

Understanding Trust Formation in GPT Services: An Empirical Study

Helena Li, Irina Rychkova*

University Paris 1, Pantheon-Sorbonne, Paris, France

Abstract

AI assistants grounded on Large Language Models (LLM) and Generative Pretrained Transformers (GPT), with ChatGPT as a popular example, are a new generation of technology, simulating human-like behavior, social interaction, and provoking emotional feedback from users. In this study, we examine the determinants affecting users' trust in GPT services. Using structural equation modeling, we analyze data collected from 124 respondents. Compared to previous studies grounded on theories of technology acceptance, we focus on social aspects of trust formation. Our conceptual framework integrates key Trust determinants from the Integrative Model of Organizational Trust and examines the role of knowledge in technology in general and AI awareness in particular in trust formation.

Keywords

trust, generative AI, GPT

1. Introduction

Generative Pretrained Transformer (GPT) marks a significant breakthrough in the NLP domain, revolutionizing our way to interact with technology and opening avenues for applications in numerous industries. With ChatGPT considered the fastest-growing consumer application in history numerous concerns and challenges related to GPT adoption are raised [1]. Technology trustworthiness is one of them [2][3].

Trust is a social construct that emerges from relationships and interactions between individuals or groups. It involves a willingness to rely on others [4, 5] and is influenced by factors such as past experience, reputation, and social norms. Trust in technology is described by a situation in which an individual user or an organization (trustor) is willing to rely on technology (trustee) to accomplish a specific task [6]. In contrast to interpersonal (or social) trust, trustworthiness of technology is mainly identified with its specific technical properties (e.g., security, fault tolerance etc)[6]. However, for the modern technologies, including GPT, technical properties are not sufficient predictors of trust: Users are often unable to objectively reason about technical properties due to complexity of a system or service [7]. Moreover, AI assistants powered by GPT technology "can understand and communicate using language in a manner that closely resembles that of humans" [1]. They simulate human-like behavior, social interaction, and provoke emotional feedback, comparable to interpersonal relationships. This shows an importance of factors of interpersonal (or social) trust in technology. In this study, we closely examine the trust formation process in the context of GPT services and the role of social trust in GPT services acceptance and use.

We define a theoretical model for trust in GPT services grounded on the Integrative Model of Organizational Trust [5]. While this model has been extensively applied in organizational settings, there is currently limited explicit application of this model in studies on trust in ChatGPT or Generative AI services. In particular, we explore how trust and decision to use GPT services depend on (a) the factors of interpersonal trust [5] and (b) the socio-demographic characteristics of a user, including education

RCIS 2025 Workshops and Research Projects Track. 20 - 23 May, 2025. Seville, Spain.

*Corresponding author.

✉ helena.li@univ-paris1.fr (H. Li); irina.rychkova@univ-paris1.fr (I. Rychkova)

ORCID 0000-0002-1100-0116 (I. Rychkova)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

and awareness in AI. We provide an empirical validation of our model, conducting a survey among 124 participants. This study aims to contribute to the existing literature by:

- Reaffirming the role of trust as a significant predictor of user engagement with GPT services, aligning with previous research on the influence of trust in technology adoption.
- Expanding the conceptualization of trust in GPT services by recognizing the role of its anthropomorphic characteristics in shaping the user-technology relationship.

The remainder of this paper is organized as follows: in Section 2, we present our foundational concepts and discuss the related works; in Section 3 we present our theoretical model for trust in GPT services and detail our research methodology; in Section 4, we discuss the results of our analysis; in Section 5, we present our conclusions.

2. Background and Related Works

2.1. Generative AI and GPT technology

Artificial Intelligence (AI) is defined as "intelligence exhibited by machines, particularly computer systems" [8]. Generative AI (GenAI) are AI models able to generate content, whether in textual, audio, image or video form. GenAI models use neural network models and usually answer to a 'prompt' which is the textual input from a user. Transformer model in particular, developed by Google in 2017 [9], allowed for shorter training as it is a model with no recurrence and worked by transforming text into numerical values (tokens).

Generative Pretrained Transformer (GPT) is a deep learning model based on the Transformer architecture, designed for natural language processing tasks. GPT models are pretrained on a large corpus of text data and fine-tuned for specific tasks. They can understand and generate contextually relevant text, which makes them widely used in AI assistants, chatbots, and other NLP applications [10]. The first GPT model was introduced in 2018 by OpenAI. Their product 'ChatGPT' is considered a breakthrough in the field. ChatGPT takes the form a chatbot, in which the user is given a field to input text, and from which the service will answer in a human natural language.

Other products such as Microsoft Bing Chat, Github Copilot, DeepSeek etc., while based on different implementations and architectures, use GPT technology as their backbone. We refer to the products of this family as 'GPT services' in this study.

While popularity of GPT services grows, is still much to be explored about their capabilities and limitations [1]. Trustworthiness is one of the main challenges in GPT acceptance [2]. "Low trust in a highly capable technology would be a huge productivity loss, whereas high trust in a less performant technology can lead to over-reliance and misuse of a technology" [11].

