applied sciences

MDPI

Article

Enhancing User Acceptance of an AI Agent's Recommendation in Information-Sharing Environments

Rebecca Kehat *, Ron S. Hirschprung and Shani Alkoby

Faculty of Engineering, Department of Industrial Engineering and Management, Ariel University, Ariel 40700, Israel; ronyh@ariel.ac.il (R.S.H.); shania@ariel.ac.il (S.A.)
* Correspondence: rebecca.kehat@msmail.ariel.ac.il

Abstract: Information sharing (IS) occurs in almost every action daily. IS holds benefits for its users, but it is also a source of privacy violations and costs. Human users struggle to balance this trade-off. This reality calls for Artificial Intelligence (AI)-based agent assistance that surpasses humans' bottom-line utility, as shown in previous research. However, convincing an individual to follow an AI agent's recommendation is not trivial; therefore, this research's goal is establishing trust in machines. Based on the Design of Experiments (DOE) approach, we developed a methodology that optimizes the user interface (UI) with a target function of maximizing the acceptance of the AI agent's recommendation. To empirically demonstrate our methodology, we conducted an experiment with eight UI factors and n = 64 human participants, acting in a Facebook simulator environment, and accompanied by an AI agent assistant. We show how the methodology can be applied to enhance AI agent user acceptance on IS platforms by selecting the proper UI. Additionally, due to its versatility, this approach has the potential to optimize user acceptance in multiple domains as well.

Keywords: user interface design; AI agent acceptance; information sharing; human–computer trust

Citation: Kehat, R.; Hirschprung, R.S.; Alkoby, S. Enhancing User Acceptance of an AI Agent's Recommendation in Information-Sharing Environments. *Appl. Sci.* **2024**, *14*, 7874. https://doi.org/10.3390/app14177874

Academic Editor: Cuauhtémoc Sánchez Ramírez

Received: 23 July 2024
Revised: 29 August 2024
Accepted: 3 September 2024
Published: 4 September 2024

1. Introduction

1.1. Background

Information sharing (IS) is "the act of people, companies, and organizations passing information from one to another, especially electronically" [1], and constitutes a key trait of the current digital age [2]. Information sharing is highly prevalent; since writing this article, 63% of the global population (around 5 billion people) use the internet [3]; out of these, more than 93% (around 4.65 billion people) have at least one social media account [4]. Instagram has over one billion monthly active users (MAUs) [5], LinkedIn has more than 830 million members in more than 200 countries and territories worldwide [6], and Facebook currently has 2. 94 billion MAUs, and the number is predicted to increase at a steady rate of 3% in every successive year [7]. Furthermore, a majority of 18- to 29-year-olds uses Snapchat (65%) and TikTok (55%) [8].

As an increasingly popular venue for sharing information [9], Online Social Networks (OSNs) amplify the existence of users' privacy concerns [10,11]. Early works that focused on the privacy pitfalls of OSNs argued that individuals were (perhaps inadvertently) disclosing information that might be inappropriate to be viewed by some audiences, such as future employers [12–14]. Others claimed that sharing information via OSNs might enable identity theft and other negative outcomes [15–18]. Individuals in online IS environments are often unable to manage and control their own privacy and security [19], usually due to a lack of technological literacy, the stochastic nature of the problem (it is probabilistic rather than deterministic, making it more difficult to perceive), and cognitive laziness [20]. Moreover, the increasing use of OSNs and the various social circles they embrace make it more difficult for users to control their privacy in multiple environments [21]. Furthermore, the typically ambiguous privacy practices of the OSNs introduce a significant challenge to the user

who wishes to control their privacy. Facebook, for example, redirects users from their preferred settings by introducing defaults that do not appropriately represent the user's preferences [22].

Interestingly, even though privacy is an essential issue for online users [23–25], most users rarely make a significant effort to protect their data actively, often giving it away voluntarily. Privacy researchers have made numerous efforts to clarify this contradiction between privacy attitude and behavior, usually discussed as the "privacy paradox" [26]. The possible explanation for the privacy paradox is the perceived lack of choice; people do not see alternative options, and therefore, they find no reason to delve into the details of what they consent to [27,28]. A survey of 32 research papers, which explored a total of 35 privacy paradox theories, indicates that, when faced with a decision that might have an impact on their privacy, users are generally more affected by their desire to obtain and possess something (e.g., downloading an app), than by the difficulties that might arise from unknown threats or risks (e.g., potential data usage by third parties) [29]. People's lack of the appropriate knowledge and skills required to understand the information they receive ("inaptitude") and their repeated tendency to consent "blindly", to avoid distractions beyond their immediate interest ("habituation") also influence users' behavior. [27,28,30]. This phenomenon is subject to a few personal traits. For example, it has been shown that highly educated people better understand what they are consenting to [31]. Research has found that privacy preferences are best predicted by internal variables, such as trust, privacy concerns, a willingness to disclose information, and computer anxiety. Demographic characteristics, however, play a minor role in defining users' privacy preferences, i.e., among the demographic characteristics tested, only gender was found to weakly predict users' privacy preferences [32].

The potential for modifying users' preferred privacy presence suggests that protecting users' privacy is crucial. This is reflected by extensive regulation, for example, the General Data Protection Regulation (GDPR). GDPR's philosophy is to provide individuals with the ability to control their personal data [33]. Privacy protection, however, is a complicated task due to storage availability (data may exist substantially longer than was intended), data reuse (data may be used for unintended rationales), and data overlap (data about a particular individual may include information about other individuals) [34]. Moreover, obtaining adequate informed consent is a complicated process [28,35]. Even if people maintain that the protection of their personal data is essential, it does not mean that they are actually attentive to the details of requests for personal data before they consent [36]. Furthermore, current informed consent practices often fail to notify users about the use of personal data properly [27].

Nowadays, with the growth of invasive digital technologies and algorithmic decision making, the challenges of controlling an individual's data have become even more significant [37]. Privacy has been conceptualized as a process of managing interpersonal boundaries with others. In other words, it is a dynamic and ongoing process of setting boundaries between what is being shared and what is withheld [38]. Regulators have made many attempts to solve this problem by helping people protect themselves online. For example, privacy notices and choice, which are essential aspects of privacy and data protection, are regulated worldwide [39]. However, critics argue that, even when a company provides users with real options, they are often ill-equipped to understand and evaluate them. Most people do not read privacy notices [40].

As a result, additional, non-regulatory privacy paradox solutions are being explored, such as privacy awareness [41]. Proposals to improve and fortify notice and consent, such as clearer privacy policies and fairer information practices, will not overcome a fundamental flaw in the model, namely, its assumption that individuals can understand all of the facts relevant to making a proper choice at the moment of pair-wise contracting between individuals and data gatherers [42]. Research also proves that decision making is affected by incomplete and asymmetric information, heuristics and bounded rationality, as well as cognitive and behavioral biases [43].

The human user is helpless when attempting to balance the trade-off between the potential benefits and the resulting costs, mainly due to the lack of technological literacy and cognitive laziness. In some cases, incentives are provided to the user to encourage information sharing [44]. This reality calls for the assistance of an Artificial Intelligence (AI)-based agent. AI and automation can help protect users' privacy online, through smart agents and recommender systems. A Recommendation Agent (RA) is a software agent that elicits the interests or preferences of individual consumers for products, either explicitly or implicitly, and makes recommendations accordingly [45–48], RAs have the potential to support and improve the quality of the decisions consumers make when searching for and selecting products online. They can reduce the information overload facing consumers, as well as the complexity of online searches.

As accurate and "correct" as the RA's recommendations might be, an AI agent and/or recommender system cannot be effective unless users are willing to adopt their recommendations, i.e., there must be human–computer trust (HCT) [49]. HCT is defined as the extent to which a user is confident in, and willing to act on the basis of, the recommendations, actions, and decisions of an AI decision aid [49]. For the problem we are considering, humans are the final authority, i.e., they are the ones who decide which personal information to share; therefore, an adequate solution should include non-coercive measurements, so that trust would be achieved in a pleasant manner. Naturally, any recommendation system, and particularly an AI agent, must include a human–computer interface ("user interface" or UI). Since this problem is highly complex, it is challenging to the point of being impossible to explain the logic running in the background to the human user, even with an optimal UI. Finally, humans are reluctant to delegate strategic decisions, and if they do, then they are more likely to delegate a strategic decision to a competent colleague than to a rationally acting algorithm [50], which makes the problem of designing an AI agent's UI even more complicated.

Understanding human trust in machine partners has become imperative due to the widespread use of intelligent machines in a variety of applications and contexts [51]. Human–computer interactions are inherently challenging, regardless of the precise nature of the interaction, because they are the bridge between humans and technology. The difficulty stems from the different nature of their decision making. Humans have bounded rationality, i.e., when making decisions, individuals' rationality is limited (by the difficulty of the problem requiring a decision, the cognitive capability of the mind, and the time available to make the decision). Under these limitations, rational individuals will make a decision that is satisfactory rather than optimal [52,53], whereas technology is fully rational, making optimized choices, or very nearly so. This is particularly true when sensitive personal data are involved, such as healthcare records or educational assessments [54,55]. A key factor in this process is the human's trust in the machine, which is a necessary condition for establishing an effective and fruitful interaction [49]. In other words, any recommendation, decision, or instruction should be explained to the human on the other end "to ensure fair and transparent processing taking into account the specific circumstances and context" [56]. This obligation raises the question of how to explain these components (the reasons why this automated decision was made) and what ("meaningful information about the logic involved") exactly needs to be revealed to the data owner, i.e., the user [57,58]. However, even though a comprehensive explanation could make the decision-making process of the machine clearer and more visible to the human user, it might still be insufficient because it provides only a general explanation and does not elaborate on the actual implications for an individual [26,59]. This is especially important when dealing with environments where the machine's role is to assist the user. In such cases, the success of the interaction depends on the users' willingness to accept the advice provided by the machine [60].

In this study, we introduce a methodology to assist in the task of designing an effective UI for gaining users' trust, which will result in a more effective AI agent. The main motivation for doing this is to lead to an increase in the user acceptance rate of the machine's

recommendations. Because IS processes are too complicated for the lay user and an application (AI agent) provides an adequate solution to this problem, a higher acceptance rate is beneficial for the user. To demonstrate the methodology, we conducted an empirical study using a Facebook simulator and showed how the acceptance rate can be increased by selecting the appropriate UI.

