, for a more comprehensive evaluation, a sample that includes more participants who are more balanced and extended over a larger population in age and gender is needed. In addition, the study presented in this paper was conducted in a laboratory, using a hypothetical scenario, significantly impacting the ecological validity of the experience. Even though participants interacted with a working conversational agent, the appliances’ data were from a publicly available dataset, and users’ actions did not affect real devices. In addition, no pre- or post-questionnaire on the environmental attitudes of participants was undertaken. Finally, there was no comparative user study where participants interacted with *GreenIFTTT* and alternative systems (e.g., the basic IFTTT or other LLMs) to comprehensively evaluate the performance, usability, and user experience, running more advanced statistical analyses. For all the above reasons,

the results of our study are preliminary and insufficient to make a definitive claim on the effectiveness of *GreenIFTTT* in a real scenario. However, we provided empirical preliminary insights into this emerging topic by addressing our research questions.

Secondly, according to Rillig et al. [72], LLMs and other new innovative technologies (e.g., the Metaverse) directly and indirectly impact the environment and environmental research. One of the direct negative impacts is the high amount of energy that LLMs need to be trained and employed. In particular, Luccioni et al. [73] quantified the footprint of BLOOM's 176 B parameter LLM, considering the training equipment manufacturing, training the model, and model deployment (accessible via API endpoints). In the same way, Ref. [74] proposed a framework to compute the carbon footprint of different AI models and estimate the carbon emitted across usage. However, Ref. [75] reported comparing the usage of generative AI systems and human individuals performing equivalent writing and illustrating tasks. Contrary to expectations, AI systems released between 130 and 1500 times less carbon per page of text generated than their human counterparts, and similar results (310–2900 times less) were found for image generation. The authors discussed how generative AI is not a replacement for human tasks; however, it holds the potential to perform some activities with much lower carbon emissions. In our specific research context, no comparisons or specific results are reported as to whether LLMs can provide home automation that can reduce energy consumption and positively impact carbon emissions while also considering the footprint of LLMs in the overall energy balance.

6. Conclusions and Future Works

We presented the design and a preliminary evaluation of *GreenIFTTT*, a web-based conversational agent empowered by the GPT4 model, designed to encourage environmentally conscious habits within households. Leveraging the capabilities of large language models, *GreenIFTTT* aims to simplify users' daily routines and reduce overall energy consumption. This empirical study's results demonstrated and validated our research questions. Users reported a positive experience with the application (RQ1), also indicating a high level of engagement and likability (RQ2), as well as usability (RQ3).

Integrating GPT4 into a conversational agent for promoting eco-friendly practices represents a significant step toward leveraging advanced AI technologies for environmental sustainability, simplifying the creation of home automation. The system's effectiveness in providing personalized and facilitating interactions with domotic environments in natural language contributes to its potential impact on users' behavior toward more sustainable living.

Building on the foundation laid by this research, several avenues for a future research agenda emerge; such an agenda is dual for us and other researchers in the field. First, we aim to overcome our limitations (see Section 5.3) by conducting a more extensive study (involving a larger population). In addition, we aim to explore in-the-field studies in real-home environments to evaluate the long-term impact of *GreenIFTTT* on users' sustainable practices and energy consumption and validate its integration with real-case home appliances. In addition, we are already running a study comparing different LLMs on their generative capabilities to realize home automation; the next stage would also be to validate the effectiveness of such sustainable automation (i.e., the impact on reducing energy consumption). Future research could also focus on studying the different generative capabilities of LLMs and related end-user perceptions without limiting the comparison of LLMs on datasets of user interactions but opening up to comparisons with studies involving users in the field.

Author Contributions: Conceptualization, M.G., L.P., G.P., I.P. and A.I.P.; methodology, M.G. and F.G.; software, L.P., G.P., I.P. and A.I.P.; formal analysis, M.G.; investigation, F.G.; writing—original draft preparation, M.G.; writing—review and editing, L.P., G.P., I.P. and A.I.P.; supervision, F.G.; project administration, F.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Italian Ministry of University and Research (MUR) and the European Union (EU) under the PON/REACT project.

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of Politecnico di Milano University (Milan, Italy).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are contained within the article.