2.2. Social Trust and Trust in Technology

In the research literature on trust, the act of trust is often represented as a relationship between a subject (the trustor) and an object of trust (the trustee) [12][13]. It is characterized by the trustor's willingness to be vulnerable, to rely upon trustee in some situation where risk is involved. Outcome of trust is defined as an actual engagement or interaction between trustor and trustee.

The Integrative Model of Organization Trust, developed by Mayer, Davis, and Schoorman [5], provides a theoretical framework for understanding how trust between organizations and/or individuals is formed. It is composed of three factors of perceived trustworthiness, which contribute to interpersonal trust:

- *Perceived Ability*: The skills, competencies, and characteristics that enable a trustee to have influence within some specific domain.
- *Perceived Benevolence*: The extent to which a trustee is believed to want to do good to the trustor, aside from an egocentric profit motive.

- *Perceived Integrity*: The perception that the trustee adheres to a set of principles that the trustor finds acceptable.

Those factors are directly affecting trust and are moderated by the trustor's propensity, that is, the general willingness of the trustor to trust others which can be explained by the personality, experience, etc, and by the perceived risk that comes from the nature of this interaction [5].

The authors of [14] argue that the dimensions of trustworthiness proposed by Mayer [5] are poorly suited for studying trust relationships between users and IT artifacts, since they are defined to fit the human character traits and human decision making. For example, to assess a 'perceived benevolence' of an IT artifact one has to assume that this artifact is able to actively decide whether to act in the interest of the user (trustor) or not.

As an alternative to social trust, *Trust in specific technology* is widely addressed in the literature [11, 15, 16, 7, 14, 17]. McKnight et al. [6] provides a framework for understanding how trust in technology is formed and its effects on technology usage. The authors put forward performance, functionality and reliability as the factors of trust in specific technology. Institution-based trust, including situational normality and structural assurance, exerts a mediated positive effect on post-adoption technology use according to [6].

According to [1], GPT services exhibit strong resemblance to human behavior and ability to simulate human decision making and character traits. This makes us reconsider the question 'Do people rely on the same dimensions of trustworthiness when deciding whether or not to trust other people compared to deciding whether or not to trust an IT artifact?' debated in [14].

2.3. Theoretical Models of Technology Acceptance

In empirical research, theoretical models specify a set of constructs and the relationships between them that explain a phenomenon of interest. The constructs can play a role of predictors, mediators or moderators for the examined phenomena. Predictors are independent variables that directly influence the dependent variable. Mediators are variables that explain the mechanism through which predictors influence the dependent variable. They act as intermediaries in the causal chain, helping to clarify how or why a certain effect occurs. Moderators are variables that affect the strength or direction of the relationship between predictors and the dependent variable. They provide insights into when or under what conditions certain effects occur.

Theoretical models such as the Technology Acceptance Model (TAM) [18] or its extension, the Unified Theory of Acceptance and Use of Technology (UTAUT) [19] are often applied within the context of information systems to understand predictors of human behavior toward potential acceptance or rejection of the technology.

TAM defines two predictors of acceptance (or acceptance factors): Perceived usefulness is the degree to which a person believes that using a particular system would enhance his or her job performance. Perceived ease of use is the degree to which a person believes that using a particular system would be free of effort.

The Unified Theory of Acceptance and Use of Technology (UTAUT) was created in 2004 [19]. Built upon the TAM, it reviewed multiple TAM models and identified the four factors of acceptance: performance expectancy, effort expectancy, social influence and facilitating conditions. Whereas the first two factors are drawn upon the original TAM, the last two elaborate on the role of social relationships and environment in acceptance. Social Influence is defined as the perceived importance for an individual to use new technology by other people. Facilitating conditions is defined as the belief of an individual that "an organization and technical infrastructure exists to support the use of a system [technology]" [19]. Those four factors are influenced by four moderators: age, gender, experience, and voluntariness of use, which result in possible gaps between the intentions to use a technology and the actual use of the technology.

Acceptance predictors from TAM and UTAUT show the importance of combining technical factors and social factors, related to user's personality and the context of use. While trust is placed among the

second-order factors of acceptance, the authors agree that without trust, users are unlikely to engage with a technology or adopt it for their needs [18].

2.4. Trust and Acceptance of GPT services in the Literature

Factors of rapid acceptance and adoption of GPT services and ChatGPT in particular draw a lot of attention in the research community. The authors of [20] and [21] use extended TAM to examine factors of ChatGPT adoption. In [20], the study is conducted among 352 students from 12 higher education institutions. The results highlight the response quality and user-friendliness of ChatGPT as main factors of its adoption. In [21], the study is conducted among Chinese university students. The findings highlight the importance of trust in the adoption process of ChatGPT: in particular, the study shows that perceived trust moderates the relationship between awareness about ChatGPT and perceived ease of use, usefulness - main concepts of TAM. The study reported in [22] focuses on user trust and its influence on the intention and actual use of ChatGPT (no reference to a particular theoretical model is provided). The survey reveals that trust has a significant direct effect on both the intention to use and actual use of ChatGPT. The authors of [23] and [24] explore factors influencing acceptance and use of ChatGPT using an extended UTAUT model. The work in [23] shows that performance expectancy, effort expectancy, hedonic motivation, facilitating conditions, and habit positively impact the behavioral intention to use ChatGPT. Trust moderates the relationship between behavioral intention and actual use behavior. The findings presented in [24] reveal that relative risk perception and emotional factors play significant roles in predicting behavioral intentions toward ChatGPT.