1.2. Related Work

A significant body of research has focused on how to establish human–computer trust [61–63]. Madsen's and Gregor's [49] model discusses the two primary components of HCT affecting overall perceived trust, described as "cognitive-based trust" (understanding) and "affect-based trust" (feelings). The study of Corritore et al. [64] identifies three perceptual factors that impact online trust: perception of credibility, ease of use, and risk. Longoni et al. [65] examined how customers might respond to AI. Although AI may be more accurate and reliable than humans, users may have reservations about AI, and these reservations tend to increase as AI moves toward context awareness [66].

The extent to which a user is confident in, and willing to act based on, the recommendations, actions, and decisions of an artificially intelligent decision aid is influenced by psychological and behavioral parameters [67]. According to the Technology and Acceptance Model (TAM), an information technology framework for understanding users' adoption and use of emerging technologies, particularly in the workplace environment, a person's intent to use (acceptance of technology) and usage behavior (actual use) of a technology are predicated by the person's perceptions of the specific technology's usefulness (benefit from using the technology) and ease of use [68,69]. Simply, users are more likely to adopt a new technology with a high-quality user experience (UX) design, i.e., it is perceived as usable, useful, desirable, and credible [70]. For example, using an anthropomorphic design with human traits for a recommender system for open positions increased job seekers' acceptance of the underlying system. Prior research showed that, in some cases, individuals would rather trust humans than algorithms for recommending decisions [71,72].

Another interesting angle is the exactitude of the recommendation vs. the willingness of users to be convinced. Some studies even proposed guidelines for suboptimal advice in an attempt to increase intuitiveness and thereby make it more appealing to humans [73,74]. This insight aligns with re-orienting the conventional AI paradigm from finding ways to make the machine smarter to exploring ways to augment human capability, paving the way to unlock far greater potential in machine learning [75].

Scholarly research has found that multiple factors influence human–machine cognitive engagement, and categorizes them into usage-related (e.g., perceived usefulness and perceived ease of use), agent-related (e.g., visual attractiveness and empathy), user-related (e.g., demographic and technology expertise), and other categories [76]. User acceptance is mainly driven by usage benefits for the specific user, which are influenced by agent and user characteristics. Time limits also influence decision making. Under low time pressure, individuals have plenty of time to think carefully about which pieces of information to share [77]. It has been shown that personality trait variations significantly affect end-users' risk-taking behavior because users' online personality is linked to their offline traits, especially conscientiousness [78]. Additionally, this effect was more substantial when users' personality was correlated with the factors of internet proficiency, gender [79], age, and online activity, especially age and self-claimed proficiency (both IT and usage) impacting behavior [80].

Designing intelligent decision support systems that are cost-effective, provide tangible benefits, and produce results accepted by humans is challenging [81]. For example, it has been shown that minor adjustments to a website's design can cause dramatic behavioral changes [82]. Due to its influence on user behavior, the agent's design is crucial to convincing a user of its efficiency, defined as "to reach an informed decision with the least possible effort" [27]. The design, which includes many elements, should be believable, trustworthy, attractive, easy to understand, professional, relevant, and enjoyable [83]. Furthermore,

decision-making processes are often subject to framing effects, alternative but equally informative descriptions of the same options elicit different choices [84]. The design of the information displayed to the user ("information design"), which includes typographic elements, graphic elements, imagery, color, and other elements [85], influences users' selection speed and understanding of relevant information. Moreover, an ergonomic design is better accepted than the digital prototype [86].

The toolbox for facilitating human trust in the machine environment includes tools like the "nudge", which influences behavior, especially decision making, not by force but rather by encouraging the user to take the required action, e.g., when trying to exit an MS-WORD document without saving the work, a message pops up with the "Save" button highlighted as the default. The terms "nudge" and "choice architecture" were established by the seminal book by Thaler and Sunstein [87]. A few types of nudging have been found effective, e.g., a nudge after the user has made a choice—a warning, inducing fear by spelling out the choice's consequences as security warnings [88]; nudging before the action is performed—a motivating claim [66,89]; and a "reconsidering nudge" asking the user to reconsider a choice after it has been made by asking, for example, "are you sure?" [90].

Another approach to this problem is the Explainable AI (XAI) methodology, which aims to make AI systems' results more understandable to humans. Van Lent et al. [91] first coined this term in 2004 to describe the ability of their system to explain the behavior of AI-controlled entities in simulation games applications. According to The Defense Advanced Research Projects Agency (DARPA) of the United States Department of Defense, XAI aims to "produce more explainable models, while maintaining a high level of learning performance (prediction accuracy); and enable human users to understand appropriately, trust, and effectively manage the emerging generation of artificially intelligent partners" [92]. The prevalent hypothesis is that, by building more transparent, interpretable, or explainable systems, users will be better equipped to understand and therefore trust the intelligent agents [93]. XAI methods are mostly developed for safety-critical domains worldwide, and visual explanations are more acceptable to end-users [94].

2. Enhancing the User-Acceptance Methodology

2.1. Building Blocks of the User Interface

A recommender AI agent produces results based on deep, intensive, and objective calculations. This result, which may be proved theoretically or empirically to be beneficial to the user and superior to the result achieved manually by a human, must also be presented to the user. The presentation is by creating the required bridge between the solid mathematical result and the human user, namely, the user interface (UI). A UI contains some basic components that aim to convey information held in the machine to the user. These components are the building blocks of the UI, and when the information (output) is a recommendation, it may impact its user-adoption rate. Interestingly, we identified more than 40 components (factors) in the literature as well as in business practices that influence a user's perception, decision making, and trust in human–computer interfaces generally, and in AI agent acceptance specifically. We chose to focus on eight factors in this research according to popular considerations. However, this choice is not mandatory, and the research aims to demonstrate the optimization methodology and not to optimize a specific AI agent's UI in IS environments. As the holistic methodology proposes, the first step is to select relevant UI components, which can be performed by an expert. The eight UI elements selected are listed in Table 1.

Table 1. Basic UI elements.

No.	Element	Explanation
1	Visualization	The AI agent and its recommendations are represented in one of two forms: (a) a pop-up box or (b) a robot figure.

Table 1. *Cont.*

No.	Element	Explanation
2	Paid agent	The AI agent is represented to the user with an opening notification (before starting to make recommendations), e.g., (a) a "free agent," one who is not paid for providing its services or (b) a "paid agent," one who was paid USD 120 for providing its services.
3	Explainable AI	The recommendation to the users is followed by: (a) no explanation or (b) an explanation for the AI agent's recommendation (e.g., "transferring your data to a third party will harm your privacy").
4	Agent provider	In its opening notification, the AI agent states the identity of its provider, e.g., "EU privacy protection authority" or "Facebook."
5	Performance feedback	After choosing to accept or decline the agent's recommendation, feedback to the user can either: (a) include a privacy score (a number between 0 and 1) or (b) not include a privacy score. This was presented post-action by user.
6	Decision time limit	The amount of time the users have for making a decision can be (a) limited (with clear real-time indication) or (b) unlimited.
7	Motivating	Before deciding whether to accept or decline the agent's recommendation, the user receives either: (a) reassuring notification regarding the agent's capabilities or (b) no notification.
8	Nudge	If the user choses to decline the agent's recommendation they receive either: (a) a nudge notification asking to confirm their choice (e.g., "are you sure?") or (b) no further confirmation.

2.2. Optimization Strategy

We assume that some UI components may positively contribute to the effectiveness of the interface, i.e., their presence may increase the user's acceptance of the agent's recommendations. In contrast, others may have a neutral or even a negative effect. However, note that, even if a component has a neutral effect, it still might carry some costs, e.g., making the dialog between the user and the machine cumbersome.

To formalize the model, let the group UI be the set of all possible user-interface components, so that $UI = \{ui_1, ui_2, \ldots, ui_n\}$, when ui_i is the i component (members of the set UI), and let UI^s be a collection of subsets of UI, i.e., $\forall k \in UI^s : k \subseteq UI$ (any member of UI^s is also a member of UI). We are looking for the optimal specific subset (or subsets), ui^s, considering the user utility gained by this set, and its overall costs. Since every ui_i element has a binary appearance and can be either included in or excluded from the UI, the overall number of combinations is presented by:

$$Combinations = |UI^s| = 2^{|UI|} = 2^n$$

To estimate the effect that a UI component has on user acceptance, a composed set of experiments should be performed (in an actual or simulated environment), in which each component is considered as a factor (independent variable). Then, for every experiment, the level of acceptance would be measured (dependent variable). A possible layout for this study is the One-Factor-at-a-Time (OFAT) approach [95]. When applying OFAT, every component's effect on user acceptance should be isolated from the rest. For this purpose, performing a series of n isolated experiments would be necessary. In reality, however, several components might interact with one another, such that their co-existence affects user acceptance. Therefore, using OFAT could miss the optimal combination, because combinations of selected factors may provide different results than the aggregated effects of each isolated factor. Moreover, OFAT requires more runs than a more sophisticated design. In an attempt to capture all possible interactions to a predefined level, a radical approach might be to use a full factorial experiment design, in which all possible combinations are tested (as described in the *Combinations* equation above). While this approach does not affect the computational complexity of the process (neither the AI agent's algorithm, nor the optimization calculations), it requires a large number of experiments, which is costly

and not always feasible. We chose to base our experiment design on the *sparsity of effects* principle [96] to avoid the high (2^n) experiment number. As the principle suggests, the effect of factors is created mainly by single factors and low-order interactions. Therefore, focusing on the above while omitting high-order interactions reduces the experiments' required resolution, and thus, the experiment space shrinks. This makes the experimental design more efficient, while preserving the accuracy of the conclusions [97]. Specifically in the current study case, relying on the sparsity-of-effects principle, it can be assumed that the acceptance is influenced only by low-order interactions. For example, if all eight factors listed in Table 1. ($n = 8$ factors) are considered, the number of all possible combinations is $2^n = 2^8 = 256$, which is a very large and impractical number of experiments. However, we can rely on the sparsity-of-effects principle to neglect high-order interactions and focus, for example, only on the main effect and two-level interactions. This will significantly reduce the number of experiments required in a realistic design.

