Bachelor thesis

in the study program **Business informatics**
at the **Hochschule Aalen**

User Acceptance of Explainable AI: A TAM-Based Investigation of Evaluation Models and Influencing Factors

Submitted by:
Leon Wyrembeck
Matriculation number: 82890

First supervisor: Prof. Dr. Theissler
Second supervisor: Mr. Gerschner

Submission date: October 29, 2025

Abstract

This bachelor's thesis investigates the acceptance of explainable artificial intelligence (XAI) from a user perspective using a combined technology acceptance model (TAM). While previous research has mainly focused on the technical side of XAI, less attention has been paid to the psychological factors that determine human behavior. To close this gap, the study combines Davis' original TAM with Shin's XAI-specific model and empirically evaluates their suitability using a survey with a credit decision scenario. The results show that the classic TAM factors have little influence on usage intention when it comes to XAI. Instead, other factors prove to have a stronger influence on usage intention. The results underline the need to supplement existing TAMs with specific XAI factors such as trust, fairness, and transparency in order to further advance research on XAI from a user perspective.

Contents

List of Figures

List of Tables

1 Introduction

In recent years, artificial intelligence (AI) has experienced massive growth. AI is now being used or advertised in almost every area of human life [58, 2]. Nevertheless, AI has brought with it challenges and obstacles that cannot be easily dismissed [2].

Above all, the interpretability of AI decisions is a major problem [40]. Due to the complexity of their structure, which today's AI models have, it is impossible for humans to understand how the AI arrives at its output and the associated solution [29]. This leads to people developing a rather negative attitude towards AI because they cannot understand how AI works. This creates resistance among users where none should actually exist.

AI was developed to facilitate or take over decisions in order to support humans. An example of this would be decision-making or the anticipation of possible future events, such as the maintenance of machines before they break down and lead to longer down times [48].
Nevertheless, as just mentioned, there is the problem of a possible aversion or even resistance to AI, which is triggered by its complexity. In AI science, this problem is referred to as the "black box problem" [57]. The problem describes the fact that AI models are so complex and non-linear in their decision-making that even if you can look inside the model and its structure, you cannot understand it [7].

Companies have recognized the importance of transparency and comprehensibility that their AI models must offer when interacting with people in order to be accepted by the general public. To counteract the problem of the so-called black box, there is therefore the research area of explainable AI (XAI) [18].
XAI is responsible for making the output of an AI model more user-friendly, i.e. more comprehensible and transparent [24]. However, for a XAI explanation to be accepted, it is necessary to know which influencing factors are the most important and largest from the user perspective. To design the explanation with the contents of the AI system so that people accept a new technology such as AI. Instead of having a negative attitude towards it.

In order to determine this acceptance of new technologies, the so-called Technical Acceptance Models (TAM) were developed. The best-known TAM is that of Davis, which was developed at the time to determine user acceptance of information technologies [15]. It is still widely used today, known worldwide and highly regarded. Such a TAM is well suited to finding out which factors are most important for increasing openness, comprehensibility and trust. Over the years, there have been repeated adaptations and spin-offs from Davis' original TAM [54, 52, 53, 47, 10]. In most cases, the Davis TAM has been adapted to suit a specific target group in conjunction with a new technology in the sector. And this is exactly what should happen with AI, especially in the XAI sector.
AI and XAI are no longer a minor topic for tech companies like Google or Meta. These companies are investing more and more in AI research and development. These two giant tech companies alone are investing billions respectively [50, 1]. In the meantime, they have also launched their own AI models on the market. For the XAI "layer" to function properly in the respective model and achieve the desired effect, the XAI "layer" must be well-tailored to the needs of the users in order to gain their trust.

There are many papers that deal with the technical aspects of XAI [18]. On the other hand, the part of research that deals with the influencing factors from the user's point of view is small, especially in connection with the well-known TAMs, which are mostly easy to use. This paper examines which TA models exist that can be applied to AI systems with XAI explanations. But it also addresses which factors in XAI explanations are most important for increasing user acceptance of AI systems.

At the beginning of this paper, the most important terms are explained and the current state of research is defined in more detail. This is intended to create a good basis, which will be important for the main part of the work later on. The following research questions are answered in this paper:

- Which Technology Acceptance Models are suitable for evaluating XAI explanations?
- Which factors, according to the TAM, influence the acceptance of AI-generated explanations by users?

- How influential are the respective factors in the acceptance of XAI explanations?

After establishing a solid foundation and knowledge about XAI and TAM, the first part of the main section focuses on identifying a suitable TAM that includes the most important influencing factors mentioned in XAI research. The selection of TAMs to be examined in more detail is based on research in the field of TAMs and XAI. The TAMs are then compared against each other, and the most suitable model is selected in order to develop and conduct the empirical study in the second part of the main section.
In this study, a survey is created in which the most important factors in the area of XAI user acceptance are examined by test subjects to determine how important they are from the user's point of view. The TAM model that performed best in the comparison in the previous chapter is used for this purpose.
At the end of this thesis, the results are presented, a conclusion is drawn, and it is shown how future work can build on this work and what the limitations of this work were.

2 Theoretical Background

2.1 Black-Box-Problem

The so-called "black box problem" is not to be understood literally. Rather, this term is intended to better illustrate the fact that AI models are so complicated due to their non-linear structure and complex nesting. Even if you could look into the model like a box, with all the parameters and weights, it would be too complex for the normal human mind to understand logically [20].