These studies indicate the significant role trust plays in acceptance of ChatGPT. While using the theoretical acceptance models (i.e., TAM, UTAUT) as underlying theories, these works do not specifically address trust formation process. Moreover, in these studies, trust is often co-notated with technical aspects (e.g., privacy and security) of the considered technology, ignoring the social and emotional aspects.

Earlier works examine trust in AI using social conceptualization of trust: The authors of [16] examine the nature of trust in AI and discuss a model inspired by interpersonal trust. The authors associate the AI trustworthiness with the AI model correctness and commitment to some contract (contractual trust). In [11], the authors propose a theoretical framework and discuss the determinants of human trust in AI. This study identifies tangibility, transparency, reliability and immediacy of AI technology with formation of cognitive trust (based on rational thinking) in users, whereas the AI's anthropomorphism plays an important role in emotional trust (based on affection).

This study goes one step further from the conventional research exploring ChatGPT acceptance factors and is designed to investigate the process of trust formation and the role of social predictors of trust in GPT services. We propose a theoretical model for trust in GPT services grounded upon Mayer et al. [5]. We extend the existing body of knowledge explicitly linking trust in ChatGPT to the established social factors of trust: Ability, Benevolence and Integrity.

3. Methodology

3.1. Theoretical Model and Hypotheses Development

Table 1 summarizes the theoretical models of trust and technology acceptance discussed in the previous section. We draw upon [5] to define the trust formation process. Figure 1 presents the theoretical model for this study.

3.1.1. Social factors of trustworthiness as predictors of Trust in GPT services.

Perceived Ability, Benevolence and Integrity are the main factors of trust according to [5]. In our context, we specify these construct integrating the trust indicators from our earlier works [17] and from the theories of technology acceptance from Table 1. Perceived Ability construct is defined as

Table 1
Summary on the theoretical models of trust and acceptance

Relationship	Subject	Object	Predictors/ Mediators	Moderators	Outcome
Social Trust [5]	Organization / Individual	Organization / Individual	Ability, benevolence, integrity	Propensity to trust, perceived risk	Interaction, collaboration
Trust in Techn. [6]	Organization / Individual	Technology (system, service)	Performance, functionality, reliability	Propensity to trust, institution-based trust	Acceptance, use
Techn. acceptance [18, 19]	Organization / Individual	Technology (system, service)	Performance expectancy, effort expectancy, social influence facilitating conditions	Age, experience, voluntariness of use	Acceptance, use

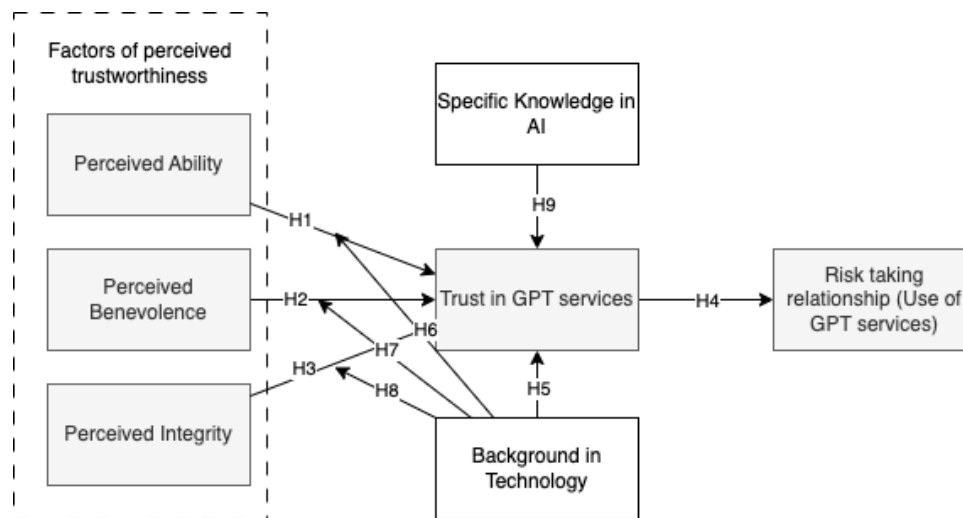


Figure 1: Theoretical model of Trust in GPT services based on [5]

competence (including technical functionality, performance, usability [25]) of a trustee (GPT services in our case) to have influence within some specific domain. It is closely related to perceived usefulness in TAM [18] and performance expectancy in UTAUT [19]. Perceived Benevolence is defined as the extent to which GPT services are believed to act in the interest of trustor (user), for example, by protecting her personal data or ensuring exactitude, fairness and objectivity of provided answers. In digital world, benevolence is closely associated with perceived privacy and security of data [25]. Perceived Integrity is defined as the perception that GPT services adhere to a set of principles (ethical, legal or other) that the trustor finds acceptable, for example, by ensuring verifiable and traceable answers. This construct can be also associated with credibility and transparency in technology [26][27].