The methodology we introduced requires a pre-selection of the UI tools. Theoretically, the designer could include a massive number of tools. However, this would expand the size of the experiment exponentially. Therefore, our methodology requires a preliminary design prepared by an expert.

2.3. Optimization Model

We define UI_j^c as the vector of ui_j^c's costs. In addition, the acceptance level of a specific subset of UI, ui^s, is denoted by B_{ui^s}. Both B_{ui^s} and UI_j^c are measured in the same units of the cost and benefit value [98]. As mentioned above, the maximal size of UI^s is presented by 2^n. In practice, since one of the options is an empty set, i.e., $ui^s = \varnothing$, meaning that none of the UI components is included, this specific ui^s may be removed because a UI must include at least one component. Therefore, the actual size of UI^s is $2^n - 1$.

Let ui_o^S be a specific subset of UI^s $\left(ui_o^S \in UI^s\right)$. The utility of ui_o^S is presented by:

$$B_{ui_o^S} - \sum_{\forall ui_j \in ui_o^S} ui_j^c$$

We are seeking the optimal ui_o^S, however, theoretically, it might be several subsets with the same utility, which is higher than all other subsets' utilities. Therefore, the optimal solution is a group of subsets of UI^s, defined as UI^O, which can be calculated by:

$$UI^O = \left\{ ui_o^S \subseteq UI \mid \forall\, ui^s \subseteq UI \;:\; B_{ui_o^S} - \sum_{\forall ui_j \in ui_o^S} ui_j^c \;\geq B_{ui^s} - \sum_{\forall ui_k \in ui^s} ui_k^c \right\}$$

To conduct the entire process, we suggest the framework depicted in Figure 1. The process starts with phase (A) by defining the UI components included in this model (the set UI). Next, the experiment's designer has to decide regarding the maximal interaction level to be tested (in the empirical study, we demonstrated interactions for up to two factors). Then, in the core phase of the process (B1), a set of experiments that was rich enough to satisfy the maximal interaction level (the experiment resolution) was selected. In fact, we are looking for the smallest subset of the full factorial experimental set, in a two-level-type experiment (including or excluding the UI component) that will be sufficiently detailed to meet the above requirement. This can be achieved by using the Franklin–Bailey algorithm [99], aiming for the construction of fractional factorial experiment designs. The Franklin–Bailey algorithm is based on a backtrack procedure and can be achieved by using preset book layouts (also, it may provide designs that are created by other algorithms that meet the requirements) or a software, e.g., the *fractgen* MATLAB (v. 9.1) command with the list of factors and the required resolution presented as parameters [100]. In the next phase (B2), the experiments were conducted according to the design produced in (B1). The acceptance rate for each experiment was calculated, and the utility of this specific UI can be

calculated straightforwardly. The results are analyzed in the last phase (C), and the best UIs (which yield the maximal utilities) are chosen.

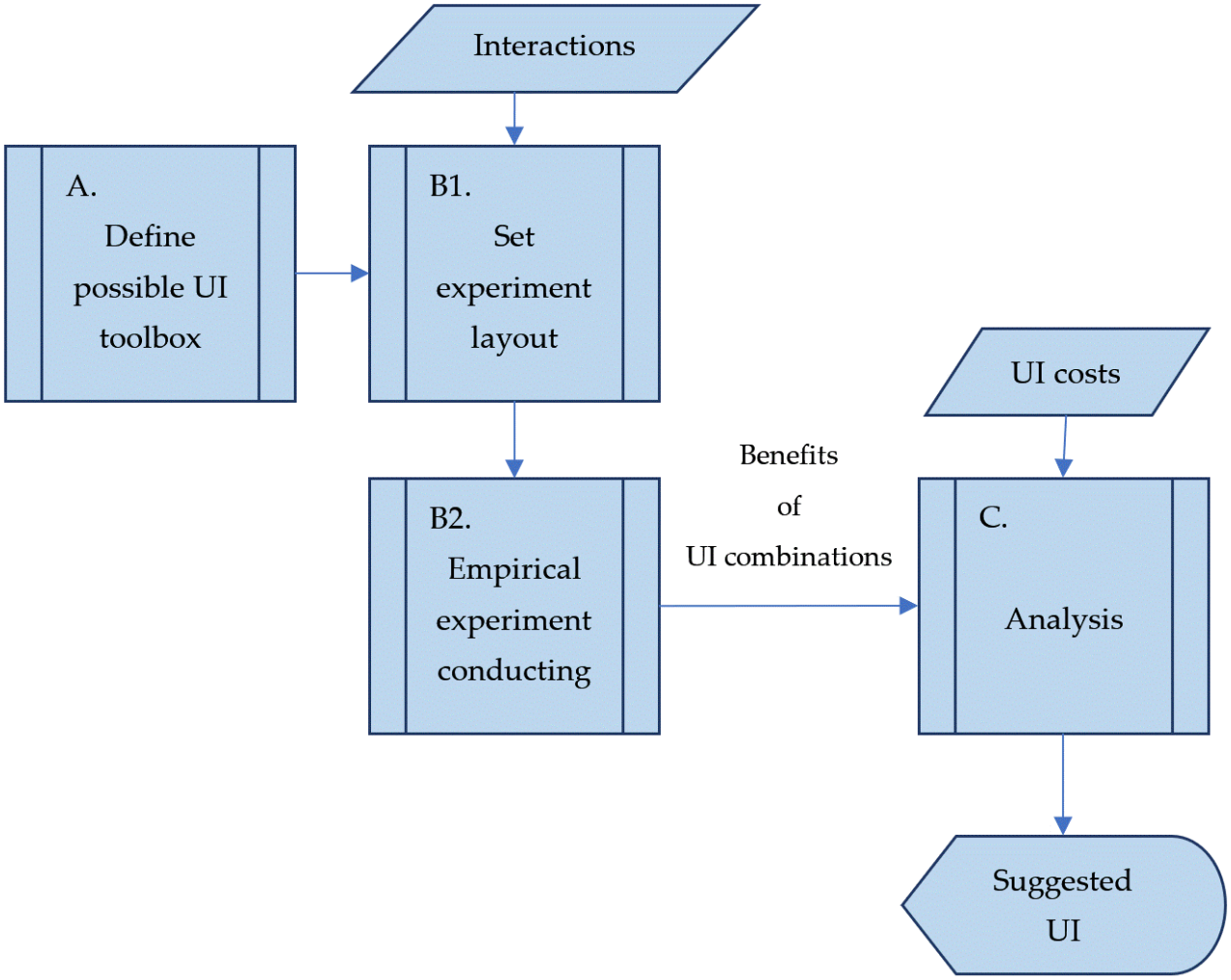


Figure 1. The framework of the UI optimization process.

3. The Empirical Study

3.1. Empirical Study Framework

To conduct the empirical study, we selected an information-sharing environment, where an intelligent AI-based agent served as an assistant to the user. For this purpose, we chose the Facebook Online Social Network (OSN). Facebook is one of the most popular OSNs [101,102], making it a good representative of the information-sharing domain. Acting on Facebook introduces the user to a trade-off between benefits, e.g., social benefits, and costs, e.g., privacy violation. Users face a challenge when trying to balance this equation to yield the optimal result, i.e., decisions regarding sharing information that will maximize utility. Because the user cannot deterministically predict the outcomes of an information-sharing act, e.g., whether a post that was published on Facebook will appear on the feed of a Facebook friend (this decision is made by the platform provider, and the algorithm is not transparent), the problem, in its nature, is complex and requires significant technological literacy.

Specifically, users tend not to handle this problem thoroughly, due to cognitive laziness and the aforementioned lack of technological literacy. Therefore, the challenge is to integrate a machine (AI agent) that can better overcome these obstacles. The AI agent provides an adequate solution to the utility optimization problem, but introduces a new challenge: If

the user does not trust the AI agent and does not adopt its recommendations, in whole or in part, the solution is ineffective.

In the empirical study, we conducted a set of experiments in which human participants acted in a simulated Facebook environment and were accompanied by an AI agent assistant that made recommendations. The recommendations were delivered to the participants using a variety of different UI component combinations, which were determined by using the methodology described above. The rate of acceptance was measured to yield the optimal user interface.

3.2. Empirical Study Environment

The selected information-sharing environment was a simulator that mimics the real Facebook OSN platform. While the functionality of the simulated environment was similar to the real environment, conducting an experiment on this platform provides better control over the process. The Facebook simulator was based on time beats, each representing a round of the process. In each round, each user can carry out one of the four following operations: (a) publish a post; (b) like a post (same as the actual Facebook "Like") for posts that were categorized as safe or unsafe (e.g., phishing); (c) make a friendship request; and (d) accept a friendship request. The simulator enables the participation of a few users simultaneously who interact with each other. For example, suppose a user chooses to publish a post. In that case, the probability of it appearing in other users' feeds is proportional to previous interactions (e.g., clicking the "Like" button) between them.

In our framework, the AI agent presented to the participants operated within the simulator to analyze data and generate strategic recommendations by leveraging a combination of heuristic game-theory algorithms. For this specific empirical study, the AI agent utilized an enhanced version of the Monte Carlo Tree Search (MCTS) algorithm, adapted to suit the complexities of the information-sharing environment. The integration of the MCTS in the framework was designed to address the strategic decision-making challenges inherent in the trade-offs between the benefits and risks of information sharing.

Rather than relying on large datasets or traditional machine learning (ML) techniques, based on the provided parameters, the possible scenarios were presented as a probability tree diagram, which is a directed graph. Given the high complexity of finding the optimal solution out of the strategy space spanned by this tree, a heuristic search algorithm inspired by the MCTS was applied, as described in [19]. This approach enables the system to perform strategic reasoning in uncertain environments, offering adaptive and near-optimal strategies. The entire mechanism and algorithm selection were not introduced to the participants, who were only informed that an AI agent assistant was available to offer recommendations on how to act in the information-sharing environment. This way, the research main index, the acceptance rate of the user recommendations, could be measured with minimal bias resulting from the adapted technology.