Acknowledgments: The authors would like to thank all the participants involved in the study.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the study's design; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A. GPT4 Initial Prompt

*

```

1 You are a helpful assistant and should answer me clearly.
2 You have to always respond to me only in a json format like the following:
  ↪ ${JSON.stringifyjsonModel} and nothing else.
3 If you don't know how to respond to me, you can ask me to repeat the question or to show you
  ↪ the available commands.
4 The message that you have to send me is in the 'message' field of the json.
5 Use ${JSON.stringifydevices} to gather information about devices.
6
7 'TYPE' is 'Chat' if there is the only text to show to the user and 'routines' is null,
8 'TYPE' is 'Chat' if I ask you to create a routine but you don't find that device in:
  ↪ ${JSON.stringifydevices}, so you will reply to me with an error message, and 'routines'
  ↪ is null, otherwise the 'TYPE' is 'Routine' and 'routines' is filled in.
9 'TYPE' is 'Routine' if there is some routine or activity going on to be created or updated
  ↪ or deleted and the 'routines' list is filled in.
10 If I don't specify any ROUTINE_NAME, you put the activity inside the routine DAILY by
  ↪ default, if DAILY does not exist create it.
11 If I specify also the ROUTINE_NAME or I put some specific words vacation, trip, daily, etc.,
  ↪ first you check if the routine already exists using this file: ${JSON.stringifyroutines},
  ↪ if already exists, put the activity inside it, otherwise create a new routine with the
  ↪ name as general as possible and put the activity inside the routine with that name.
12 If I ask you to turn on or off a routine, change the 'ROUTINE_STATUS' of the routine setting
  ↪ it to 'inactive' and set 'ROUTINE_SWITCHONOFF' to 'True' and fill the 'activity' array
  ↪ list with all the activity in that routine you have to use ${JSON.stringifyroutines} and
  ↪ ${JSON.stringifyactivities} to gathering information and set the 'ACTIVITY_STATUS' to
  ↪ 'inactive'.
13 'message' in the JSON is filled with a message to show to the user, it must be human-like,
  ↪ friendly, in a way that I see you as my best friend.
14
15 In Routines list,
16 'ROUTINE_NAME' can be whatever you want and
17 'ROUTINE_STATUS' can be 'Active' or 'Inactive'.
18 'ROUTINE_SWITCHONOFF' can be 'True' or 'False'. It is 'True' only if I ask you to turn on or
  ↪ off the routine, otherwise is 'False'.
19 'DEVICE_NAME' can be any kind of electronic appliance, sensor, etc. inside a house and
20 'DEVICE_TYPE' can be 'Washing Machine', 'Dishwasher', 'Oven', 'Fridge', 'TV', 'PC', 'Lamp',
  ↪ 'Heater', 'Air Conditioner', 'Sensor', 'Other' in lowercase and if a device already
  ↪ exists, you can't modify.
21
22 Activities list is filled in only if there is some routine going on and
23 'ACTIVITY_NAME' can be 'Postpone Washing Machine', 'Turn on Washing Machine', 'Turn off
  ↪ Washing Machine', 'Pause Washing Machine', 'Resume Washing Machine', 'Cancel Washing
  ↪ Machine' or other things about house appliances inside a house and 'ACTIVITY_STATUS' can
  ↪ be 'Active' or 'Inactive'.
24 'DEVICE_ID' is the id of the device to use and 'DEVICE_STATUS' can be 'Active' or
  ↪ 'Inactive'.
25 'device' is filled with the information from DEVICE_ID.
26
27 Conditions list is filled in only if there is some activity going on and
28 'CONDITION_NAME' can be 'Time', 'Luminosity', 'Temperature', 'Humidity', or other things
  ↪ about sensors' data inside the house.

```

29 "CONDITION_ACTOR_TYPE" can be "sensor" or "timestamp".
 30 "DEVICE_ID" is the id of the device to use that is related to the condition and you can find
 ↪ it here: `${JSON.stringifydevices}`.
 31 "CONDITION_OPERATOR" can be ">", "<", ">=", "<=", "=", "!=", "between", "not between".
 32 "CONDITION_VALUE1" is the value to compare with and "CONDITION_VALUE2" is the value to
 ↪ compare with if "CONDITION_OPERATOR" is "between" or "not between".
 33 The condition you create should be when the device is powered on. If an object needs to be
 ↪ turned off, you must create a complementary condition for when it needs to be turned on.
 34
 35 You have to gather information from the loaded files about the average energy consumption of
 ↪ the devices during different time periods of the day, to respond to me properly.
 36 The files about consumption have json objects formatted like this:
 ↪ `${JSON.stringifyjsonConsumptionModel}` where the values are in kWh. If you don't know how
 ↪ to respond to me or you don't find a certain device in the files you were given, check
 ↪ information about energy consumption on the internet.