First of all, we need to clarify and differentiate between the terms machine learning (ML) and deep learning (DL). DL is a subfield of ML and both are subfields of AI. The main differences between ML and DL are that ML uses different algorithms, requires smaller amounts of data but therefore features often have to be determined manually, and ML is therefore suitable for simpler tasks. While DL consists of neural networks, it extracts features automatically, which in turn means that it requires large amounts of data, but is therefore also well suited to complex tasks such as image recognition [36]. From these differences you can already guess that DL models probably has more "black box problems" than ML models [21], since the complexity of the model architecture with non-linear structure is significantly higher. Many ML models are simpler and more understandable in their structure than DL models.
ML and DL models should therefore be viewed more as a family of AI methods. Be it in the form of Recurrent Neural Networks (RNN) which is used in DL. While a support vector machine or a decision engine is used for ML models [57, 36]. To imagine it better by a DL, you can think of it like a child growing up and learning with its parents. However, the parents determine how and what the child is allowed to do, but the child makes the more precise experiences with the respective situation itself and learns from them. It's the same with a DL program. The only difference is that the developers only specify certain starting parameter values, while the program itself determines these values through the learning algorithm and the learning environment, which is important for differentiating things. Because the developer doesn't directly specify the differences themselves, they can no longer directly understand the precise effect of individual parameters or input patterns on the output or specifically change them [57]. This creates the black box problem, which means that complex operations or decisions are output by the AI model without the user being able to properly understand them [26].

In practice, the problems in the medical field are particularly clearly visible. In medicine, ML and DL models are used with good success and are being used more and more [42]. Nevertheless, there are legal and ethnic problems that arise. Namely due to the complexity and the resulting lack of transparency. There are concerns about whether this technology is ethnically and culturally acceptable in medicine. In the medical field, AI models must be able to explain their decisions plausibly and logically in order to gain the trust of doctors, nurses and patients. AI models that cannot explain their decisions transparently. This can lead to problems with accountability and liability in the event of complications for the patient if the medical treatment is based on AI. Furthermore, it can complicate medical problems if the healthcare professionals themselves cannot accurately understand the output of the AI [34].

But it is not only in the medical sector that AI models are used for better decision-making. AI models are also used in the financial sector to determine, for example, the creditworthiness of a customer and whether or not they will receive a loan [45]. The customer data that the bank has is fed into the AI model and processed. The AI model uses the data, statistics and ML to calculate the so-called credit score. This is used to determine the credit risk of the respective customer with a high degree of accuracy. However, these models also lack transparency. It is often not possible to understand how the model came to the decision not to give a borrower the loan they wanted and what the exact reasons for this are [38]. This example also leads to skepticism in the financial sector towards AI models that are supposed to make decisions easier [45]. AI models that work with ML or DL lack mostly transparency. Both the end user and the developer themselves are often unable to understand exactly how decisions are made, as the models are too opaque and complex. This makes it difficult for a person to logically explain the decision or output of the model [26]. Sooner or later, this inevitably leads to mistrust and even resistance to AI on the part of humans. Although AI should help to support a person and make decisions easier or less difficult.

2.2 Aim and role of XAI

To tackle the black box problem, there is a field of AI research that is not new but has been gaining in importance for some time and rightly so. We are talking about XAI, which stands for "explainable AI". The XAI "layer" is responsible for explaining the output of the AI model to the user in an understandable way [18, 5]. This is important to minimize the black box problem, especially with DL models that are not linear and too complicated for the individual human. It is important that the AI justifies its decisions so that the respective person can better understand and cross-check the decision. After all, even AI models are not error-free and their data sets with which they are fed and trained can lead to decisions that are not correct. And this is precisely where XAI should help to better explain AI model decisions, but also to increase the transparency of the respective model so that the user develops a higher level of acceptance and thus trust and openness towards the AI [5]. The user can also better understand the model's decision and, if necessary, check it again [27]. XAI can be seen as an interface between the complex AI model and the user.

There are two main categories of XAI methods according to Arrieta et al. [5].

- Post-hoc explanations
- Intrinsic models

A post-hoc explanation is an explanation procedure in which the method is intended to explain the model's decision retrospectively [5]. This method is mainly used for models that work with a neural network but is also suitable for other models such as support vector machines (SVM). The best-known post-hoc methods are feature attribution methods such as SHAP or LIME [33, 43]. These methods show which input features had an influence on the prediction of the model and its decision. However, there are also text-based explanations which are becoming more and more popular, as they can be understood by everyone in their own language without any technical knowledge [19]. However, text-based explanations are not an automatic part of XAI methods.

Then there is the intrinsic category. These are easier to understand based on their structure, as they are not nonlinear like a neural network, which usually uses a post-hoc method. Instead, intrinsic models use decision trees or linear regressions [44]. These are therefore easier to understand, but are not used in today's practice for complex tasks, as they do not have the performance of a post-hoc method which, as already mentioned, should make the decision of a black box model such as a neural network understandable. It can therefore be said that post-hoc explanations are the most common, as post-hoc explanations are mostly used in DL, which is used to tackle complex tasks [5].

However, the focus of this work is on the user perspective and thus on increasing the comprehensibility, usefulness, fairness and transparency of the respective explanation types for the user. This increases trust and acceptance.

2.3 Relevant user factors in XAI research

This chapter focuses on the psychological factors that influence whether users accept and adopt the explanations generated by XAI "layer". In contrast to technical considerations, which focus on how it works technically, this chapter is concerned with how explanations are perceived, understood and evaluated by users, i.e. the psychological level of XAI.

It is important to understand how users interact with and react to AI model explanations. XAI serves as a kind of bridge between the AI model and the human with the aim of promoting trust, understanding and acceptance [1].

To achieve this, it is necessary to identify the key factors that influence users' perceptions of and behavior toward AI. Even if a system provides technically good explanations, it can still fail if these explanations do not match the expectations, level of understanding or emotional reactions of the users.

One of the most important factors in people is **trust**. When a person trusts someone or something, the person is more open and receptive to the other person or technology. A good example in normal life is trust in our parents or partners. When they give us advice in a life situation, we listen to it without skepticism and engage with it. This is because we have a strong trust in these people and do not mistrust them because we assume that they do not want anything bad for us. As a result, we are more open and receptive to their advice. This is why trust plays an essential role in XAI research, because if a person does not develop trust in an AI model, acceptance is minimized and the AI model is viewed critically. To create trust in AI models, it is important to understand how humans develop trust in any machine. Hoffmann et al. describe in their research that humans always have a mixture of justified and unjustified trust, as well as justified and unjustified distrust in machines [28]. These two opposing poles are constantly changing and are not constantly rising or falling. As with a friendship between people, trust can be built up for a long time but can quickly disappear again. It is the same with man and machine [28]. In other words, it is not a one-off process that lasts forever once achieved. In order to maintain people's trust in the long term, it will be necessary to constantly adapt the respective parameters to different influences, as is the case with a person-to-person friendship, which can be changed by a wide variety of influences over the years.