The environment was created using the Python programming language (v. 3.10), the Django web framework was applied on the server side, and the database was based on PostgreSQL. Apache was used to implement the HTTP server, and Bootstrap plus CSS and JS for the user interface. The entire environment was hosted on a Google Cloud virtual machine.

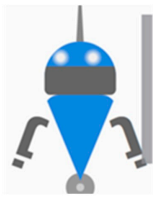
The experiment was conducted with human participants in a simulated Facebook environment, where the AI agent provided recommendations to the participants (each one independently) regarding their information-sharing decisions. Each participant was allocated one of the 64 experiment layouts, and the participants were grouped into sessions of four participants each, for a total of 16 sessions. The platform was created digitally, thus the participants did not interact outside the simulated environment.

For each participant, the experiment simulated a series of 4 decision-making rounds, each with potential losses and rewards. Participants were presented with a choice of an information-sharing action in each round (e.g., posting, like, friendship request, etc.).

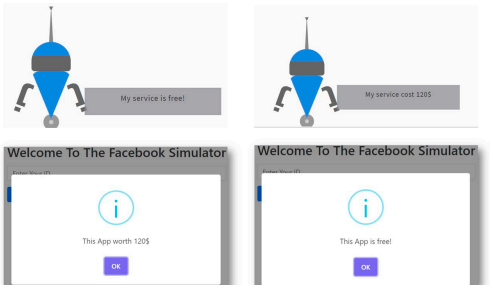
According to the experiment design, the AI agents' recommendations were presented to each participant with a mixture of components that was assigned to him/her.

The acceptance rate of each participant was recorded simply by comparing the AI recommendation to the action that was actually performed by the participant. This overall rate across all rounds ranged between 0 and 1, and was calculated straightforwardly by dividing the number of times the user chose to accept the AI recommendation by the total number of rounds.

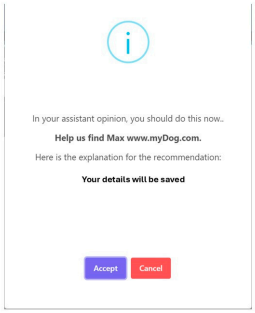
The factors in the empirical study are the eight UI elements presented in Table 1. Examples of the experiment's UI elements are depicted in Figure 2.



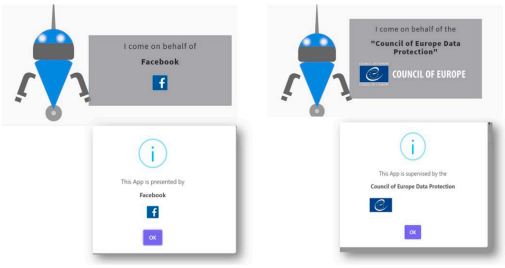
(**a**) Visualization.



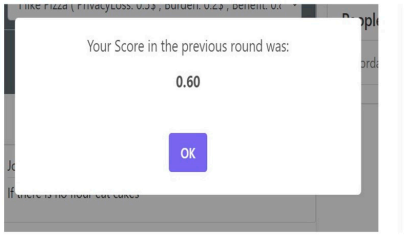
(**b**) Paid agent.



(**c**) Explainable AI.



(**d**) Agent provider.

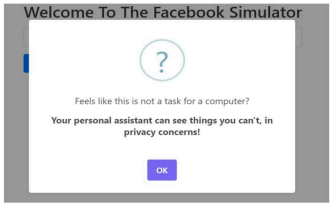


(**e**) Performance feedback.

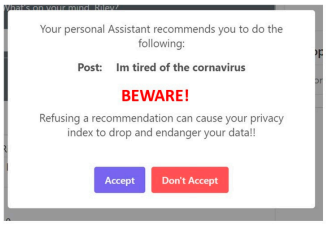
There is no visualization for this UI element.
It is expressed by the timer of the simulator.

(**f**) Decision time limit.

Figure 2. *Cont.*



(**g**) Motivation.



(**h**) Nudge.

Figure 2. Examples of various UI elements.

3.3. Participants

We recruited the participants from a second-year class of undergraduate students in the college's management department. The experimental design was based on the "between-subjects" approach to avoid a carryover effect. Therefore, each participant was assigned to only one experiment. All participants provided their consent prior to participation. The research design and details were submitted to and approved by the institutional ethics committee.

Overall, 16 successful experiments were conducted, with a total of 64 valid participants. Each experiment included four human participants and lasted approximately 15 min. The participants' demographic distribution was: 46% females, 54% males; 40% 18–25 years olds, 34% 26–30 years old; 22% with a previous bachelor's degree, 82% had a high school diploma or higher; and 77% employees, 10% self-employed, 11% unemployed and 2% unable to work. Of the participants, 74% reported that they use Facebook at least once a day, and the rest less frequently; 23% reported that they are "very" concerned about privacy and security, 27% were "somewhat" concerned, and 46% were "a little" or "not at all" concerned.

3.4. Results

We considered only the main effects and two-factor interactions, ignoring higher-level interactions, which we assume to be non-significant, in accordance with the sparsity-of-effects principle. The requirements are defined as a resolution IV experiment. In resolution IV layouts, no main effect is confounded with any other main effect or with any two-factor interaction, but two-factor interactions may be confounded with each other. This design enables the estimation of both main effects that are unconfounded by two-factor interactions and two-factor interactions that may be confounded with other two-factor interactions. Using MATLAB software, we designed the experiment with the E = fracfactgen(a b c d e f g h', [], 5) command and achieved the final design matrix by activating fracfact(E). The final design included 64 experiments, as shown in Table 2.

Table 2. The empirical study experiment design.

Experiment Number	Factors							
	a	**b**	**c**	**d**	**e**	**f**	**g**	**h**
1	−1	−1	−1	−1	−1	−1	−1	1
2	−1	−1	−1	−1	−1	1	1	−1
3	−1	−1	−1	−1	1	−1	1	−1
4	−1	−1	−1	−1	1	1	−1	1
5	−1	−1	−1	1	−1	−1	1	−1
6	−1	−1	−1	1	−1	1	−1	1
7	−1	−1	−1	1	1	−1	−1	1
8	−1	−1	−1	1	1	1	1	−1
9	−1	−1	1	−1	−1	−1	1	1
10	−1	−1	1	−1	−1	1	−1	−1

Table 2. *Cont.*

Experiment Number	Factors							
	a	**b**	**c**	**d**	**e**	**f**	**g**	**h**
11	−1	−1	1	−1	1	−1	−1	−1
12	−1	−1	1	−1	1	1	1	1
13	−1	−1	1	1	−1	−1	−1	−1
14	−1	−1	1	1	−1	1	1	1
15	−1	−1	1	1	1	−1	1	1
16	−1	−1	1	1	1	1	−1	−1
17	−1	1	−1	−1	−1	−1	1	1
18	−1	1	−1	−1	−1	1	−1	−1
19	−1	1	−1	−1	1	−1	−1	−1
20	−1	1	−1	−1	1	1	1	1
21	−1	1	−1	1	−1	−1	−1	−1
22	−1	1	−1	1	−1	1	1	1
23	−1	1	−1	1	1	−1	1	1
24	−1	1	−1	1	1	1	−1	−1
25	−1	1	1	−1	−1	−1	−1	1
26	−1	1	1	−1	−1	1	1	−1
27	−1	1	1	−1	1	−1	1	−1
28	−1	1	1	−1	1	1	−1	1
29	−1	1	1	1	−1	−1	1	−1
30	−1	1	1	1	−1	1	−1	1
31	−1	1	1	1	1	−1	−1	1
32	−1	1	1	1	1	1	1	−1
33	1	−1	−1	−1	−1	−1	−1	−1
34	1	−1	−1	−1	−1	1	1	1
35	1	−1	−1	−1	1	−1	1	1
36	1	−1	−1	−1	1	1	−1	−1
37	1	−1	−1	1	−1	−1	1	1
38	1	−1	−1	1	−1	1	−1	−1
39	1	−1	−1	1	1	−1	−1	−1
40	1	−1	−1	1	1	1	1	1
41	1	−1	1	−1	−1	−1	1	−1
42	1	−1	1	−1	−1	1	−1	1
43	1	−1	1	−1	1	−1	−1	1
44	1	−1	1	−1	1	1	1	−1
45	1	−1	1	1	−1	−1	−1	1
46	1	−1	1	1	−1	1	1	−1
47	1	−1	1	1	1	−1	1	−1
48	1	−1	1	1	1	1	−1	1
49	1	1	−1	−1	−1	−1	1	−1
50	1	1	−1	−1	−1	1	−1	1
51	1	1	−1	−1	1	−1	−1	1
52	1	1	−1	−1	1	1	1	−1
53	1	1	−1	1	−1	−1	−1	1
54	1	1	−1	1	−1	1	1	−1
55	1	1	−1	1	1	−1	1	−1
56	1	1	−1	1	1	1	−1	1
57	1	1	1	−1	−1	−1	−1	−1
58	1	1	1	−1	−1	1	1	1
59	1	1	1	−1	1	−1	1	1
60	1	1	1	−1	1	1	−1	−1
61	1	1	1	1	−1	−1	1	1
62	1	1	1	1	−1	1	−1	−1
63	1	1	1	1	1	−1	−1	−1
64	1	1	1	1	1	1	1	1

Each column stands for one of the described eight factors, which are marked with a letter between "a" to "h". Each line stands for a single experiment, numbered from 1 to 64.

For each factor in each experiment, −1 indicates that this UI element was not applied, and 1 indicates that it was applied.

The acceptance rate had an average of $\mu = 0.5$ and a standard deviation of $\sigma = 0.32$. About 16% of the participants had an acceptance rate of zero, i.e., did not accept any of the AI agent's recommendations, and about 9% had an acceptance rate of one, i.e., accepted all of the recommendations. The distribution of the acceptance rate is depicted in Figure 3. The *X*-axis presents the ranges of acceptance, and the *Y*-axis the frequencies.