H1: Perceived Ability has a significant effect on Trust in GPT services.

H2: Perceived Benevolence has a significant effect on Trust in GPT services.

H3: Perceived integrity has a significant effect on Trust in GPT services.

3.1.2. Trust in GPT services and Use of GPT services.

Trust is a core variable in our model. We define Trust in GPT services as willingness of user to rely on these services in a situation where a risk or a negative outcome is possible. Risk Taking in relationship [5] defines the actual Use of GPT services in our model.

H4: Trust has a significant positive effect on Use of GPT services.

3.1.3. The role of technical background in trust formation.

The Integrative model of organizational trust defines Propensity to trust as an individual characteristic of a trustor which predicts interpersonal Trust and also moderates the impact of PA, PB, PI on Trust. In our model, we consider that the role of propensity to trust can be fulfilled by the trustor's knowledge about technology: we suggest that the users exposed to technology in general (via their academic background and professional experience) are more likely to rely on a new technology such as GPT services and trust it. Moreover, we suggest that a user with a strong Background in Technology will be more aware about technical functionalities and other characteristics of GPT services constituting Perceived Ability, Benevolence and Integrity. Therefore Background in Technology can moderate the effect of these factors of trust.

H5: User's Background in Technology has a significant effect on Trust in GPT services.

H6-H8: User's Background in Technology has a moderating effect on her Perceived Ability, Benevolence, and Integrity of GPT services.

3.1.4. The role of specific knowledge in AI in trust formation.

We consider that Specific Knowledge in AI can also improve trust in GPT services.

H9: Specific Knowledge in AI has a significant effect on Trust in GPT services.

3.2. Measurement Model

We use reflective latent constructs to model Perceived Ability, Benevolence and Integrity as well as Trust and Use (see Table 2). A construct is modeled as latent if it cannot be measured directly.

In our study, Perceived Ability is associated with perceived functionality, efficiency [28], usefulness [18], performance expectancy and effort expectancy [19]. It is measured as a latent variable, using 5 items PA1-PA5. Perceived Benevolence - 'willingness to do good' according to [5] - is associated with confidentiality, assertion that GPT services will not share or disclose user personal data or chat details. It is also associated with objectivity (non bias) of responses [17]. Perceived Integrity is associated with transparency and traceability in our model. Each of these constructs is measured with 4 indicators PB1-PB4, PI1-PI4 (see Table 2).

Dependent variables Trust and Use are also modeled as latent variables. We measure Trust by suggesting respondents a specific situation and asking whether they would trust GPT Service in such situation (T1-T4). The indicator T5 is used to measure how important for the respondents is the fact that the content is produced by a human (it is reverse coded). To measure Use, we define the items that are not Likert-coded. In U1, the respondents are invited to checkbox GPT services they know/use in the list. In U2, U3, we measure frequency of use for leisure and work as ordinal variables.

Each item is measured using a 7 point Likert scale. Background in Technology and Specific Knowledge in AI are calculated from the categorical variables in socio-demographic data (see Table 3). For example, Background in Technology is calculated as: *Degree in Tech. x Education*.

3.3. Data Collection and Sample Characteristics

The questionnaire has been designed iteratively and tested on a small convenience sample for a feedback. Final version of the questionnaire has been distributed among the students (bachelor and master) of MIAE master program of Sorbonne University and on the professional LinkedIn network of the authors (non-probabilistic accidental sampling [29]). Data has been collected in 2 periods: between February and April 2024 and between November and December 2024. Collected through Google Forms, the data was exported into CSV, cleaned and coded for further analysis.

In total, 124 responses have been collected (N=124). Table 3 summarizes the socio-demographic data on the sample.

Table 2
Measurement model

<i>Construct</i>	<i>Item</i>	<i>Measurement</i>
Perceived Ability	PA1	Cover letters written by GPT services are better than if I wrote them myself
	PA2	The use of GPT services for candidate screening makes the recruitment process more efficient
	PA3	I think GPT services are more efficient than Tabnine (or other similar tool for coding assistance) to generate/review code and save me time
	PA4	I think GPT services are more convenient than Tabnine (or other similar tool for coding assistance) for coding
	PA5	Working with GPT services is easier than configuring a query on a searching engine like Google scholar
Perceived Benevolence	PB1	I would not mind typing personal details like my location, dietary restriction, allergies in a GPT service
	PB2	I think GPT Services would give a restaurant recommendation that is less biased than a person's
	PB3	GPT services for candidate screening make less biased decisions than a human actor
	PB4	I would use code, or details about my work/project when prompting GPT services for help
Perceived Integrity	PI1	GPT services are more transparent as a recruiting tool than human actors
	PI2	I think the sources GPT services return are traceable
	PI3	I think GPT algorithm and/or data training is transparent.
	PI4	The fact that Wikipedia provide the references to the sources is important to me (reversed coding)
Trust	T1	I would trust a GPT service to find a restaurant for a friend meeting
	T2	I would trust a GPT service (e.g., ChatGPT) to conduct a literature research and text editing for an essay/article/thesis
	T3	I would trust a GPT service (e.g., copilot) to help me correcting a bug in the code / getting the code ready for production today
	T4	I would trust a GPT service (e.g., chatbot) to help me with my CV and/or cover letter.
	T5	The fact that Wikipedia articles are written by human contributors is important to me (reversed coding)
Use	U1	What kind of GPT services do you use?
	U2	How often do you use ChatGPT for work?
	U3	How often do you use ChatGPT for fun/personal stuff?