Another important influencing factor is **understanding**. As described in detail in the chapter "Black Box", it is important that people understand the explanation of the model. If people do not understand something, it is like "magic" to them. This is also how Herrera et al. explain it when they define Arthur C. Clarke's third law of engineering. It states: "Any sufficiently advanced technology is indistinguishable from magic" [27]. Herrera et al. explain that when we don't understand something, we associate it with magic [27]. And as was the case in the Middle Ages, if people cannot comprehend or understand something, they feel uncomfortable and reject it. Which is why this influencing factor should not be ignored and is one of the most important factors. Miller et al. describe it as follows: A good explanation is only good if it is tailored to the individual user and understandable to them [37]. Because every person is different and thinks differently, it is a very complex factor, but one that has an immense influence on the topic of XAI. Every person is different simply because of the level of education they have achieved. Another point is the prior knowledge in the respective topic. Does the user have prior knowledge of the question posed to the model yes or no? If yes, how much prior knowledge is there? If no, at what level of knowledge does the user understand the explanation of the model? So you can see that this influencing factor is not easy and cannot be clearly defined like some others.

Perceived usefulness (PU) is another important factor. This factor defines how useful a user finds a new technology and how much it helps them in their work or life. An example would be the cell phone. For everyone, the cell phone is an important tool and helps us in our lives. Whether it's the standard mobile calling or quickly surfing the internet for information that helps you save time. PU was first described and noted in Davis' TAM [15]. To this day, this factor is present in most TAMs.

Another important factor that needs to be considered in XAI explanations is **fairness**. Black-box models tend to make unfair decisions from the user's perspective. This can happen when race, age, or gender are taken into account in the decision. That can lead to unfair decisions that are perceived as discriminatory by the user. Therefore, according to Arrieta et al., fairness should be considered an important factor in XAI [5]. Fairness is a particularly important concept that is frequently mentioned in XAI and AI research [17]. This factor is crucial in order to prevent users from developing distrust and rejection toward the model. Fairness is especially important in sensitive domains such as the judicial system (e.g., convictions, sentencing) and the economy (e.g., loan approvals), where AI models are increasingly being used or will be used in the future [17, 46].

The last important influencing factor that is of great importance in XAI research is **satisfac-**

tion. Satisfaction is difficult to achieve because people are not automatically satisfied when they receive a correct explanation. Satisfaction must be seen as a variable that does not depend on the correctness of the output and is not easy to measure. A user can be dissatisfied even if they have understood an explanation, and vice versa. Herrera et al. describe this very well in their study: "Satisfaction reflects how well the explanation matches the user's expectations, even if it is technically correct." [27]. In other words, no matter how perfect the explanation is and how many influencing factors are taken into account. Whether the user is satisfied with the answer depends on his state of mind and his point of view. This also means that this influencing factor is almost impossible to examine clearly, as almost anything can influence the user's satisfaction.

According to the scientific literature, these factors are among the most important influencing factors in XAI research. According to current research, if these factors are taken into account when generating XAI explanations, there will be a positive influence on the adoption and use of AI models by a user. The factors listed are also important in the subsequent model selection and are tried to be taken into account as much as possible.

2.4 State of research & research gap

Research in the field of XAI is constantly being driven forward and is no longer a marginal topic in the field of AI. Nevertheless, the research has a larger share of technical research than from the psychological perspective, i.e. from the user's point of view [18]. More specifically, this means that more research and development is being done on the XAI methods themselves than on the actual core question of why the field of XAI exists in AI in the first place. Namely, to increase people's understanding, perception and acceptance of an AI model. Especially in connection with the well-known and respected TAM by Davis or any other TAMs, there is almost no research in contrast to technical research [15]. However, the TAMs has been used for decades in various technological developments to determine the respective user acceptance. Over time, there have been repeated improvements and extensions to Davis' original TAM. One of these examples is the Unified Theory of Acceptance and Use of Technology (UTAUT), but there are also other acceptance models [52]. In most cases, Davis' original TAM serves as the so-called basis on which a new model is developed. These models then differ in their areas of application and are often tailored to a technology.

In the AI technology area that uses XAI explanations, the most important factors that influence usage intention should be investigated [32]. Most research in this area, when conducting user studies, does not use a TAM model for their studies. This means that existing research cannot be easily compared with each other, as there are no uniform framework conditions. Furthermore, the problem is that the original TAM does not examine all influencing factors that are important in XAI research. Nevertheless, there are researches that have addressed this problem and developed TAMs based on Davis' original TAM. An example of this is Baroni et al., who build a model based on Davis' TAM and with additional constructs that examine the intention to use in car damage claims [4]. Another work comes from Shin, who also investigates the most important factors and their interrelationships that influence XAI from the user perspective. He also builds a model, but does not derive it from the TAM [47]. Another example is Panagoulias et al., who use TAM to develop a framework for personalized explanations in the medical field. This is referred to as PINXEL by Panagoulias et al [39]. However, it has only been tested and applied in the medical field, so it cannot be considered a general solution. Even though the medical sector is one of the areas that has a great need for solving the black box problem. It is crucial that the explanations provided and the associated decisions made by an AI model are comprehensible and acceptable to the user. Whether that is the patient or the attending physician [46].