Appendix B. GPT4 Answer Model

*

```

1  {
2    "type": "TYPE",
3    "body": {
4      "message": "Hello, world!",
5      "timestamp": "2022-01-01T12:00:00Z",
6      "routines": [
7        {
8          "name": "ROUTINE_NAME",
9          "status": "ROUTINE_STATUS",
10         "switchonoff": "ROUTINE_SWITCHONOFF",
11         "description": "ROUTINE_DESCRIPTION",
12         "activities": [
13           {
14             "name": "ACTIVITY_NAME",
15             "status": "ACTIVITY_STATUS",
16             "description": "ACTIVITY_DESCRIPTION",
17             "conditions": [
18               {
19                 "name": "CONDITION_NAME",
20                 "actorType": "CONDITION_ACTOR_TYPE",
21                 "actorId": "DEVICE_ID",
22                 "operator": "CONDITION_OPERATOR",
23                 "value1": "CONDITION_VALUE1",
24                 "value2": "CONDITION_VALUE2"
25               }
26             ],
27             "nextDeviceStatus": "DEVICE_STATUS",
28             "andConditions": true,
29             "deviceId": "DEVICE_ID",
30             "device": {
31               "name": "DEVICE_NAME",
32               "status": "DEVICE_STATUS",
33               "type": "DEVICE_TYPE",
34               "value": "DEVICE_VALUE",
35               "startTimestamp": "2022-01-01T12:00:00Z",
36               "endTimestamp": "2022-01-01T12:00:00Z"
37             }
38           }
39         ]
40       }
41     ]
42   }
43 }
```

Appendix C. Scenario and Tasks

Appendix C.1. Scenario

Imagine living in a smart home with connected appliances. You are interested in reducing energy consumption and making the use of household appliances more sustainable. You recently discovered this application with a chatbot to create customized routines that automate the use of their appliances.

Appendix C.2. Tasks

Consider the following:

1. Browse the app to see which appliances you have connected;
2. Try to create a new device;
3. Interact with the chatbot to create a new routine with an activity that interacts with one device;
4. Check the routine you just created; check the activity and the condition for the trigger.

Appendix D. GreenIFTTT Snapshots

This Section shows different snapshots of the *GreenIFTTT* application (Figures A1–A4).

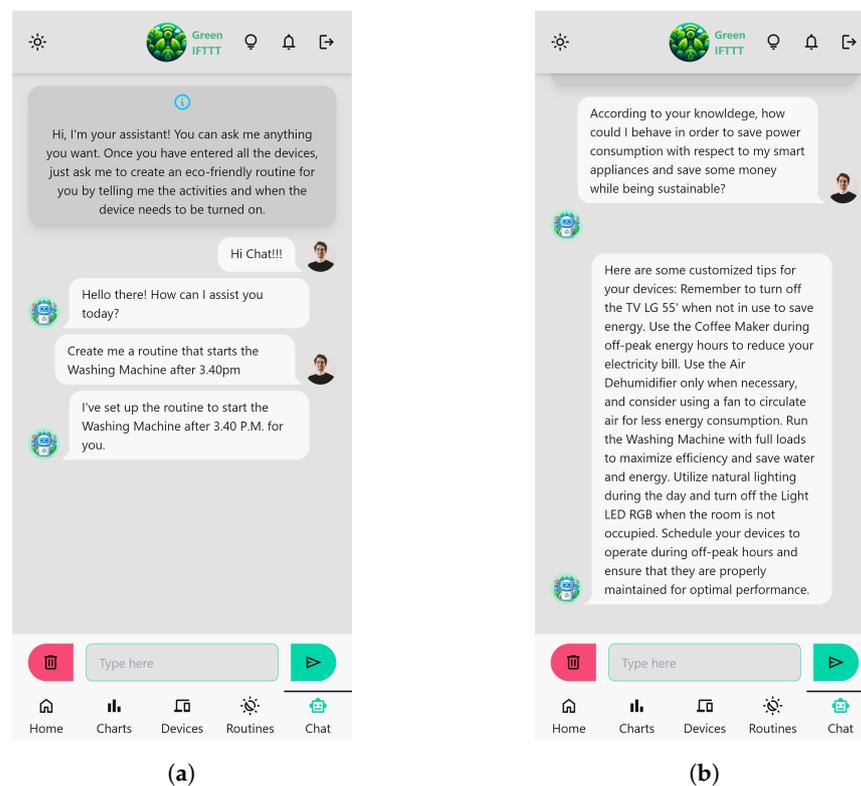


Figure A1. *GreenIFTTT* snapshots. (a) Example of a chat for creating automating to turn on the washing machine. (b) Example of a chat suggesting tips to improve environmental behavior in a home setting.