3.4. Data Analysis

We conducted the data analysis using JASP - an open-source program for statistical analysis supported by the University of Amsterdam [30].

We use Structural Equation Modeling (SEM) method to test the defined hypotheses [31]. Following the recommendations from [32], we choose to use Covariance-Based Structural Equation Modeling (CB-SEM) for the following reasons: the goal of our study is to test and confirm a well-established

Table 3
Socio-Demographic Data

Variable		n (N=124)	%	Variable		n (N=124)	%
Age	18-	1	1%	Education	High school	14	11%
	18-25	96	77%		Bachelor	49	40%
	26-35	14	11%		Master	48	39%
	36-50	8	6%		PhD	10	8%
	51+	5	4%		Other	3	2%
Degree in Tech.	Yes	113	91%	Knowledge in AI	None	6	5%
	No	11	9%		Some	39	31%
					Fairly good	26	21%
					Expert	53	43%

theory and not a not to explore a new theory; we consider that our model consists predominantly of reflective constructs (where indicators are manifestations of the latent variable).

An independent sample t-test with mean comparison verification is conducted to analyze differences in Trust and Use of GPT services between two samples, corresponding to the two evaluation periods in our survey, separated by a nine-month interval.

4. Results

In this section, first, we evaluate our measurement model and determine if the collected data represents reliably the theoretical constructs; than we discuss the hypotheses testing results.

4.1. Measurement Model Assessment

We asses factor loadings, convergent and discriminant validity of the model. Factor loadings represent the correlation between latent variables (factors) and their measured items. The values of factor loadings are generally expected to be ≥ 0.7 - for strong loading (good representation of the factor) and 0.4 - 0.7 for acceptable loading. For our model, all latent variables show moderate to strong loading, with the lowest value for PI2 = 0.404. All factor loadings are statistically significant with p-value < 0.001 .

We use Cronbach's Alpha to asses model reliability (≥ 0.7 for good reliability). The overall reliability $\alpha = 0.828$ shows that the full set of items is reliable even though some individual constructs are weak (e.g., Trust: $\alpha = 0.296$).

Discriminant validity measures pairwise how each construct is distinct from other constructs in the model. We use Heterotrait-Monotrait (HTMT) ratio to asses discriminant validity in our model. Most construct pairs have HTMT < 0.85 (validity threshold), indicating that they are sufficiently distinct. However, Benevolence and Integrity (HTMT = 0.872) indicates higher correlation between the two factors.

We conclude, that convergent validity is supported in the model as the factor loadings are high, indicating that indicators are strongly related to their respective latent constructs. Discriminant validity is partially supported. The high correlation between Benevolence, and Integrity suggests some overlap.

4.2. Path Estimates and Hypotheses Testing

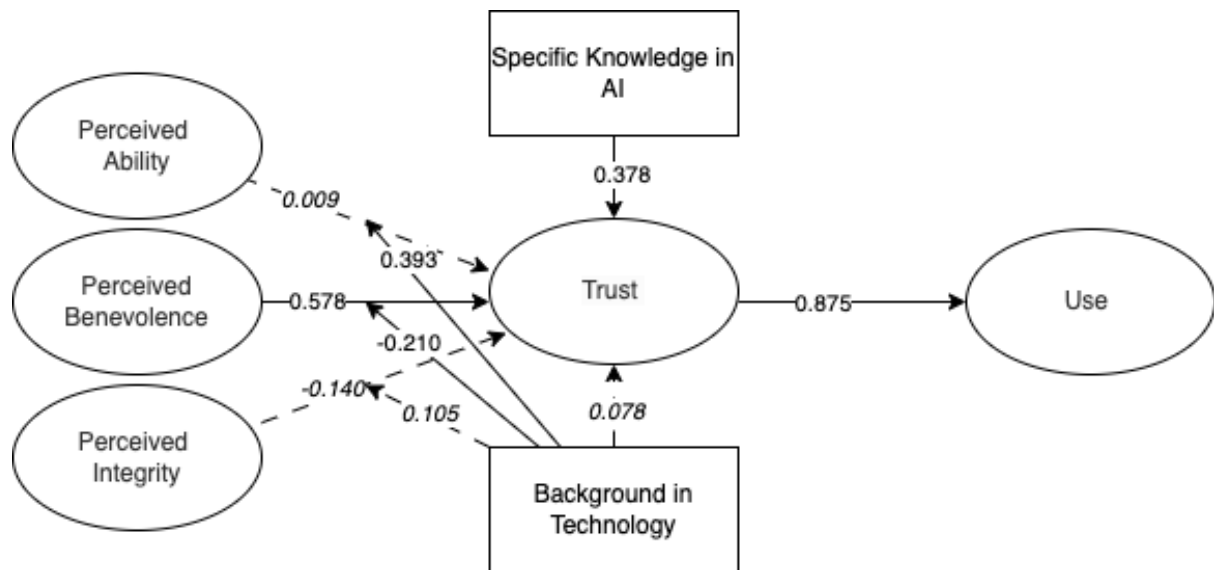
A significant path coefficient (β and p-value < 0.05 in Table 4) is one of the primary measures to evaluate the causal relationship in our structural model. Figure 2 summarises the hypotheses testing results. The findings of this study provide empirical support for the hypotheses H1-H9 as follows:

H1: Perceived Ability shows a small positive effect on Trust in GPT services ($\beta = 0.009$); however, this relationship is not statistically significant ($p > 0.05$). This suggests that, based on the current data,

Table 4

Regression coefficients

Outcome	Predictor	Std.Estimate β	Std.Err	p(<0.05)
Trust	Ability	0.009	0.255	0.925
	Benevolence	0.578	0.226	< .001
	Integrity	-0.140	0.215	0.444
	Spec.Kn.AI	0.378	0.081	< .001
	Bgr. In Tech.	0.078	0.088	0.349
	INT_A_TBgr	0.393	0.252	< .001
	INT_B_TBgr	-0.210	0.227	0.032
	INT_I_TBgr	0.105	0.207	0.566
Use	Trust	0.875	0.067	< .001

**Figure 2:** Structural model: Hypotheses testing

there is no strong evidence that users' perception of the system's capability directly influences their trust.

H2: Perceived Benevolence has a significant positive effect on Trust in GPT services ($\beta = 0.578$, $p < 0.001$). This finding suggests that potential subjectivity (bias) and confidentiality of user interactions with GPT services represent an important concern for the users, inhibiting trust.

H3: This study does not confirm a statistically significant effect of Perceived Integrity on trust in GPT services ($\beta = -0.140$, $p > 0.05$). The lack of statistical significance implies that perceived integrity does not play a decisive role in shaping user trust, or that other factors may have a stronger influence on trust formation in this context.

H4: The study confirms that Trust has significant positive effect on Use of GPT services ($\beta = 0.875$, $p < 0.001$). The high effect size suggests that trust is a primary determinant of adoption and continued use, indicating that users who perceive GPT services as reliable, capable, and well-intentioned are much more inclined to integrate them into their activities.

H5: The study does not confirm a statistically significant direct effect of User's background in technology on Trust ($\beta = 0.078$, $p > 0.05$). Nevertheless, we were able to confirm its moderating effect:

H6: User's background in technology significantly moderates the relationship between Perceived Ability and Trust ($INT_{A-TBgr} : \beta = 0.393$, $p < 0.001$). This suggests that the higher the user's level of education in technology, the stronger the impact of Perceived Ability on Trust in GPT services. Specifically, users with a stronger technological background are more likely to develop trust in GPT services when they perceive the system as capable.

H7: User's background in technology significantly moderates the relationship between Perceived Benevolence and Trust ($INT_{B-TBgr} : \beta = -0.210, p = 0.032$). Negative value of β suggests that the higher the user's level of education in technology, the weaker the impact of Perceived Benevolence on Trust in GPT services. Specifically, users with a stronger technological background are less likely to rely on the perceived benevolence (fairness, objectivity, confidentiality etc) of the exchanges with GPT services when forming trust.

H8: User's background in technology does not significantly influence the relationship between Perceived Integrity and Trust in GPT services. The non-significant moderating effect ($INT_{I-TBgr} : \beta = 0.105, p = 0.566$) suggests that factors such as honesty, transparency, and adherence to ethical principles impact trust in GPT services consistently across users, regardless of their technical background.

H9: Specific Knowledge in AI has a significant positive effect on Trust in GPT services ($\beta = 0.378, p < 0.001$), indicating that users with greater AI-related knowledge are more likely to trust these services. This finding suggests that familiarity with AI concepts, mechanisms, and limitations enhances users' confidence in GPT systems. Users who understand AI may better assess its capabilities, interpret its outputs more accurately, and manage their expectations, leading to increased trust.

4.3. Further Observations

We conducted the independent samples t-test in order to evaluate whether there is a significant difference between the means for Trust and Use of GPT services between two groups. The first group of respondents participated in the survey on February 2024 (Period 1) and the second group - on November 2024 (Period 2). The results are illustrated in Fig.3.

We did not find significant evidence of increase of Trust in GPT services over time, with the t-statistics $t = -1.817$ and $p = 0.072$ (conventional significance threshold - $p < 0.05$). In contrast, we found significant increase of self-reported Use of GPT services between February 2024 and November 2024, with the t-statistics $t = -3.876$ and $p < 0.001$. This suggests that factors other than Trust (e.g., habit formation, external influences, peer pressure, availability of new services etc.) may be driving the growing use and adoption of GPT services.