However, there is also research that does not use Davis' TAM or another acceptance model based on the original TAM to investigate the acceptance of XAI explanations and their influencing factors. Examples include Herrera et al., who conducted a user study but did not use a TAM, instead measuring trust and satisfaction as individual variables [27]. Another example is the research by Westphal et al., who did not use the original TAM as the basis for their two studies on improving user perception and compliance in human-AI collaboration, although they briefly discuss the TAM in their paper [55].
It quickly becomes apparent that there is no uniform solution for evaluating XAI explanations in terms of their understanding, perception, and acceptance, nor for defining the most important influencing factors for users. This thesis contributes to addressing this gap. There is ongoing research and development into improved explanation methods and better model architectures. Despite this

immense growth in the industry, the user acceptance, which should be a key component, has not yet been sufficiently investigated [46]. However, this is incredibly important for understanding how users perceive XAI explanations in the first place and whether they understand and accept them correctly. These points are essential for creating a better relationship between AI and users so that users are willing to engage with this relatively new technology. This problem is also described by Miller in his work [37]. He states that many XAI studies do not use socially grounded explanations and should rely more on insights from the social and cognitive sciences [37]. The targeted TAM approach should help define a uniform standard as a kind of framework for future research that also deals with XAI from a user perspective. This should create a large and framework-based knowledge base, like a large knowledge database, where results can be compared with each other without much effort. The TAM subsequently selected, or the combination of models, which are best suited for XAI explanations, is intended to provide a kind of foundation. Like Davis' TAM, which has evolved over the years. As in all other research areas, new insights are constantly being added to XAI research.

2.5 Definition of "explainability" for AI

An important concept that needs to be clarified in order to avoid misunderstandings in the literature and in subsequent empirical study is the definition of "explainability" in AI models. Section 2.1 discussed the black box problem, which occurs particularly in deep learning models due to their nonlinear and complex structure. As a result, the decisions made by the model are difficult or impossible for the user to understand. But what exactly is meant by "explainability" in science when it comes to AI models?

As already discussed in detail in various chapters, there is the research field of XAI, which aims to make the output of an AI model more understandable, comprehensible, and explainable for the user. But how exactly is the term "explainability" defined in the field of AI and XAI? Arrieta et al. also addressed this question [5]. They concluded that "explainability" cannot be defined as a single word, but rather as the interplay of several concepts. To ensure consistent terminology and avoid confusion, Arrieta et al. differentiate between several terms and clearly differentiate them from one another [5]:

- Interpretability: Means that the model's decision should be understandable to a person without any tools, based solely on their own logic [5].
- Explainability: This refers to the behavior or output of the model that makes its functionality more understandable for a specific target group. This means that explainability depends on the level of knowledge of the respective user or group [5].
- Transparency: A model is only transparent if one has access to all of its mechanisms, from the architecture to the weights. And with this insight, the user should understand how the model works architecturally. Arrieta et al. therefore divide transparent models into three categories (simulatable models, decomposable models, and algorithmically transparent models), since a model can have different degrees of understandability [5].

An important point that Arrieta et al. emphasize in their work is that the explanation must be adapted to the target group [5]. What seems understandable to one person may not be so for another. Explainability therefore serves as the so-called interface between people and model, like a kind of interpreter.
This also means that explainability not only includes understanding the technical processes but also aims to make the decision made understandable, acceptable, and trustworthy for the user.

2.6 What are technology acceptance models?

As this work is strongly concerned with **technology acceptance models**, it is also important to create a basic understanding of the term TAM.
The first TAM was invented in 1989 by Fred Davis, who was an information scientist. The foundation for the first TAM is the socio-psychological model Theory of Reasoned Action (TRA), which was developed in 1975 by Fishbein and Ajzen [22]. It was invented to predict and measure the acceptance of a new technology or system. The aim was to find out why people accept or reject a new technology. And what their exact reasons are [15]. Davis uses two different constructs in his model that influence the third major construct:

- ***Perceived Usefulness (PU)***
  - The PU value indicates how helpful a person finds the technology and supports them in achieving their goals [15]. A simple example would be: That a businessman has an autonomous car and by not driving himself he has more time to complete his work tasks. In this example, the PU value increases.
- ***Perceived Ease of Use (PEOU)***
  - PEOU refers to how effortless a person thinks it is to learn the new technology [15]. As an example, let's take the same as PU with the businessman and the autonomous car. In that the businessman only has to sit in the car and tell the car where he wants to be taken, without having to take a course in theory and practice as with real driving. The PEOU value is high because you don't have to learn a lot of new things.

Both points together influence the "Behavioral Intention of Use". While PEOU has a certain influence on PU. This is because people automatically find something more useful if it is easy to learn [15].

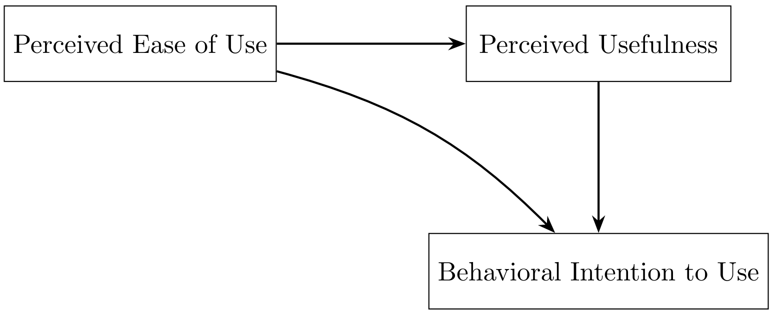


Figure 1: Core Constructs of the Technology Acceptance Model (Davis, 1989)

From these three constructs, Davis developed the original TAM, which is still highly regarded today and serves as the foundation for later TAMs. The model is simple and quickly understood by anyone, yet the added value it can offer for a new technology is immense. Every major technological innovation, whether online banking, smartphones, or social networks, was directly or indirectly researched using the Davis TAM after its adoption [30, 31, 16]. Many technologies have a history with the TAM or a modified form of the TAM.