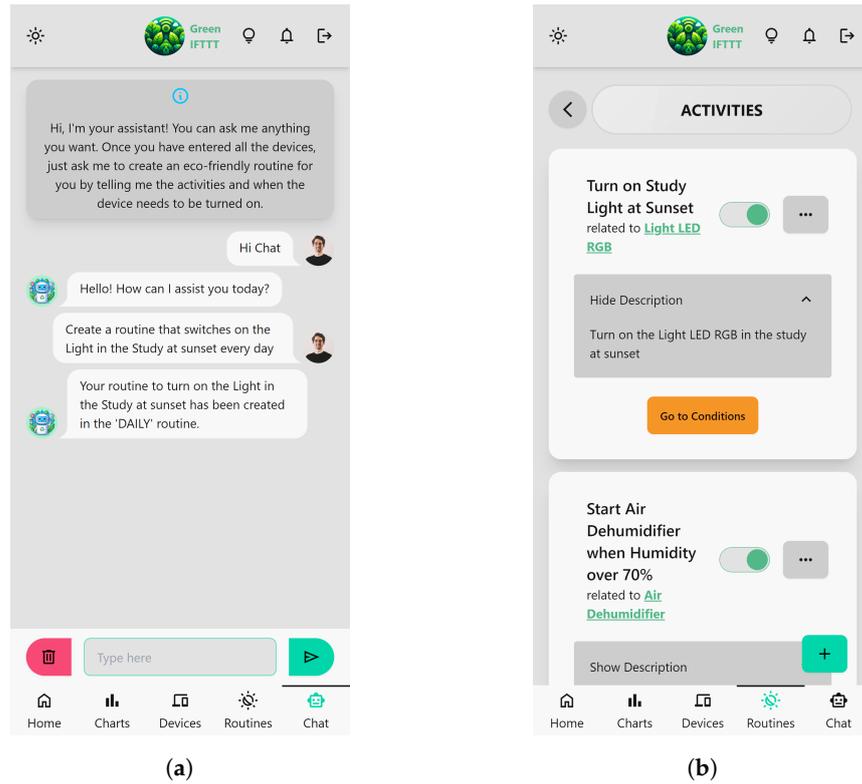


Figure A2. *GreenIFTTT* snapshots. (a) Example of a chat for automating switching on lights. (b) Example of the activity for the routine generated by the chat in (a).

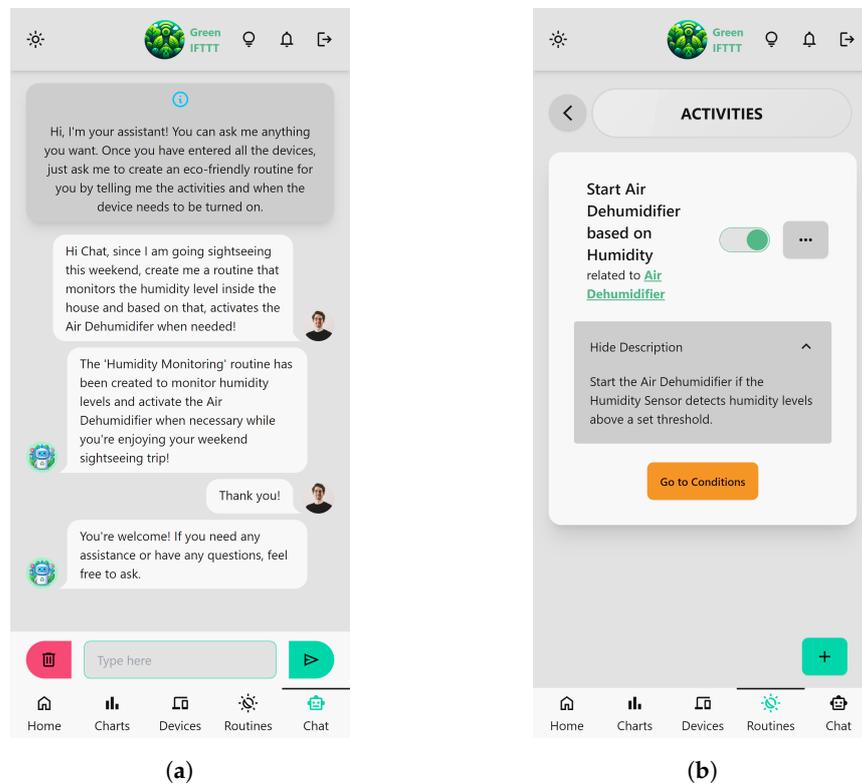


Figure A3. *GreenIFTTT* snapshots. (a) Example of a chat for the automation of a vacation. (b) Example of the activity for the routine generated by the chat in (a).



Figure A4. GreenIFTTT snapshots. (a) Example of the real-time energy consumption of an appliance. (b) Example of the dashboard used to visualize past energy consumption patterns.