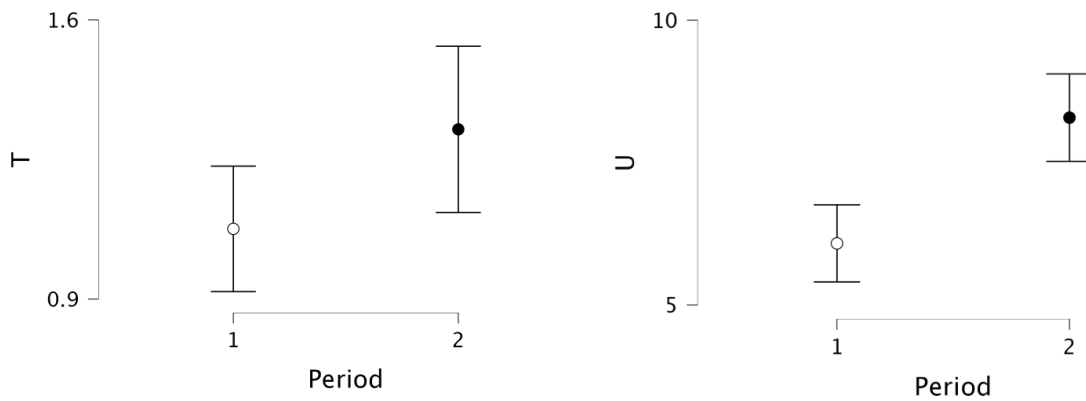


Figure 3: Independent samples t-test on Trust (T) and Use (U)

5. Conclusion

This study aimed to provide a deeper understanding of the social dimensions of trust in increasingly human-like AI technologies. We explored user trust in GPT services by examining social factors influencing its formation as defined by the Integrative Model of Organizational Trust [5]. The study empirically validated a proposed theoretical model through a survey of 124 participants. Data analysis was conducted using Structural Equation Modeling (SEM) in JASP.

While our study was not able to validate all the hypotheses confirming the Integrative Model of Organizational Trust, it confirmed that (social) trust has a significant positive effect on the use of GPT services (H4). In particular, our findings show that perceived benevolence of GPT services has a significant positive effect on trust in GPT services (H2), highlighting users' concerns about potential bias and the confidentiality of their interactions.

Perceived ability and perceived integrity showed small positive and negative effects respectively (H1, H3). However, based on our data, those effects were not statistically significant. The absence of significant positive effect of perceived integrity on trust in GPT services raises a question: To what extent do users consider transparency and explainability of results important when deciding to engage with a service? This question will be a subject of our further investigation.

The role of user background in technology was also explored. While it did not have a direct statistically significant effect on trust (H5), it did exhibit significant moderating effects. In particular, we were able to confirm that a higher level of education in technology strengthens the positive impact of perceived ability on trust (H6). Similarly, a higher level of education in technology weakens the positive impact of perceived benevolence on trust (H7). The impact of perceived integrity on trust was not significantly moderated by the user's technological background (H8).

We also confirmed that specific knowledge in AI has a significant positive effect on trust in GPT services, suggesting that familiarity with AI enhances user confidence (H9).

Finally, our study shown that, while the self-reported use of GPT services significantly increased between February and November 2024, there was no significant increase in trust during the same period. This implies that factors beyond trust might be driving the growing adoption of GPT services.

This study contributes to the literature by reaffirming the significant role of trust in user engagement with GPT services and by explicitly linking trust in GPT services to the established factors of social trust.

In our future research, we are going to further explore the factors driving adoption of GPT services and delve deeper into the interplay between social trust and technical understanding in the context of rapidly evolving AI technologies.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

- [1] G. Yenduri, M. Ramalingam, G. C. Selvi, Y. Supriya, G. Srivastava, GPT (generative pre-trained transformer)—a comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions, *IEEE Access* 12 (2024) 1–1. doi:10.1109/ACCESS.2024.3389497.
- [2] K. Church, Emerging trends: When can users trust gpt, and when should they intervene?, *Natural Language Engineering* (2024) 1–11.
- [3] J. Buchanan, W. Hickman, Do people trust humans more than chatgpt?, *Journal of Behavioral and Experimental Economics* 112 (2024) 102239. URL: <https://doi.org/10.1016/j.socrec.2024.102239>. doi:10.1016/j.socrec.2024.102239.
- [4] D. Gambetta, et al., Can we trust trust, *Trust: Making and breaking cooperative relations* 13 (2000) 213–237.
- [5] R. C. Mayer, J. H. Davis, F. D. Schoorman, An integrative model of organizational trust, *The Academy of Management Review* 20 (1995) 709–734. doi:10.2307/258792.
- [6] M. C. J. B. T. Mcknight, D. Harrison, P. F. Clay, Trust in a specific technology: An investigation of its components and measures, *ACM Transactions on management information systems* 2 (2011) 1–25.