Over the years, Davis' TAM was continually researched and developed further. This resulted in new TAMs, some of which were well-known, such as **TAM2**, **TAM3**, **UTAUT**, and others [54, 53, 52]. However, models have also been developed that are not intended for the general public, but rather to cover a specific area in greater detail, for example in the fields of medicine, law, or economics.
So it's easy to see why TAMs are relevant for XAI. The reason is simple: users who come into contact with AI models that provide XAI explanations should not only understand them, but also accept them and be open to the model's output. AI, and especially XAI, explanations in the general public are still in the midst of a revolution where AI is the pioneer. Over the past few years, there has been a tremendous increase in AI patents [51]. This is why it is important to measure influencing factors using acceptance models so that the models can be adapted and improved if necessary. Factors such as explainability, understanding, trust, fairness, interpretability, and transparency, which are strongly emphasized in XAI research, should not be missing in new TAM models tailored to XAI.

After discussing the key concepts and the relevance of this work, the next chapter introduces models that are either suitable for the field of XAI or well-established in TAM research. From these models, an analysis will be conducted to determine which one provides the best foundation for the study carried out later. The study will then be based on the results of the next chapter.

3 Literature research on TA-Models in the XAI context

3.1 Purpose and Approach of the Literature Review

This chapter aims to provide a good overview of the best TA models in the context of XAI and to compare them with each other. This chapter is essential for laying a solid foundation for the subsequent empirical study, which will serve to identify and interpret the most relevant influencing factors in XAI explanations. Therefore, this chapter aims to find suitable TA models or model components that can later be used in the empirical study for evaluation purposes.

3.2 Overview of Technology Acceptance Models

As mentioned in the previous chapter, there are a number of TA models that evolved from Davis original TAM and have also gained widespread acceptance. This overview analyzes the best-known models that have made a name for themselves in the field of technology acceptance models. However, models that are not as well known in this field of research but have a closer connection to XAI, which is essential for this work, are also considered. Some of these were developed exclusively for XAI, such as Shin's model. But as already mentioned, models such as the original TAM by Davis or UTAUT, which are well known in acceptance model research, are also considered [47, 15, 52].

3.2.1 Technology Acceptance Model by Davis

The TAM, which is considered the origin of all TAMs, was developed and published by Fred Davis in 1989. Davis' intention in developing the TAM was to build a framework that could be used to predict, explain, and measure the acceptance of new information technologies by users. This area of research is even more important today than it was back then, because the 21st century is full of super innovative technologies (smartphones, AI, blockchain, VR, AR, etc.). But these developments aren't just happening in industry. They're also happening in e-commerce application, in the financial sector with online banking, and in medicine [11, 30]. TAM and its constructs can be applied to virtually any new technology to determine user acceptance of, for example, enterprise software, healthcare systems, and much more, not just in the original use case of emails [11, 15]. Like most research, TAM is based on earlier research on which it builds. TAM is based on the Theory of Reasoned Action and uses its basic principles [22]. TAM aims to model the factors that determine user acceptance or rejection.

As already mentioned in earlier chapters, the TAM is based on two central constructs PU and PEOU:

- PU stands for Perceived Usefulness and describes the point of how a user thinks a new technology improves and or supports their work.
- PEOU Perceived Ease of Use and is intended to describe how much effort a user thinks it takes to learn the new technology

Both constructs influence the user's intention, which in turn influences actual use, known as behavioral intention to use. It is also important to note that PEOU has an indirect influence on PU. According to Davis, the reason for this is that people are more open to a technology if it is not too difficult to learn and use [15].

Davis TAM is highly relevant in the context of XAI, as the explanations generated represent a new level of interaction that is intended to increase user acceptance, which is also the intention of Davis TAM. It is important to understand how users assess the usefulness and relevance of XAI explanations in order to improve acceptance and trust in AI models.

However, Davis' original TAM has some issues with XAI because it doesn't take into account factors like trust, fairness, transparency, or understanding, which are really important in XAI research. Although Davis' original TAM provides a very good and solid foundation and has therefore been well known and appreciated for decades, it needs to be expanded or combined with other TAMs for use in XAI explanations in order to fully capture the user experience in AI explanations.

3.2.2 TAM2

The second model will examine more closely is the direct successor to the original TAM and is called TAM2. It was published in 2000 by Venkatesh and Davis and is intended as an extension of Davis's first model. TAM2 was primarily developed to improve explanatory power in relation to voluntary usage contexts and complex information systems. As already mentioned, TAM2 is an extension of Davis' first TAM and thus also builds on its PU and PEOU constructs. However, the extension of TAM2 is that the PU construct contains additional predictors that focus on social influence processes and cognitive processes [54].

The new constructs include:

- Subjective Norm (SN): the perceived social pressure to use the system.
- Image: the degree to which system use enhances one's status in a social group.
- Job Relevance: how relevant the system is to the user's work tasks.
- Output Quality: perceived quality of the system's performance.
- Result Demonstrability: how observable the benefits of system use are.

These extensions integrated into TAM2 are intended to help determine how users evaluate and perceive a particular technology in terms of its usefulness based on external influences, such as those from fellow human beings or colleagues [54].
In the context of XAI, TAM2 is definitely more relevant than the first TAM. After all, XAI are designed to explain AI decisions to users. Factors such as task relevance or output quality, where questions such as: "Does the explanation help me to complete my work?" or "Is the output/explanation clear and accurate?" play an important role in acceptance.

Compared to the original TAM, its extension TAM2 offers a framework in which the understanding of acceptance in professional, demanding or social environments is more precisely focused. This fits well with AI systems with XAI implementation because, as mentioned in previous chapters, they also have applications in healthcare and finance.
What should be mentioned is that TAM2 is not the latest extension but there is still a TAM3 model [53]. However, this was not considered in more detail as it focuses more on system training and usability, which is not relevant for XAI explanations.