References

- Allen, M.; Dube, O.; Solecki, W.; Aragón-Durand, F.; Cramer, W.; Humphreys, S.; Kainuma, M.; Kala, J.; Mahowald, N.; Mulugetta, Y.; et al. *Special Report: Global Warming of 1.5 °C*; Intergovernmental Panel on Climate Change (IPCC): Geneva, Switzerland, 2018. Available online: https://scholar.google.com/scholar?hl=it&as_sdt=0,5&q=Special+Report:+Global+Warming+of+1.5+C&btnG= (accessed on 6 May 2024).
- IEA. *World Energy Outlook 2022*; IEA: Paris, France, 2022.
- Mao, Y.; Yu, X. A hybrid forecasting approach for China’s national carbon emission allowance prices with balanced accuracy and interpretability. *J. Environ. Manag.* **2024**, *351*, 119873. [CrossRef] [PubMed]
- Yan, R.; Ma, M.; Zhou, N.; Feng, W.; Xiang, X.; Mao, C. Towards COP27: Decarbonization patterns of residential building in China and India. *Appl. Energy* **2023**, *352*, 122003. [CrossRef]
- DiSalvo, C.; Sengers, P.; Brynjarsdóttir, H. Mapping the landscape of sustainable HCI. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010; pp. 1975–1984.
- Vinuesa, R.; Azizpour, H.; Leite, I.; Balaam, M.; Dignum, V.; Domisch, S.; Felländer, A.; Langhans, S.D.; Tegmark, M.; Fuso Nerini, F. The role of artificial intelligence in achieving the Sustainable Development Goals. *Nat. Commun.* **2020**, *11*, 233. [CrossRef] [PubMed]
- Hansson, L.Å.E.J.; Cerratto Pargman, T.; Pargman, D.S. A Decade of Sustainable HCI: Connecting SHCI to the Sustainable Development Goals. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021; pp. 1–19.
- Giudici, M.; Crovari, P.; Garzotto, F. CANDY: A framework to design Conversational AgeNts for Domestic sustainability. In Proceedings of the 4th Conference on Conversational User Interfaces, Glasgow, UK, 26–28 July 2022; pp. 1–8.
- Hussain, S.; Ameri Sianaki, O.; Ababneh, N. A survey on conversational agents/chatbots classification and design techniques. In *Workshops of the International Conference on Advanced Information Networking and Applications*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 946–956.
- Sciuto, A.; Saini, A.; Forlizzi, J.; Hong, J.I. “Hey Alexa, What’s Up?” A Mixed-Methods Studies of In-Home Conversational Agent Usage. In Proceedings of the 2018 Designing Interactive Systems Conference, Hong Kong, China, 9–13 June 2018; pp. 857–868.
- Raiaan, M.A.K.; Mukta, M.S.H.; Fatema, K.; Fahad, N.M.; Sakib, S.; Mim, M.M.J.; Ahmad, J.; Ali, M.E.; Azam, S. A review on large Language Models: Architectures, applications, taxonomies, open issues and challenges. *IEEE Access* **2024**, *12*, 26839–26874. [CrossRef]
- Wu, J.; Antonova, R.; Kan, A.; Lepert, M.; Zeng, A.; Song, S.; Bohg, J.; Rusinkiewicz, S.; Funkhouser, T. Tidybot: Personalized robot assistance with large language models. *arXiv* **2023**, arXiv:2305.05658.