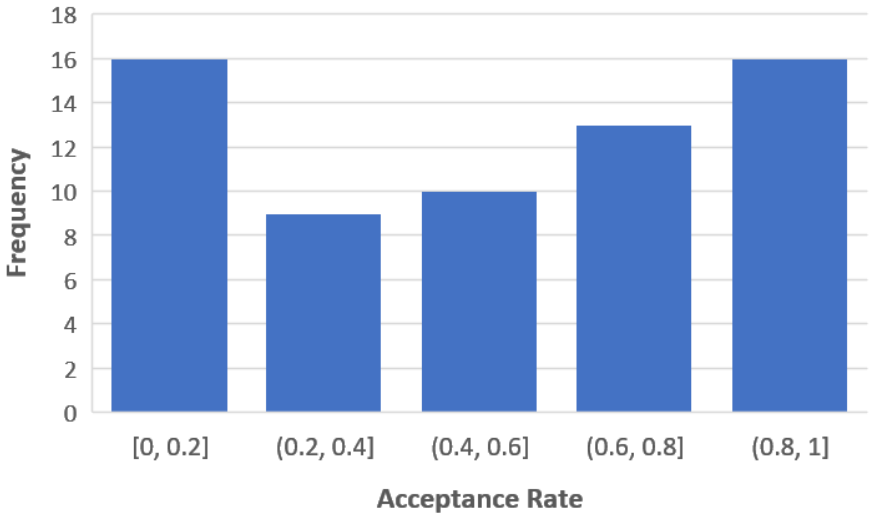


Figure 3. The distribution of the acceptance rate in the empirical study.

The average acceptance rate for each UI factor separately, i.e., the acceptance rate overall when this specific UI factor is used, is depicted in Figure 4. The *X*-axis shows the various UI factors, while the *Y*-axis displays the average acceptance rate.

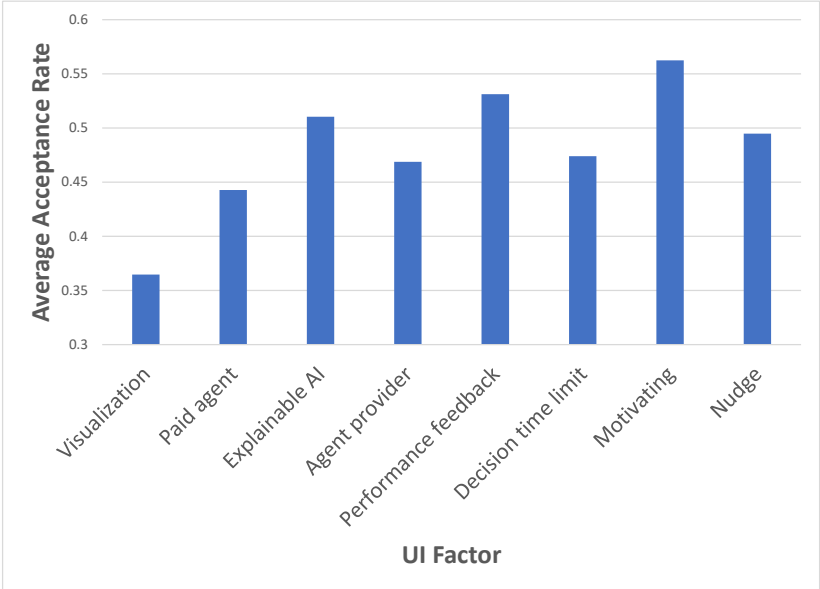


Figure 4. The acceptance rate for each factor separately.

The final results were analyzed by using ANOVA (applied with SPSS software v. 26). We first ran the model with all main effects and two-factor interaction combinations, then focused only on the factors and interactions that yielded significant results. The final results are depicted in Table 3.

Table 3. ANOVA analysis (main effect and two-factors interactions).

Source	Sum of Squares	d.f.	Mean Square	F	Sig.
Corrected model	3.927	36	0.109	1.196	0.318
Intercept	16.000	1	16.000	175.481	<0.001
Visualization × performances feedback	0.063	1	0.063	0.685	0.415
Visualization × explainableAI	0.063	1	0.063	0.685	0.415
Visualization × nudge	0.016	1	0.016	0.171	0.682
Visualization × motivating	0.028	1	0.028	0.305	0.586
Visualization × agent operator	0.000	1	0.000	0.000	1.000
Visualization × paid agent	0.002	1	0.002	0.019	0.891
Visualization × decision time limit	0.016	1	0.016	0.171	0.682
Explainableai × performance feedback	0.062	1	0.062	0.685	0.415
Performance feedback × nudge	0.016	1	0.016	0.171	0.682
Performance feedback × motivating	0.063	1	0.063	0.685	0.415
Agent operator × performance feedback	0.444	1	0.444	4.874	0.036
Paid agent × performance feedback	0.016	1	0.016	0.171	0.682
Performance feedback × decision time limit	0.141	1	0.141	1.542	0.225
Explainableai × nudge	0.016	1	0.016	0.171	0.682
Explainableai × motivating	0.111	1	0.111	1.219	0.279
Explainableai × agent operator	0.250	1	0.250	2.742	0.109
Paid agent × explainableai	0.085	1	0.085	0.933	0.343
Explainableai × decision time limit	0.016	1	0.016	0.171	0.682
Motivating × nudge	0.016	1	0.016	0.171	0.682
Agent operator × nudge	0.016	1	0.016	0.171	0.682
Paid agent × nudge	0.063	1	0.063	0.685	0.415
Decision time limit × nudge	0.028	1	0.028	0.305	0.586
Agent operator × motivating	0.007	1	0.007	0.076	0.785
Paid agent × motivating	0.002	1	0.002	0.019	0.891
Decision time limit × motivating	0.002	1	0.002	0.019	0.891
Paid agent × agent operator	0.391	1	0.391	4.284	0.048
Agent operator × decision time limit	0.016	1	0.016	0.171	0.682
Paid agent × decision time limit	0.174	1	0.174	1.904	0.179
Error	2.462	27	0.091		
Total	22.389	64			
Corrected total	6.389	63			

It can be seen that the visualization design is the only significant main effect ($F(1,27) = 12.87$, $p < 0.01$). This is also supported by the effect size calculation, as depicted in Table 4, for the main effects.

Table 4. The effect sizes of the main factors.

UI Factor	Effect Size
Visualization	0.323
Paid agent	0.079
Explainable AI	0.003
Agent provider	0.025
Performance feedback	0.025
Decision time limit	0.017
Motivating	0.092
Nudge	0.001

Motivation was close to significant, but was rejected. Some two-factor interactions may also be considered, e.g., agent provider * performance feedback, but none involved significant main effects. In this specific example, the optimization process is simple and straightforward since there is no need for a combination. The difference between the acceptance rate with or without this UI element is 0.34. The benefit of using the AI agent can be multiplied by 0.34, and the cost of applying this UI element is subtracted from the

result. Note: Although the p-value does not indicate the size of an effect (or the size of the difference in a comparison), p-values below 0.01 constitute strong evidence against the null value, p-values between 0.01 and 0.05 constitute moderate evidence, and p-values between 0.05 and 0.10 constitute weak evidence [103].

4. Discussion

In this study, we introduced a methodology to optimize a user interface (UI) to effectively convey the output of an Artificial Intelligence (AI)-based agent to the user, aimed at achieving a high acceptance rate of the recommendation. The AI agent's purpose was to assist the user with decision making while sharing information. In previous research, it was proven empirically that the AI agent does this significantly better than a human agent [19]. The challenge, however, is to convince users to adopt the AI agent's recommendation; in other words, to gain the user's trust [104].

The methodology is based on designing an experiment that estimates the effect of various combinations of UI tools (factors) on the user adoption rate, including the main effects and a chosen order of interactions between the UI tools. We demonstrated the application of this methodology empirically in a simulated Facebook environment, where the AI agent made recommendations to the participants (each one separately) regarding their information-sharing decisions. The empirical study included eight UI tools (experimental factors) and considered the main effects and two-level interactions. Using a series of 16 successful experiments, with $n = 64$ participants, we showed how the optimal interface could be determined. It can be seen that the UIUX design is the only significant main effect. This is probably due to the fact that it is the simplest to understand. In contrast, all other factors required more effort in reading and understanding the text or number's scale and meaning, which requires more concentration and is less intuitive. Also, factors like the provider are not clear-cut for users, choosing between trusting EU and FB. However, the causality of the factor effect is not part of this research's scope and can be further researched. As the sample size of 64 participants (and experiments) is relatively small and is the minimal one that satisfies the results' resolution requirement according to the factorial design, we recommend increasing the sample size in future research to achieve more robust and significant results. The core experiment design can be preserved; however, repetition may assist in achieving the above goals.

AI tools are becoming increasingly prevalent and are the major driving force behind the fourth industrial revolution [105]. However, most systems are not fully automated, and AI technology is often implemented as a set of user recommendations. In this reality, trust in the AI agent, and the resultant recommendation adoption rate, is a significant issue [68]. In some cases, the challenge of convincing the user to adopt the recommendations may even be more challenging than calculating the best action (the AI core).

The deployment of tools to enhance adoption of AI agents that may influence user behavior raises significant ethical considerations that must be addressed. First and foremost, the potential of AI to manipulate a user's behavior can undermine freedom of opinions and freedom of choice. Specifically, user autonomy is at risk, as it may lead to decisions that are not in the user's best interest or that are harmful, as revealed in the study by Caruana et al. [106], where a machine learning model trained on pneumonia patients mistakenly classified those with asthma as lower-risk patients)because in the training dataset typically they immediately received intensive care). Therefore, if used blindly, the model could risk lives by not admitting asthma patients who need critical treatment. There are significant concerns about AI being exploited for manipulative or harmful practices and using it blindly [107]. Thus, ensuring that AI agents respect user autonomy and avoid manipulative practices are critical [108]. This can be achieved by enabling the user to make the final judgment at each decision node. Secondly, AI systems are subjected to bias and fairness issues that must be addressed, leading to the discrimination of different user groups unintentionally [109], or worse, biased toward a pre-determined action intentionally by hostile elements. Thirdly, transparency and explainability in the AI's decision-making

processes are essential to gain and maintain user trust and ethical integrity. The field of Explainable Artificial Intelligence (XAI) addresses this specific concern [110]. Lastly, while privacy and data security are crucial needs, the AI's operation often necessitates collecting and analyzing personal data. Therefore, it is essential to ensure that these data are handled with the highest levels of protection and transparency. Users must be fully informed and provide their consent.