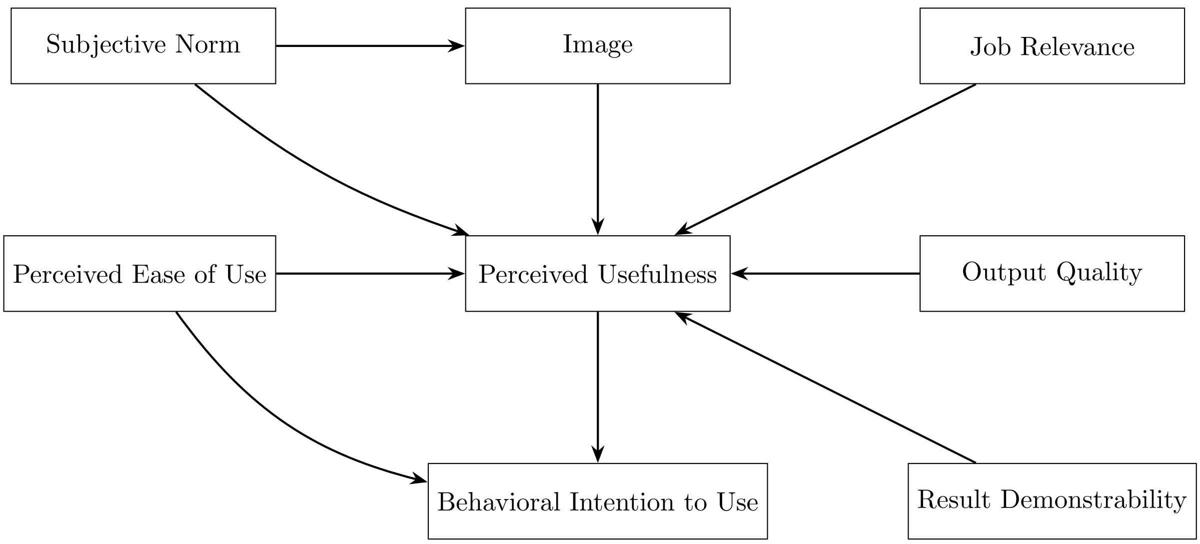


Figure 2: Extended Technology Acceptance Model 2 (TAM2)

3.2.3 UTAUT

The third model to be examined in more detail is the Unified Theory of Acceptance and Use of Technology (UTAUT), which, like TAM2, was co-developed by Venkatesh in 2003 and designed as an integration of various recognized acceptance models, including TAM, the Theory of Planned

Behavior (TPB), and the motivation model. The purpose of developing UTAUT was to unify acceptance research by combining the best empirical parts of well-known models in UTAUT. UTAUT comprises four important constructs for user intention and behavior [52].

- Performance Expectancy (PE): the degree to which using the system is perceived to improve job performance.
- Effort Expectancy (EE): the degree of ease associated with the use of the system.
- Social Influence (SI): the degree to which users perceive that important others believe they should use the new system.
- Facilitating Conditions (FC): the degree to which users believe that an organizational and technical infrastructure exists to support system use.

The special feature of UTAUT is that it can differentiate between certain groups such as gender, age, experience and voluntariness of use. This makes it very flexible to use in order to analyze certain groups more precisely [52].

In collaboration with XAI, UTAUT is not a bad model either; quite the contrary. UTAUT covers important aspects that are relevant to XAI. PE, for example, can refer to how useful a user finds the explanation. In other words, does the technology help the user perform tasks and support them in doing so? EE can describe how easy a user finds the explanation in terms of using the system. SI is also important because decisions in important areas such as health and finance are often made in teams, so it is important how the user thinks about how others think when using the system.

However, the UTAUT model also has a major disadvantage, namely its generality. This can lead to important influencing factors in the XAI context, such as trust, satisfaction with explanations, or perceived transparency, not being adequately taken into account. Nevertheless, UTAUT provides a solid basis for studies in complex environments such as the healthcare and financial sectors with multiple users.

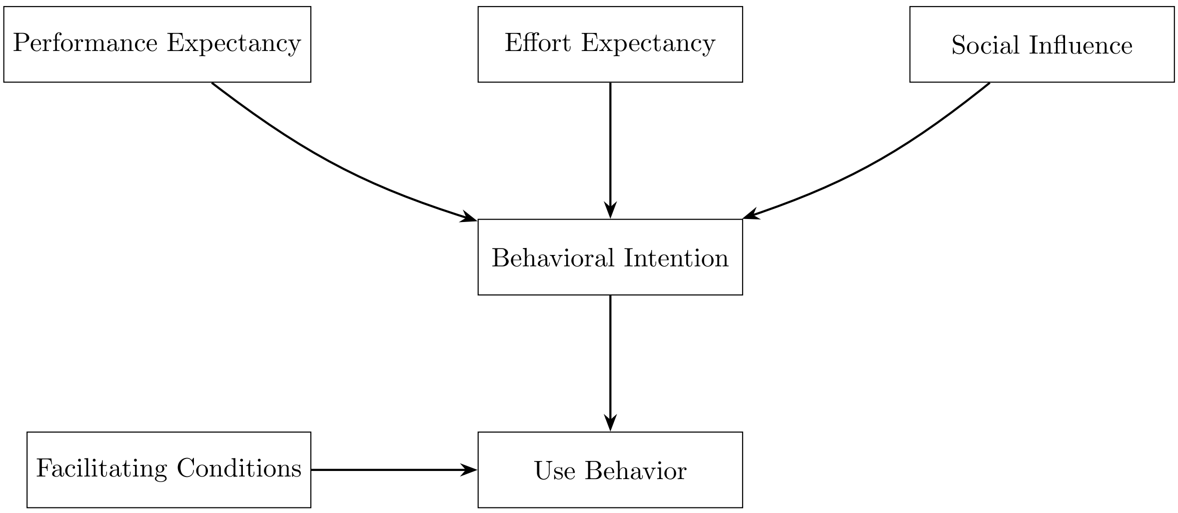


Figure 3: Unified Theory of Acceptance and Use of Technology (UTAUT)

3.2.4 Trust-enhanced TAM

The fourth model examined and analyzed in this paper is a model that is also a direct extension of Davis' original TAM, which was developed for mobile payments [14]. It is called Trust-enhanced TAM and was developed in 2003 by Dahlberg et al. [14]. As the name suggests, this model takes "trust" into account as an important influencing factor and adds it to the model.