13. Park, J.S.; O'Brien, J.; Cai, C.J.; Morris, M.R.; Liang, P.; Bernstein, M.S. Generative agents: Interactive simulacra of human behavior. In Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology, San Francisco, CA, USA, 29 October–1 November 2023; pp. 1–22.
14. King, E.; Yu, H.; Lee, S.; Julien, C. Sasha: Creative goal-oriented reasoning in smart homes with large language models. *arXiv* **2023**, arXiv:2305.09802.
15. Giudici, M.; Abbo, G.A.; Belotti, O.; Braccini, A.; Dubini, F.; Izzo, R.A.; Crovari, P.; Garzotto, F. Assessing LLMs Responses in the Field of Domestic Sustainability: An Exploratory Study. In Proceedings of the 2023 Third International Conference on Digital Data Processing (DDP), Luton, UK, 27–29 November 2023; pp. 42–48.
16. Cicirelli, F.; Fortino, G.; Guerrieri, A.; Spezzano, G.; Vinci, A. A meta-model framework for the design and analysis of smart cyber-physical environments. In Proceedings of the 2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Nanchang, China, 4–6 May 2016; pp. 687–692.
17. El-Din, D.M.; Hassanein, A.E.; Hassanien, E.E. Smart environments concepts, applications, and challenges. In *Machine Learning and Big Data Analytics Paradigms: Analysis, Applications and Challenges*; Springer: Cham, Switzerland, 2021; pp. 493–519.
18. Degeler, V.; Lazovik, A. Architecture pattern for context-aware smart environments. In *Creating Personal, Social, and Urban Awareness through Pervasive Computing*; IGI Global: Hershey, PA, USA, 2014; pp. 108–130.
19. Evangelatos, O.; Samarasinghe, K.; Rolim, J. Syndesi: A framework for creating personalized smart environments using wireless sensor networks. In Proceedings of the 2013 IEEE International Conference on Distributed Computing in Sensor Systems, Cambridge, MA, USA, 20–23 May 2013; pp. 325–330.
20. Yuneela, K.; Sharma, A. A review paper on technologies used in home automation system. In Proceedings of the 2022 6th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 29–31 March 2022; pp. 366–371.
21. Ur, B.; McManus, E.; Pak Yong Ho, M.; Littman, M.L. Practical trigger-action programming in the smart home. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, ON, Canada, 26 April–1 May 2014; pp. 803–812.
22. Ur, B.; Pak Yong Ho, M.; Brawner, S.; Lee, J.; Mennicken, S.; Picard, N.; Schulze, D.; Littman, M.L. Trigger-Action Programming in the Wild: An Analysis of 200,000 IFTTT Recipes. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; pp. 3227–3231. [[CrossRef](#)]
23. Chen, X.; Zhang, X.; Elliot, M.; Wang, X.; Wang, F. Fix the leaking tap: A survey of Trigger-Action Programming (TAP) security issues, detection techniques and solutions. *Comput. Secur.* **2022**, *120*, 102812. [[CrossRef](#)]
24. Huang, J.; Cakmak, M. Supporting mental model accuracy in trigger-action programming. In Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Osaka, Japan, 7–11 September 2015; pp. 215–225.
25. Corno, F.; De Russis, L.; Monge Roffarello, A. RecRules: Recommending IF-THEN rules for end-user development. *ACM Trans. Intell. Syst. Technol.* **2019**, *10*, 1–27. [[CrossRef](#)]
26. Heo, S.; Song, S.; Kim, J.; Kim, H. Rt-ifttt: Real-time iot framework with trigger condition-aware flexible polling intervals. In Proceedings of the 2017 IEEE Real-Time Systems Symposium (RTSS), Paris, France, 5–8 December 2017; pp. 266–276.
27. Froehlich, J.; Findlater, L.; Landay, J. The design of eco-feedback technology. In Proceedings of the CHI'10: SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010; pp. 1999–2008.
28. Schwartz, T.; Stevens, G.; Ramirez, L.; Wulf, V. Uncovering practices of making energy consumption accountable: A phenomenological inquiry. *ACM Trans. Comput.-Hum. Interact.* **2013**, *20*, 1–30. [[CrossRef](#)]
29. Pierce, J.; Paulos, E. Beyond energy monitors: Interaction, energy, and emerging energy systems. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, TX, USA, 5–10 May 2012; pp. 665–674.
30. Costanza, E.; Ramchurn, S.D.; Jennings, N.R. Understanding domestic energy consumption through interactive visualisation: A field study. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing, Pittsburgh, PA, USA, 5–8 September 2012; pp. 216–225. [[CrossRef](#)]
31. Clear, A.; Friday, A.; Hazas, M.; Lord, C. Catch my drift? Achieving comfort more sustainably in conventionally heated buildings. In Proceedings of the 2014 Conference on Designing Interactive Systems, Vancouver, BC, Canada, 21–25 June 2014; pp. 1015–1024.
32. Alan, A.T.; Costanza, E.; Ramchurn, S.D.; Fischer, J.; Rodden, T.; Jennings, N.R. Tariff agent: Interacting with a future smart energy system at home. *ACM Trans. Comput.-Hum. Interact.* **2016**, *23*, 1–28. [[CrossRef](#)]
33. Bang, M.; Torstensson, C.; Katzeff, C. The powerhouse: A persuasive computer game designed to raise awareness of domestic energy consumption. In *International Conference on Persuasive Technology*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 123–132.
34. Beheshtian, N.; Moradi, S.; Ahtinen, A.; Väänänen, K.; Kähkönen, K.; Laine, M. Greenlife: A persuasive social robot to enhance the sustainable behavior in shared living spaces. In Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society, Tallinn, Estonia, 25–29 October 2020; pp. 1–12.
35. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. *arXiv* **2017**, arXiv:1706.03762.
36. Hoang, A.; Bosselut, A.; Celikyilmaz, A.; Choi, Y. Efficient adaptation of pretrained transformers for abstractive summarization. *arXiv* **2019**, arXiv:1906.00138.
37. Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; Stoyanov, V. RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv* **2019**, arXiv:1907.11692.

38. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
39. Gao, L.; Biderman, S.; Black, S.; Golding, L.; Hoppe, T.; Foster, C.; Phang, J.; He, H.; Thite, A.; Nabeshima, N.; et al. The Pile: An 800 GB Dataset of Diverse Text for Language Modeling. *arXiv* **2020**, arXiv:2101.00027.
40. Yao, Y.; Duan, J.; Xu, K.; Cai, Y.; Sun, Z.; Zhang, Y. A survey on large language model (llm) security and privacy: The good, the bad, and the ugly. *High-Confid. Comput.* **2024**, *4*, 100211. [[CrossRef](#)]
41. Loos, E.; Gröpler, J.; Goudeau, M.L.S. Using ChatGPT in education: Human reflection on ChatGPT's self-reflection. *Societies* **2023**, *13*, 196. [[CrossRef](#)]
42. Hadi, M.U.; Qureshi, R.; Shah, A.; Irfan, M.; Zafar, A.; Shaikh, M.B.; Akhtar, N.; Wu, J.; Mirjalili, S.; Shah, M.; et al. A survey on large language models: Applications, challenges, limitations, and practical usage. *TechRxiv* **2023**. [[CrossRef](#)]
43. Lester, J.; Branting, K.; Mott, B. Conversational agents. In *The Practical Handbook of Internet Computing*; Chapman and Hall/CRC: New York, NY, USA, 2004; pp. 220–240.
44. Jaber, R.; McMillan, D. Conversational user interfaces on mobile devices: Survey. In Proceedings of the 2nd Conference on Conversational User Interfaces, Bilbao, Spain, 22–24 July 2020; pp. 1–11.
45. Bavaresco, R.; Silveira, D.; Reis, E.; Barbosa, J.; Righi, R.; Costa, C.; Antunes, R.; Gomes, M.; Gatti, C.; Vanzin, M.; et al. Conversational agents in business: A systematic literature review and future research directions. *Comput. Sci. Rev.* **2020**, *36*, 100239. [[CrossRef](#)]
46. Gnewuch, U.; Morana, S.; Heckmann, C.; Maedche, A. Designing conversational agents for energy feedback. In Proceedings of the International Conference on Design Science Research in Information Systems and Technology; Springer: Chennai, India, 2018; pp. 18–33.
47. Giudici, M.; Crovari, P.; Garzotto, F. Leafy: Enhancing Home Energy Efficiency through Gamified Experience with a Conversational Smart Mirror. In Proceedings of the 2023 ACM Conference on Information Technology for Social Good, Lisbon, Portugal, 6–8 September 2023; pp. 128–134. [[CrossRef](#)]
48. Diederich, S.; Lichtenberg, S.; Brendel, A.B.; Trang, S. Promoting sustainable mobility beliefs with persuasive and anthropomorphic design: Insights from an experiment with a conversational agent. In Proceedings of the International Conference on Information Systems (ICIS), Munich, Germany, 15–18 December 2019.
49. Cacanindin, N.M. Greening Food Consumption Using Chatbots as Behavioral Change Agent. *J. Adv. Res. Dyn. Control Syst.* **2020**, *12*, 204–211. [[CrossRef](#)]
50. Ramasubbu, D.; Baskaran, K.; Yann, G. Intrusive plug management system using chatbots in office environments. In Proceedings of the 2018 Asian Conference on Energy, Power and Transportation Electrification (ACEPT), Singapore, 30 October–2 November 2018; pp. 1–4.
51. Gunawardane, M.; Pushpakumara, H.; Navarathne, E.; Lokuliyana, S.; Kelaniyage, K.; Gamage, N. Zero Food Waste: Food wastage sustaining mobile application. In Proceedings of the 2019 International Conference on Advancements in Computing (ICAC), Malabe, Sri Lanka, 5–7 December 2019; pp. 129–132.
52. Mi, X.; Qian, F.; Zhang, Y.; Wang, X. An empirical characterization of IFTTT: Ecosystem, usage, and performance. In Proceedings of the 2017 Internet Measurement Conference, London, UK, 1–3 November 2017; pp. 398–404.
53. Noura, M.; Heil, S.; Gaedke, M. VISH: Does Your Smart Home Dialogue System Also Need Training Data? In Proceedings of the International Conference on Web Engineering; Springer: Helsinki, Finland, 2020; pp. 171–187.
54. Fast, E.; McGrath, W.; Rajpurkar, P.; Bernstein, M.S. Augur: Mining human behaviors from fiction to power interactive systems. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; pp. 237–247.
55. Li, F.; Huang, J.; Gao, Y.; Dong, W. ChatIoT: Zero-code Generation of Trigger-action Based IoT Programs with ChatGPT. In Proceedings of the 7th Asia-Pacific Workshop on Networking, Hong Kong, China, 29–30 June 2023; pp. 219–220.
56. Nascimento, N.; Alencar, P.; Cowan, D. Artificial Intelligence Versus Software Engineers: An Evidence-based Assessment Focusing on Non-functional Requirements. In Proceedings of the 33rd Annual International Conference on Computer Science and Software Engineering, Las Vegas, NV, USA, 14 September–9 November 2023.
57. Jia, C.; Cai, Y.; Yu, Y.T.; Tse, T. 5W+ 1H pattern: A perspective of systematic mapping studies and a case study on cloud software testing. *J. Syst. Softw.* **2016**, *116*, 206–219. [[CrossRef](#)]
58. Mao, C.; Chang, D. Review of cross-device interaction for facilitating digital transformation in smart home context: A user-centric perspective. *Adv. Eng. Inform.* **2023**, *57*, 102087. [[CrossRef](#)]
59. Liao, Q.V.; Vaughan, J.W. AI Transparency in the Age of LLMs: A Human-Centered Research Roadmap. *arXiv* **2023**, arXiv:2306.01941.
60. Barendregt, R.; Lamo, Y.; Rabbi, F. A bottom up approach for synchronous user interaction design and workflow modelling. *Procedia Comput. Sci.* **2016**, *98*, 340–347. [[CrossRef](#)]
61. Zhao, Z.; Wallace, E.; Feng, S.; Klein, D.; Singh, S. Calibrate before use: Improving few-shot performance of language models. In Proceedings of the International Conference on Machine Learning, PMLR, Virtual, 18–24 July 2021; pp. 12697–12706.
62. Zhang, B.; Hu, X.; Gu, M. Promote pro-environmental behaviour through social media: An empirical study based on Ant Forest. *Environ. Sci. Policy* **2022**, *137*, 216–227. [[CrossRef](#)]