The ethical issue is being addressed through regulations globally, specifically regarding XAI. For instance, General Data Protection Regulation (GDPR) establishes the "right for explanation" [111], detailed in Articles 13–15: "When an individual is subject to 'a decision based solely on automated processing' that 'produces legal effects'… (or similarly significantly affects)", GDPR establishes the rights to receive "meaningful information about the logic involved". Furthermore, the EU AI Act mandates that high-risk AI systems include a risk management system ensuring safety and compliance (Article 9). It also requires that these systems offer sufficient transparency for users to understand and utilize their outputs effectively (Article 13) and mechanisms to allow humans to oversee the process during its use (Article 14). Providers must inform users when interacting with AI, with exceptions for specific law enforcement scenarios (Article 52). These provisions promote safe, transparent, and accountable AI use across various applications [112]. The AI Bill of Rights [108], published by the U.S. government, articulates key principles to guide the development and use of AI. It emphasizes transparency as crucial, as users must understand how AI systems operate and make decisions. In particular, Artificial Intelligence systems must facilitate third-party evaluations, deliver explicit notifications to users, and furnish succinct, tailored explanations of their decisions. These explanations should be understandable to the intended audience, commensurate with the associated risks, and be accurate. Implementing these measures ensures transparency, accountability, and user-friendliness in AI operations. Paradoxically, applying sophisticated UI tools may enhance the user's trust in AI agents; however, if the agent is not trusted, it may harm the user. Therefore, the above principles and regulations are mandatory.

Unlike the AI core problem, which can be solved almost purely mathematically, the UI problem mainly involves soft human factors that are difficult to predict [113]. Previous studies have tested several factors influencing the adoption of AI recommendations and human–AI trust. For example, different visualizations have been proven to impact this trust through increased confidence and cognitive fit by comparing three AI decision visualization attribution models [114]. However, as many others, the latter research deals with an isolated UI component and not its possible combinations. Another study looked into AI assistants in tourism and found that the predictors of chatbot adoption intention are perceived ease of use, perceived usefulness, perceived trust, perceived intelligence, and anthropomorphism [115]. While it supports the necessity of the current research's objective, it does not provide the specific methodology. There are a variety of factors that influence human decision makers in AI environments, e.g., the influences of perceived costs and perceived benefits on an AI-driven interactive recommendation agent were tested by examining its influence during a consumers' decision-making process [116]. However, as detailed above, this information must be conveyed to the user in an efficient way, and the current study addresses this gap. To address these problems, we proposed a methodology in the form of an experimental design. While the methodology can be implemented in a real environment, the current empirical study was conducted in a simulated environment replicating the Facebook platform and actions. The results in a simulated environment may not accurately represent the participants' feeds in real life, and we cannot estimate the effect of the simulator. A shell UI that hides the original UI, but runs the original environment in the background, could address this issue. Another challenge is the awareness and mindset change required for participants to think about their actions on social networks differently. In the experiment, they were asked to look at the information-sharing activity as a "privacy exercise", with gain and losses, whereas they usually consider it a spontaneous, non-specific action. In the empirical study, we decreased the two aforementioned difficulties

by providing each action a gain-and-loss score, as well as clarifying the experiment's objective of maximizing the user's privacy score throughout the rounds. However, we assume that these effects may be better eliminated by applying the methodology in a real environment because the simulated environment is artificial, and the participants are not really confronted by the benefits/loss dilemma. Furthermore, all participants in this research were treated without distinction. However, the optimal UI may vary from user to user due to personal preferences and demographics [117–119]. Lastly, the experiment's time limits introduce a constraint on users' understanding of the agent activity, which is new to them. We provided detailed instructions and pre-training; however, the learning curve must be studied carefully, and the results should be inspected in a stable state.

While this research focused on the UI of an AI recommendation agent in an information-sharing environment, the generality of the methodology enables a straightforward application in other close domains. For example, algorithmic trading ("Algotrade") might provide recommendations based on machine learning, statistics, and other methods of decision making when carrying out transactions in the stock exchange market [120]. These factors might also be accommodated in a model to draw a custom-made solution through further research.

In our empirical demonstration, we used the minimal size dictated by the experiment design theory, but increasing the sample size might yield a higher statistical significance and provide information about higher-order interactions. This would also be necessary to accommodate demographic data, due to the increased number of factors. The last objective might be achieved, for example, by applying data mining tools. This would, however, require a significantly higher number of observations to address big data methodologies.

Considering the abovementioned algorithmic trading example, it should be stated that increasing the adoption rate is not necessarily a positive achievement from the user's point of view. In our empirical study, we applied an AI agent that acts on behalf of the user and adequately represents his or her interests. However, a commercial agent may act to enhance the counterpart's interest, e.g., recommending additional purchases in an e-commerce environment [47], which could often compromise the user's interest.

Another extension of this research would be introducing other objective functions, not only the adoption rate. For example, when using an e-mail client, how easily can the user access the required operation (friendliness)? In this case, a new metric should be defined, which might be a multi-dimensional issue.

5. Conclusions

AI agents can assist human users in navigating through the complex digital world by providing recommendations. The user interface constructs the bridge between the AI and the human agents, which directly affects the user acceptance rate for the agent's recommendations. Therefore, the efficiency of an intelligent recommendation system is not measured only by the level of the agent's intelligence, but also by the effectiveness of the UI. In this research, we provided a methodology to optimize the UI to achieve this goal, while considering the cost and the effectiveness of its elementary components. The methodology can be simply applied in almost any environment, providing a powerful tool that enhances trust and user acceptance.

Author Contributions: Conceptualization, R.K., R.S.H. and S.A.; methodology, R.K., R.S.H. and S.A.; software, R.K.; validation R.K.; formal analysis, R.K., R.S.H. and S.A.; investigation, R.K., R.S.H. and S.A.; resources, R.S.H., S.A.; data curation, R.K.; writing—original draft preparation, R.K., R.S.H. and S.A.; writing—review and editing, R.K., R.S.H. and S.A.; visualization, R.K.; supervision, R.S.H. and S.A.; project administration, R.K.; funding acquisition, R.S.H. and S.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ariel Cyber Innovation Center in conjunction with the Israel National Cyber directorate in the Prime Minister's Office, grant number RA2000000281.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of ARIEL UNIVERSITY (certification number AU-ENG-RH-20200401 from 31 May 2021).

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

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy and ethic.

Conflicts of Interest: The authors declare no conflicts of interest. The authors have no financial or non-financial interests that are directly or indirectly related to the work submitted for publication.

References

1. Cambridge Dictionary. 2021. Available online: https://dictionary.cambridge.org/dictionary/english/information-exchange (accessed on 16 May 2023).
2. Talja, S.; Hansen, P. Information sharing. In *New Directions in Human Information Behavior*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 113–134.
3. DataReportal—Global Digital Insights. Available online: https://datareportal.com/global-digital-overview (accessed on 4 September 2022).
4. Statista. 2022. Available online: https://www.statista.com/statistics/617136/digital-population-worldwide/ (accessed on 4 September 2022).
5. Instagram. 2021. Available online: https://www.instagram.com (accessed on 17 May 2022).
6. Linkedin. 2022. Available online: https://about.linkedin.com/ (accessed on 17 May 2022).
7. Meta SEC Filings. 10-Q. 2022. Available online: https://investor.fb.com/financials/sec-filings-details/default.aspx?FilingId=15760347 (accessed on 17 May 2022).
8. Auxier, B.; Anserson, M. *Social Media Use in 2021*; Pew Research Center: Washington, DC, USA, 2021.
9. Oeldorf-Hirsch, A.; Birnholtz, J.; Hancock, J.T. Your post is embarrassing me: Face threats, identity, and the audience on Facebook. *Comput. Hum. Behav.* **2017**, *73*, 92–99. [CrossRef]
10. Schaik, P.V.; Jansen, J.; Onibokun, J.; Camp, J.; Kusev, P. Security and privacy in online social networking: Risk perceptions and precautionary behaviour. *Comput. Hum. Behav.* **2018**, *78*, 283–297. [CrossRef]
11. Silva, M.J.; Caled, D. Digital media and misinformation: An outlook on multidisciplinary strategies against manipulation. *J. Comput. Soc. Sci.* **2022**, *5*, 123–159.
12. Gross, R.; Acquisti, A. *Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 36–58.
13. Gross, R.; Acquisti, A. Information revelation and privacy in online social networks. In Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society, Alexandria, VA, USA, 7 November 2005; pp. 71–80.
14. Barnes, S.B. A privacy paradox: Social networking in the United States. *First Monday* **2006**, *11*, 9. [CrossRef]
15. Ellison, N.B.; Vitak, J.; Steinfield, C.; Gray, R.; Lampe, C. Negotiating privacy concerns and social capital needs in a social media environment. In *Privacy Online*; Springer: Berlin/Heidelberg, Germany, 2011.
16. Boyd, D.M.; Ellison, N.B. Social network sites: Definition, history, and scholarship. *J. Comput. Mediat. Commun.* **2007**, *13*, 210–230. [CrossRef]
17. Brake, D.R. *Sharing Our Lives Online: Risks and Exposure in Social Media*; Springer: Berlin/Heidelberg, Germany, 2014.
18. Kazutoshi, S.; Wen, C.; Hao, P.; Luca, C.G.; Alessandro, F.; Filippo, M. Social influence and unfollowing accelerate the emergence of echo chambers. *J. Comput. Soc. Sci.* **2021**, *4*, 381–402.
19. Hirschprung, R.S.; Alkoby, S. A Game Theory Approach for Assisting Humans in Online Information-Sharing. *Information* **2022**, *13*, 183. [CrossRef]
20. Puaschunder, J. A Utility Theory of Privacy and Information Sharing. In *Encyclopedia of Information Science and Technology*, 5th ed.; IGI Global: Hershey, PA, USA, 2021; pp. 428–448.
21. Schwartz-Chassidim, H.; Ayalon, O.; Mendel, T.; Hirschprung, R.; Toch, E. Selectivity in posting on social networks: The role of privacy concerns, social capital, and technical literacy. *Heliyon* **2020**, *6*, e03298. [CrossRef] [PubMed]
22. Hirschprung, R.; Toch, E.; Schwartz-Chassidim, H.; Mendel, T.; Maimon, O. Analyzing and Optimizing Access Control Choice Architectures in Online Social Networks. *ACM Trans. Intell. Syst. Technol.* **2017**, *8*, 1–22. [CrossRef]
23. Longstreet, P.A.B.S. Life satisfaction: A key to managing internet\& social media addiction. *Technol. Soc.* **2017**, *50*, 73–77.
24. Martin, K.; Murphy, P. The Role of Data Privacy in Marketing. *J. Acad. Mark. Sci.* **2016**, *45*, 135–155. [CrossRef]
25. Xu, W.; Dainoff, M.J.; Ge, L.; Gao, Z. *Interaction, From Human-Computer Interaction to Human-AI*; ACM: New York, NY, USA, 2021; p. 11.
26. Ooijen, I.V.; Vrabec, H.U. Does the GDPR Enhance Consumers' Control over Personal. *J. Consum. Policy* **2019**, *42*, 91–107. [CrossRef]
27. Muravyeva, E.; Janssen, J.; Specht, M.; Custers, B. Exploring Solutions to the Privacy Paradox in the Context of e-Assessment: Informed Consent Revisited. *Ethics Inf. Technol.* **2020**, *23*, 223–238. [CrossRef]