In this model, the factor "trust" can directly influence the construct "behavioral intention to use," which defines the actual intention to use the new technology. The newly added construct

"trust" also indirectly influences the perceived usefulness construct, as trust in something reinforces the perception that the technology is useful. The expanded model was developed because Dahlberg et al. believe that even if the system is easy to learn and helps with task processing, users will only accept it if they truly trust the technology. But also because the model should be more realistic with the added construct [14].
As mentioned several times in this paper, "trust" is a very important factor in XAI research. This is because users must truly accept the decisions made by the model, even if they do not understand the model's decision-making process. For this reason, the trust-enhanced TAM was included in the model selection process.

The model provides a good basis for XAI because of the additional factor of "trust" that has been added to it. This makes the model relevant for application to XAI explanations. Nevertheless, the model has the problem that it only examines four constructs in detail (PU, PEOU, BI, and trust). However, as already mentioned, XAI research mentions even more constructs that are very relevant for the evaluation and acceptance of an XAI explanation from a user's perspective.

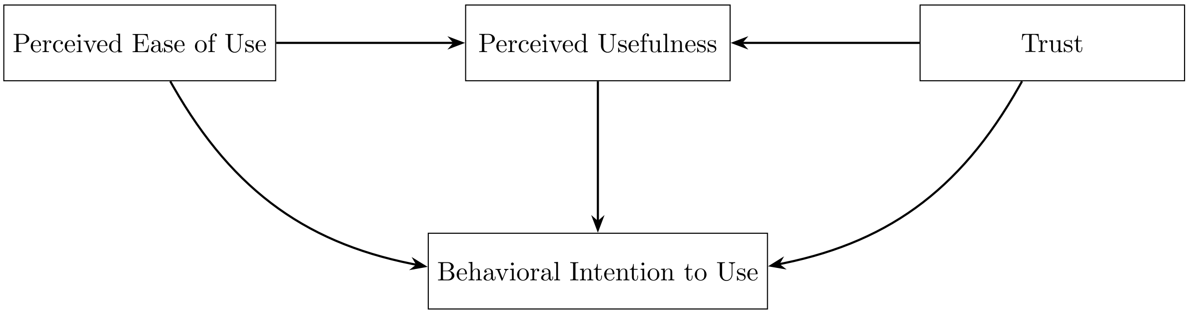


Figure 4: Trust-Enhanced Technology Acceptance Model

3.2.5 Shin's User Acceptance Model

The penultimate model examined is the model of Dongshee Shin 2021, which was developed specifically for XAI [47]. In contrast to classical TAM extensions, Shin's approach is not directly based on Davis's TAM, but on the Heuristic–Systematic Model (HSM) [9]. Nevertheless, Shin's goal was to develop a model that examines the most important influencing factors of XAI on user acceptance.

Shin includes the following most important aspects in his model:

- Explainability
- Causability
- Fairness, Accountability, and Transparency
- And Trust, which serves as the most important factor for acceptance

The model was validated in an empirical user study in connection with AI decisions [47]. Shin's model was therefore developed directly for XAI. It integrates the central influencing factors of XAI research and tries to measure users' perception regarding the explanatory quality of AI outputs. For this reason, Shin's model is very well suited for the subsequent study in this work.

However, it must be noted that Shin does not directly refer to TAM in his work. This means that Shin's model and its constructs can only be used in combination with another TA model in order to be relevant for this work. Also, Understanding is not directly shown in the model as its own construct, even though in Shin's work it is repeatedly mentioned [47]. In his discussion section, Shin points out that people want to understand how the AI system works, which makes this construct relevant [47].

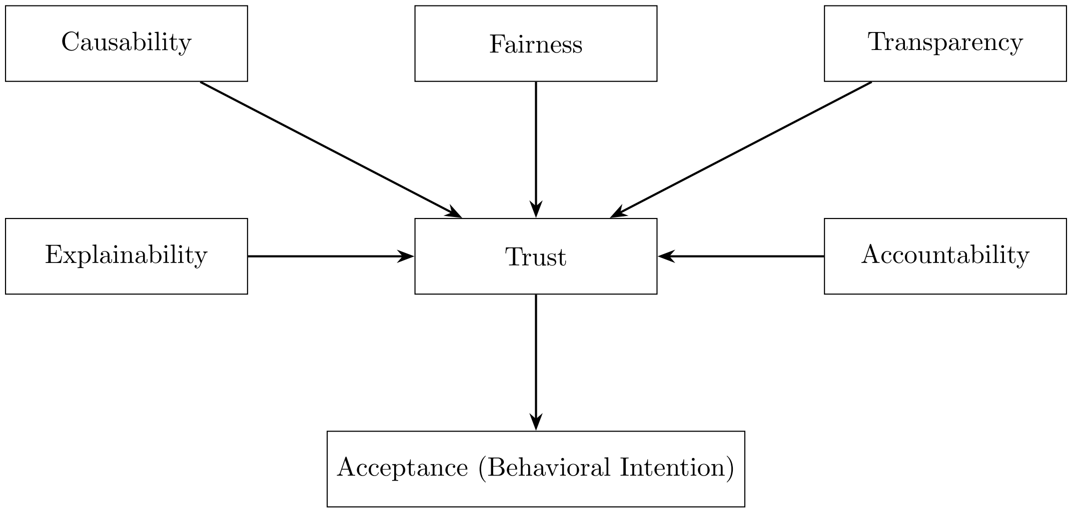


Figure 5: Shin's (2021) User Acceptance Model

3.2.6 PINXEL Framework

The final acceptance approach discussed is the PINXEL framework ("Personalized INterface for eXplainability basEd on Levels"), proposed by Panagoulias et al. (2024) for XAI in healthcare. Rather than a classical acceptance model, PINXEL represents a conceptual framework that builds on the TAM by Davis. It incorporates the core constructs of perceived usefulness and perceived ease of use, but adapts them through two key external variables: users AI literacy level and the abstraction level of the explanation. These dimensions allow the framework to dynamically tailor explanations to individual users, promoting more effective and trustworthy human AI interaction in complex domains such as medicine [39].