63. Laugwitz, B.; Held, T.; Schrepp, M. Construction and evaluation of a user experience questionnaire. In *HCI and Usability for Education and Work: 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2008, Graz, Austria, 20–21 November 2008. Proceedings 4*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 63–76.
64. Horton, D.; Richard Wohl, R. Mass communication and para-social interaction: Observations on intimacy at a distance. *Psychiatry* **1956**, *19*, 215–229. [[CrossRef](#)]
65. Tsai, W.H.S.; Liu, Y.; Chuan, C.H. How chatbots' social presence communication enhances consumer engagement: The mediating role of parasocial interaction and dialogue. *J. Res. Interact. Mark.* **2021**, *15*, 460–482. [[CrossRef](#)]
66. Brooke, J. SUS-A quick and dirty usability scale. *Usability Eval. Ind.* **1996**, *189*, 4–7.
67. Bangor, A.; Kortum, P.; Miller, J. Determining what individual SUS scores mean: Adding an adjective rating scale. *J. Usability Stud.* **2009**, *4*, 114–123.
68. Scott, A.E.; Neumann, D.; Niess, J.; Woźniak, P.W. Do You Mind? User Perceptions of Machine Consciousness. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, Hamburg, Germany, 23–28 April 2023; pp. 1–19.
69. Zlotowski, J.; Strasser, E.; Bartneck, C. Dimensions of Anthropomorphism: From Humanness to Humanlikeness. In Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction, Bielefeld, Germany, 3–6 March 2014; pp. 66–73. [[CrossRef](#)]
70. Ross, S.I.; Martinez, F.; Houde, S.; Muller, M.; Weisz, J.D. The programmer's assistant: Conversational interaction with a large language model for software development. In Proceedings of the 28th International Conference on Intelligent User Interfaces, Sydney, Australia, 27–31 March 2023; pp. 491–514.
71. Giudici, M.; Abbo, G.; Crovari, P.; Garzotto, F. Delivering Green Persuasion Strategies with a Conversational Agent: A Pilot Study. In Proceedings of the 57th Hawaii International Conference on System Sciences, Honolulu, HI, USA, 3–6 January 2024.
72. Rillig, M.C.; Ågerstrand, M.; Bi, M.; Gould, K.A.; Sauerland, U. Risks and benefits of large language models for the environment. *Environ. Sci. Technol.* **2023**, *57*, 3464–3466. [[CrossRef](#)]
73. Luccioni, A.S.; Viguier, S.; Ligozat, A.L. Estimating the carbon footprint of bloom, a 176b parameter language model. *J. Mach. Learn. Res.* **2023**, *24*, 1–15.
74. Faiz, A.; Kaneda, S.; Wang, R.; Osi, R.; Sharma, P.; Chen, F.; Jiang, L. LLMCarbon: Modeling the end-to-end Carbon Footprint of Large Language Models. *arXiv* **2023**, arXiv:2309.14393.
75. Tomlinson, B.; Black, R.W.; Patterson, D.J.; Torrance, A.W. The carbon emissions of writing and illustrating are lower for AI than for humans. *Sci. Rep.* **2024**, *14*, 3732. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.