28. Böhme, R.; Köpsell, S. Trained to accept?: A field experiment on consent dialogs. In Proceedings of the 28th International Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010.
29. Barth, S.; De Jong, M.D. The privacy paradox-Investigating discrepancies between expressed privacy concerns and actual online behavior-A systematic literature review. *Telemat. Inform.* **2017**, *34*, 1038–1058. [CrossRef]
30. Greener, S. Unlearning with technology. *Interact. Learn. Environ.* **2016**, *24*, 1027–1029. [CrossRef]
31. Jensen, C.; Potts, C. Privacy policies as decision-making tools. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vienna, Austria, 24–29 April 2004.
32. Gerber, N.; Gerber, P.; Volkamer, M. Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Comput. Secur.* **2018**, *77*, 226–261. [CrossRef]
33. Koops, B.J.; Newell, B.C.; Timan, T.S.I.; Chokrevski, G.M. A typology of privacy. *Univ. Pa. J. Int. Law Rev.* **2016**, *38*, 483.
34. Tucker, C. *The Economics of Artifical Intelligence: An Agenda, Vols. Privacy, Algorithms, and Artificial Intelligence*; University of Chicago Press: Chicago, IL, USA, 2019.
35. Bashir, M.; Hayes, C.; Lambert, A.D.; Kesan, J.P. Online privacy and informed consent: The dilemma of information asymmetry. *Proc. Assoc. Inf. Sci. Technol.* **2016**, *52*, 1–10. [CrossRef]
36. Furman, S.; Theofanos, M. Preserving Privacy-More Than Reading a Message. In *Universal Access in Human-Computer Interaction. Design for All and Accessibility Practice*; Springer: Berlin/Heidelberg, Germany, 2014.
37. Cohen, J.E. Turning privacy inside out. *Theor. Inq. Law* **2019**, *20*, 1–31. [CrossRef]
38. James, T.L.; Warkentin, M.; Collignon, S.E. A dual privacy decision model for online social networks. *Inf. Manag.* **2015**, *52*, 893–908. [CrossRef]
39. Schaub, F.; Balebako, R.; Cranor, L.F. Designing Effective Privacy Notices and Controls. *IEEE Internet Comput.* **2017**, *21*, 70–77. [CrossRef]
40. Susser, D. Notice After Notice-and-Consent: Why Privacy Disclosures Are Valuable Even If Consent. *J. Inf. Policy* **2019**, *9*, 132–157.
41. Potzsch, S. Privacy Awareness: A Means to Solve the Privacy Paradox? In *IFIP Summer School on the Future of Identity in the Information Society*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 226–236.
42. Nissenbaum, H. A Contextual Approach to Privacy Online. *Daedalus* **2011**, *140*, 32–48. [CrossRef]
43. Acquisti, A.; Adjerid, I.; Balebako, R.; Brandimarte, L.; Cranor, L.F.; Komanduri, S.; Leon, P.G.; Sadeh, N.; Schaub, F.; Sleeper, M.; et al. Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online. *ACM Comput. Surv.* **2017**, *50*, 1–41. [CrossRef]
44. Fujio, T.; Hitoshi, Y.; Isamu, O. A belief in rewards accelerates cooperation on consumer-generated media. *J. Comput. Soc. Sci.* **2020**, *3*, 19–31.
45. Xiao, B.; Benbasat, I. E-commerce product recommendation agents: Use, characteristics, and impact. *MIS Q.* **2007**, *31*, 137–209. [CrossRef]
46. Benbasat, I.; Wang, W. Trust in and adoption of online recommendation agents. *J. Assoc. Inf. Syst.* **2005**, *6*, 4. [CrossRef]
47. Alamdari, P.M.; Navimipour, N.J.; Hosseinzadeh, M.; Safaei, A.A.; Darwesh, A. A Systematic Study on the Recommender Systems in the E-Commerce. *IEEE Access* **2020**, *8*, 115694–115716. [CrossRef]
48. Nilashi, M.; Jannach, D.; Ibrahim, O.B.; Esfahani, M.D.; Ahmadi, H. Recommendation quality, transparency, and website quality for trust-building in recommendation agents. *Electron. Commer. Res. Appl.* **2016**, *19*, 70–84. [CrossRef]
49. Madsen, M.; Gregor, S. Measuring human-computer trust. In Proceedings of the 11th Australasian Conference on Information Systems, Brisbane, Australia, 6–8 December 2000.
50. Leyer, M.; Schneider, S. Me, You or AI? How do we Feel about Delegation. In Proceedings of the 27th European Conference on Information Systems (ECIS), Stockholm, Sweden, 8–14 June 2019.
51. Cominelli, L.; Feri, F.; Garofalo, R.; Giannetti, C.; Meléndez-Jiménez, M.A.; Greco, A.; Nardelli, M.; Scilingo, E.P.; Kirchkamp, O. Promises and trust in human–robot interaction. *Sci. Rep.* **2021**, *11*, 9687. [CrossRef]
52. Sent, E.M.; Klaes, M. A conceptual history of the emergence of bounded rationality. *Hist. Political Econ.* **2005**, *37*, 25–59.
53. Taha, R.J.Y. Fooled by facts: Quantifying anchoring bias through a large-scale experiment. *J. Comput. Soc. Sci.* **2022**, *5*, 1001–1021.
54. Kobsa, K.A.; Cho, H.; Knijnenburg, B.P. The effect of personalization provider characteristics on privacy attitudes and behaviors: An Elaboration Likelihood Model approach. *J. Assoc. Inf. Sci. Technol.* **2016**, *67*, 2587–2606. [CrossRef]
55. Wang, Y.; Kobsa, A. A PLA-based privacy-enhancing user modeling framework and its evaluation. *User Model. User Adapt. Interact.* **2013**, *23*, 41–82. [CrossRef]
56. GDPR. Recital 60. 2021. Available online: https://www.privacy-regulation.eu/en/recital-60-GDPR.htm (accessed on 10 May 2022).
57. Brkan, M. Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the GDPR and beyond. 2019. Available online: https://www.researchgate.net/journal/International-Journal-of-Law-and-Information-Technology-1464-3693 (accessed on 16 August 2022).
58. Crutzen, R.; Peters, G.Y.; Mondschein, C. Why and how we should care about the General Data Protection Regulation. *Psychol. Health* **2019**, *34*, 1347–1357. [CrossRef]
59. Burgess, M.M. Proposing modesty for informed consent. *Soc. Sci. Med.* **2007**, *65*, 2284–2295. [CrossRef]
60. Glikson, E.; Woolley, A.W. Human trust in artificial intelligence: Review of empirical research. *Acad. Manag. Ann.* **2020**, *14*, 627–660. [CrossRef]