- [7] T. Garry, T. Harwood, Trust and its predictors within a cyber-physical system context, *Journal of Services Marketing* 33 (2019) 407–428.
- [8] S. J. Russell, P. Norvig, *Artificial intelligence: a modern approach* (2016).
- [9] N. S. N. P. J. U. L. J. A. N. G. K. Vaswani, Ashish, I. Polosukhin, Attention is all you need, *Advances in neural information processing systems* 30 (2017).
- [10] X. Han, Z. Zhang, N. Ding, Y. Gu, X. Liu, Y. Huo, J. Qiu, Y. Yao, A. Zhang, L. Zhang, et al., Pre-trained models: Past, present and future, *AI Open* 2 (2021) 225–250.
- [11] E. Glikson, A. W. Woolley, Human trust in artificial intelligence: Review of empirical research, *Academy of Management Annals* 14 (2020) 627–660.
- [12] L. G. Zucker, Production of trust: Institutional sources of economic structure, 1840–1920., *Research in organizational behavior* (1986).
- [13] D. M. Rousseau, S. B. Sitkin, R. S. Burt, C. Camerer, Not so different after all: A cross-discipline view of trust, *Academy of management review* 23 (1998) 393–404.
- [14] A. H. H. H. A. W. Söllner, Matthias, J. M. Leimeister, Understanding the formation of trust in it artifacts (2012).
- [15] L. F. V. S. A. A. Y. A. R. B. S. G. e. a. Murtin, Fabrice, Trust and its determinants: Evidence from the Trustlab experiment (2008).
- [16] A. M. T. M. Jacovi, Alon, Y. Goldberg, Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in ai, *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (2021) 624–635.
- [17] I. Rychkova, M. Ghriba, Trustworthiness requirements in information systems design: Lessons learned from the blockchain community, *Complex Systems Informatics and Modeling Quarterly* (2023) 67–91.
- [18] F. D. Davis, Perceived usefulness, perceived ease of use, and user acceptance of information technology, *MIS quarterly* (1989) 319–340.
- [19] M. G. M. G. B. D. Venkatesh, Viswanath, F. D. Davis, User acceptance of information technology: Toward a unified view, *MIS quarterly* (2003) 425–478.
- [20] R. Kumar, S. Anu, Evaluating chatgpt adoption through the lens of the technology acceptance model: Perspectives from higher education, *International Journal of Technology and Learning in Digital Education* 15 (2024) 213–228. URL: <https://www.inderscienceonline.com/doi/10.1504/IJTLID.2024.140316>. doi:10.1504/IJTLID.2024.140316.
- [21] B. Shahzad, et al., Chatgpt awareness, acceptance, and adoption in higher education: The role of trust as a cornerstone, *International Journal of Educational Technology in Higher Education* 21 (2024) 47. URL: <https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-024-00478-x>. doi:10.1186/s41239-024-00478-x.
- [22] T. Choudhury, M. Shamszare, Investigating the impact of user trust on the adoption and use of chatgpt: Survey analysis, *Journal of Artificial Intelligence Research* 67 (2023) 1025–1043. URL: <https://pubmed.ncbi.nlm.nih.gov/37314848/>. doi:10.1016/j.artint.2023.103294.
- [23] S. Bhat, et al., Examining chatgpt adoption among educators in higher educational institutions using extended utaut model, *Journal of Information, Communication and Ethics in Society* 22 (2024) 78–95. URL: <https://www.emerald.com/insight/content/doi/10.1108/JICES-03-2024-0033/full/html>. doi:10.1108/JICES-03-2024-0033.
- [24] S. Lee, S. M. Jones-Jang, M. Chung, N. Kim, J. Choi, Who is using chatgpt and why? extending the unified theory of acceptance and use of technology (utaut) model, *Information Research* 29 (2024) 54–72. URL: <https://doi.org/10.47989/ir291647>. doi:10.47989/ir291647.
- [25] J. G. P. Yousafzai, Shumaila Y., G. R. Foxall, A proposed model of e-trust for electronic banking, *Technovation* 23 (2003) 847–860.
- [26] C. P. Pfleeger, S. L. Pfleeger, Analyzing computer security: A threat/vulnerability/countermeasure approach (2012).
- [27] B. Killinger, Integrity: Doing the right thing for the right reason (2010).
- [28] D. H. McKnight, N. L. Chervany, What trust means in e-commerce customer relationships: An interdisciplinary conceptual typology, *International journal of electronic commerce* 6 (2001) 35–59.

- [29] S. M. S. R. M. d. M. Linåker, Johan, M. Höst, Guidelines for conducting surveys in software engineering (2015).
- [30] JASP Team, JASP (Version 0.19.3) [Computer software], 2024. URL: <https://jasp-stats.org/>.
- [31] D. Russo, K.-J. Stol, PLS-SEM for software engineering research: An introduction and survey, ACM Computing Surveys 54 (2021) 1–38. URL: <https://doi.org/10.1145/3447580>. doi:10.1145/3447580.
- [32] E. E. Rigdon, M. Sarstedt, C. M. Ringle, On comparing results from CB-SEM and PLS-SEM: Five perspectives and five recommendations, Marketing: Zeitschrift für Forschung und Praxis (ZFP) 39 (2017) 4–16. URL: <https://rsw.beck.de/docs/librariesprovider3/default-document-library/10-15358-0344-1369-2017-3-4.pdf>. doi:10.15358/0344-1369-2017-3-4.