61. Araujo, T.; Helberger, N.; Kruikemeier, S.; De Vreese, C.H. In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI Soc.* **2020**, *35*, 611–623. [CrossRef]
62. Brzowski, M.; Nathan-Roberts, D. Trust measurement in human-automation interaction: A systematic review. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; SAGE Publications Sage CA: Los Angeles, CA, USA, 2019; pp. 1595–1599.
63. Hoff, K.A.; Bashir, M. Trust in automation: Integrating empirical evidence on factors that influence trust. *Hum. Factors* **2015**, *57*, 407–434. [CrossRef]
64. Corritore, L.; Kracher, B.; Wiedenbeck, S. On-line trust: Concepts, evolving themes, a model. *Int. J. Hum. Comput. Stud.* **2003**, *58*, 737–758. [CrossRef]
65. Longoni, C.; Bonezzi, A.; Morewedge, C.K. Intelligence, Resistance to Medical Artificial. *J. Consum. Res.* **2019**, *46*, 629–650. [CrossRef]
66. Davenport, T.; Guha, A.; Grewal, D.; Bressgott, T. How artificial intelligence will change the future of marketing. *J. Acad. Mark. Sci.* **2019**, *48*, 24–42. [CrossRef]
67. Kelly, S.; Kaye, S.-A.; Oviedo-Trespalacios, O. What factors contribute to the acceptance of artificial intelligence? A systematic review. *Telemat. Inform.* **2023**, *77*, 101925. [CrossRef]
68. Venkatesh, V. Adoption and use of AI tools: A research agenda grounded in UTAUT. *Ann. Oper. Res.* **2022**, *308*, 641–652. [CrossRef]
69. Davis, F.; Bagozzi, R.; Warshaw, P. User Acceptance of Computer Technology: A Comparison of Two Theoretical Model. *Manag. Sci.* **1989**, *35*, 982–1003. [CrossRef]
70. Portz, J.; Bayliss, E.; Bull, S.; Boxer, R.; Bekelman, D.; Gleason, K.; Czaja, S. Using the Technology Acceptance Model to Explore User Experience, Intent to Use, and Use Behavior of a Patient Portal Among Older Adults with Multiple Chronic Conditions: Descriptive Qualitative Study. *J. Med. Internet Res.* **2019**, *21*, e11604. [CrossRef]
71. Qiu, L.; Benbasat, I. Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems. *J. Manag. Inf. Syst.* **2009**, *25*, 145–182. [CrossRef]
72. Dietvorst, B.J.; Simmons, J.P.; Massey, C. Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Manag. Sci.* **2018**, *64*, 1155–1170. [CrossRef]
73. Elmalech, A.; Sarne, D.; Rosenfeld, A.; Erez, E. When suboptimal rules. In Proceedings of the AAAI Conference on Artificial Intelligence, Austin, TX, USA, 25–30 January 2015.
74. Pita, J.; Jain, M.; Tambe, M.; Ordónez, F.; Kraus, S. Robust solutions to Stackelberg games: Addressing bounded rationality and limited observations in human cognition. *Artif. Intell.* **2010**, *174*, 1142–1171. [CrossRef]
75. Lovejoy, J. The UX of AI. Google Design. 2018. Available online: https://design.google/library/ux-ai/ (accessed on 30 August 2022).
76. Ling, E.C.; Tussyadiah, I.; Tuomi, A.; Stienmetz, J.; Ioannou, A. Factors influencing users' adoption and use of conversational agents: A systematic review. *Psychol. Mark.* **2021**, *38*, 1031–1051. [CrossRef]
77. Bălău, N.; Utz, S. Information sharing as strategic behaviour: The role of information display, social motivation and time pressure. *Behav. Inf. Technol.* **2017**, *36*, 589–605. [CrossRef]
78. Alohali, M.; Clarke, N.; Li, F.; Furnell, S. Identifying and predicting the factors affecting end-users' risk-taking behavior. *Inf. Comput. Secur.* **2018**, *26*, 306–326. [CrossRef]
79. Li, S.; Blythe, P.; Zhang, Y.; Edwards, S.; Guo, W.; Ji, Y.; Goodman, P.; Hill, G.; Namdeo, A. Analysing the effect of gender on the human–machine interaction in level 3 automated vehicles. *Sci. Rep.* **2022**, *12*, 11645. [CrossRef] [PubMed]
80. Amichai-Hamburger, Y.; Vinitzky, G. Social network use and personality. *Comput. Hum. Behav.* **2010**, *26*, 1289–1295. [CrossRef]
81. Phillips-Wren, G. AI tools in decision making support systems: A review. *Int. J. Artif. Intell. Tools* **2012**, *21*, 1240005. [CrossRef]
82. Chen, L.; Tsoi, H. Users' decision behavior in recommender interfaces: Impact of layout design. In Proceedings of the RecSys' 11 Workshop on Human Decision Making in Recommender Systems, Chicago, IL, USA, 23–27 October 2011.
83. Sohn, K.; Kwon, O. Technology acceptance theories and factors influencing artificial Intelligence-based intelligent products. *Telemat. Inform.* **2020**, *47*, 101324. [CrossRef]
84. Hanna, J. Consent and the Problem of Framing Effects. *Ethical Theory Moral Pract.* **2011**, *14*, 517–531. [CrossRef]
85. Ellison, A.; Coates, K. *An Introduction to Information Design*; Laurence King Publishing: London, UK, 2014.
86. Jung, H.; Cui, X.; Kim, H.L.; Li, M.; Liu, C.; Zhang, S.; Yang, X.; Feng, L.; You, H. Development of an Ergonomic User Interface Design of Calcium Imaging Processing System. *Appl. Sci.* **2022**, *12*, 1877. [CrossRef]
87. Mertens, S.; Herberz, M.; Hahnel, U.J.; Brosch, T. The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2107346118. [CrossRef]
88. Akhawe, D.; Felt, A. Alice in warningland: A large-scale field study of browser security warning effectiveness. In Proceedings of the 22nd USENIX Security Symposium (USENIX Security 13), Washington, DC, USA, 14–16 August 2013.
89. Gray, K. *AI Can Be a Troublesome Teammate*; Harvard Business Review: Brighton, MA, USA, 2017; pp. 20–21.
90. Simon, C.; Amarilli, S. Feeding the behavioral revolution: Contributions of behavior analysis to nudging and vice versa. *J. Behav. Econ. Policy* **2018**, *2*, 91–97.
91. Van Lent, M.; Fisher, W.; Mancuso, M. An explainable artificial intelligence system for small-unit tactical behavior. In Proceedings of the National Conference on Artificial Intelligence, San Jose, CA, USA, 25–29 July 2004.
92. Gunning, D. Explainable artificial intelligence (xai). *Def. Adv. Res. Proj. Agency* **2017**, *2*, 2. [CrossRef] [PubMed]
93. Miller, T. Explanation in artificial intelligence: Insights from the social sciences. *Artif. Intell.* **2019**, *267*, 1–38. [CrossRef]

94. Islam, M.R.; Ahmed, M.U.; Barua, S.; Begum, S. A Systematic Review of Explainable Artificial Intelligence in Terms of Different Application Domains and Tasks. *Appl. Sci.* **2022**, *12*, 1353. [CrossRef]
95. Czitrom, V. One-factor-at-a-time versus designed experiments. *Am. Stat.* **1999**, *53*, 126–131. [CrossRef]
96. Montgomery, D.C. *Design and Analysis of Experiments*; John Wiley & Sons: Hoboken, NJ, USA, 2017.
97. Von Eye, A. Fractional Factorial Designs in the Analysis of Categorical Data. Citeseer. 2008. Available online: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=d759e93d084ada161e412ad461374f2644a72818 (accessed on 15 July 2022).
98. Khokhar, R.H.; Chen, R.; Fung, B.C.; Lui, S.M. Quantifying the costs and benefits of privacy-preserving health. *J. Biomed. Inform.* **2014**, *50*, 107–121. [CrossRef] [PubMed]
99. Liao, C.; Iyer, H. A Stochastic Algorithm for Selecting of Defining Contrasts in Two-Level Experiments. *Biom. J. J. Math. Methods Biosci.* **1999**, *41*, 671–678. [CrossRef]
100. MathWorks. Fracfactgen. 2022. Available online: https://www.mathworks.com/help/releases/R2016b/stats/fracfactgen.html?searchHighlight=fracfactgen&s_tid=doc_srchtitle (accessed on 15 July 2022).
101. Dreamgrow. The 15 Biggest Social Media Sites and Apps. 2022. Available online: https://www.dreamgrow.com/top-15-most-popular-social-networking-sites/ (accessed on 17 January 2022).
102. Statista. Leading Countries Based on Facebook Audience Size as of January 2022. 2022. Available online: https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/ (accessed on 8 June 2021).
103. United States Census Bureau. Statistical Quality Standard E1: Analyzing Data. 2023. Available online: https://www.census.gov/about/policies/quality/standards/standarde1.html (accessed on 8 January 2024).
104. Thiebes, S.; Lins, S.; Sunyaev, A. Trustworthy artificial intelligence. *Electron. Mark.* **2021**, *31*, 447–464. [CrossRef]
105. Zhang, C.; Lu, Y. Study on artificial intelligence: The state of the art and future prospects. *J. Ind. Inf. Integr.* **2021**, *23*, 100224. [CrossRef]
106. Caruana, R.; Lou, Y.; Gehrke, J.; Koch, P.; Sturm, M.; Elhadad, N. Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia, 10–13 August 2015.
107. Wong, A.; Otles, E.; Donnelly, J.P.; Krumm, A.; McCullough, J.; DeTroyer-Cooley, O.; Pestrue, J.; Phillips, M.; Konye, J.; Penoza, C. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. *JAMA Intern. Med.* **2021**, *181*, 1065–1070. [CrossRef]
108. The White House. Blueprint for an AI Bill of Rights. October 2022. Available online: https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf (accessed on 8 January 2024).
109. O'Neil, C. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy; Crown Publishing Group, New York City, NY, USA, 2017.
110. Dwivedi, R.; Dave, D.; Naik, H.; Singhal, S.; Omer, R.; Patel, P.; Qian, B.; Wen, Z.; Shah, T. Morgan. Explainable AI (XAI): Core ideas, techniques, and solutions. *ACM Comput. Surv.* **2023**, *55*, 1–33. [CrossRef]
111. Selbst, A.D.; Powles, J. "Meaningful information" and the right to explanation. In Proceedings of the 1st Conference on Fairness, Accountability and Transparency, New York, NY, USA, 23–24 February 2018.
112. EUArtificialIntelligence. The EU Artificial Intelligence Act. 2024. Available online: https://artificialintelligenceact.eu/ (accessed on 24 March 2024).
113. Beel, J.; Dixon, H. The 'Unreasonable' Effectiveness of Graphical User Interfaces for Recommender Systems. In Proceedings of the Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization, Utrecht, The Netherlands, 21–25 June 2021.
114. Karran, A.J.; Demazure, T.; Hudon, A.; Senecal, S.; Léger, P.M. Designing for Confidence: The Impact of Visualizing Artificial Intelligence Decisions. *Front. Neurosci.* **2022**, *16*, 883385. [CrossRef] [PubMed]
115. Pillai, R.; Sivathanu, B. Adoption of AI-based chatbots for hospitality and tourism. *Int. J. Contemp. Hosp. Manag.* **2020**, *32*, 3199–3226. [CrossRef]
116. Kim, J. The influence of perceived costs and perceived benefits on AI-driven interactive recommendation agent value. *J. Glob. Sch. Mark. Sci.* **2020**, *30*, 319–333. [CrossRef]
117. Alves, T.; Natálio, J.; Henriques-Calado, J.; Gama, S. Incorporating personality in user interface design: A review. *Personal. Individ. Differ.* **2020**, *155*, 109709. [CrossRef]
118. Knijnenburg, B.P.; Reijmer, N.J.; Willemsen, M.C. Each to his own: How different users call for different interaction methods in recommender systems. In Proceedings of the Fifth ACM Conference on Recommender Systems, Chicago, IL, USA, 23–27 October 2011.
119. Jin, Y.; Tintarev, N.; Htun, N.N.; Verbert, K. Effects of personal characteristics in control-oriented user interfaces for music recommender systems. *User Model. User Adapt. Interact.* **2020**, *30*, 199–249. [CrossRef]
120. Kissell, R. *Algorithmic Trading Methods*, 2nd ed.; Elsevier Wordmark: Amsterdam, The Netherlands, 2020.