The most important points of the model are:

- **perceived usefulness** and **perceived ease of use** (from TAM)
- **Behavioral Intention** from TAM
- **AI literacy** and **the level of abstraction**: These allow for personalized explanations and thereby foster trust in the system

In an empirical study, PINXEL is validated in the medical field with medical professionals [39]. It must be said, however, that even though the Pinxel framework is based on TAM, it is only difficult to use because it is difficult to adapt to the needs of this work. This framework is also tailored to the medical field, which limits its general application. The focus is also more on factors that are considered important in the field of medicine.

3.3 Model selection

All six models differ from each other in their own way and each introduces additional constructs designed to cover specific areas of application. The main differences between the models lie in their areas of application, but above all in the influencing factors that they additionally consider and examine.

When comparing the models in terms of their relevance and suitability for the study in this article, it becomes clear that the model developed by Shin is best suited for the following study [47]. There are two main reasons for this: First, the model addresses key constructs such as trust, fairness, and transparency which are among the most important and frequently mentioned factors in XAI research and have already been addressed several times in this paper.
Second, Shin's model was explicitly developed and empirically validated in the context of explainable AI, which makes it particularly suitable for this study, as the model does not need to be significantly modified.

Other models such as UTAUT or TAM2 also have their very strong points, which is why the

model decision was not easy. The decisive factor is that Shin's model was developed for XAI and thus comes closest to the requirements. Nevertheless, Shin's model should be used in combination with Davis' original TAM, since Shin's model does not explicitly mention that it is based on TAM and he also does not classify his model as a TA model, even though his model comes close to a TA model. Another reason why the models should be combined is that Shin also does not examine the original constructs PU and PEOU, which are the main constructs in Davis' model. By combining the models, the respective constructs can be integrated, allowing an examination of, for example, whether perceived usefulness has an influence on trust or which constructs have the strongest impact on usage intention. Also, it is important to consider and examine the original constructs.
In addition, Davis specifies the exact scale values in his model, which are also recognized. Although these important parameters are missing in Shin's model.

Another reason is that Davis' TAM is recognized and well-known worldwide, so future readers are likely to be more open to a combined model than if only Shin's model is used, which is not yet widely used or has not yet been externally validated. In addition, other works that also deal with the topic of XAI explanations from a user perspective and are also based on Davis' TAM can be more easily compared with each other, as the scales correspond to the standard used by Davis. By combining the two models and evaluating them later, an attempt is also made to make the results more comparable with other results from other works that have used a different model for their research in this area.

XAI Influencing Factor	**TAM**	**TAM2**	**UTAUT**	**Trust-enh. TAM**	**Shin**	**PINXEL**
Perceived Usefulness (PU)	x	x	x	x		x
Perceived Ease of Use (PEOU)	x	x	x	x		x
Trust			x	x	x	x
Understanding						x
Fairness					x	
Transparency					x	
Behavioral Intention to Use	x	x	x	x	x	x

Table 1: Presence of the most important influencing factors of the models examined

Therefore, after the model investigation, the decision was made to combine Shin's model with Davis' TAM as well as possible. Furthermore, the construct Understanding, which is frequently mentioned in the XAI literature and was also noted in earlier chapters, is incorporated. This combined approach aims to identify and examine the most important needs of users when interacting with XAI explanations in AI models.

4 Methodology

4.1 Research Design

The following study uses a structured questionnaire that aims to empirically measure users' acceptance of XAI explanations and to identify the factors that are most important to users. The questionnaire is based both on the constructs from Shin's model and the construct understanding as well as on those from Davis. Davis' work also serves to define the framework for the study.

4.2 Model & Constructs Used

The constructs and the model itself consist of two models. Most of the constructs examined in this study originate from the Shin model, which was developed for XAI explanations and thus covers the most important constructs in XAI research. The second model used is Davis' original model. By adding the Davis model, the constructs PU and PEOU are also directly considered. Even though Shin attempts to replace PU and PEOU with the new constructs in his model and make it more applicable to XAI explanations, the addition of PU and PEOU as separate constructs allows us to examine whether connections arise between Shin's new constructs and Davis's old constructs, or whether new insights can be gained that are useful for this field of research.

The Davis model also serves to define the framework for the study. The aim is for the study to use the same scale values as most other research in this field. Since most research is based on Davis' model, it can be assumed that it also adopts Davis' scale values, which range from 1 to 7. This should make it easier to compare and examine research, thus providing a larger data set for future research that can potentially build on it.

This subsection lists all the constructs that are queried and examined in the study through specific questions:

Perceived Usefulness: This design comes from the original Davis TAM but is also built into the Shin model. "Perceived Usefulness" has the definition of measuring how useful the researched technology appears to the user in performing their tasks.

Perceived Ease of Use: This construct also originally comes from the Davis TAM and is the second important point there. This construct is also examined in Shin's model. The exact explanation of "Perceived Ease of Use" is already in the name, because it is about how easy the user finds it to learn the new technology.

Transparency: This construct was also taken from Shin's work. This construct is about finding out how transparent the respondents find the AI model and their answer. Transparency is seen as one of the most important influencing factors in XAI research.

Trust: This influencing factor was built into the Shin model. Everyone generally defines "trust" in the same way, i.e. the extent to which a user doubts or does not doubt the new technology in this study XAI explanations.

Understanding: As already mentioned, this construct is not found in either of the two models. Nevertheless, it was included because, on the one hand, Shin points out that Understanding is important, and on the other hand, the reviewed literature repeatedly highlights Understanding and regards it as an important factor in the field of XAI [8]. It measures whether the user has better understood how the AI arrives at its decision after an XAI explanation.

Fairness: This is the last newly added construct in Shin's model. It is about whether the decision of the model is comprehensible for the user and whether this decision is also fair.

Behavioral Intention to Use: The last structure was first used again in the original TAM and is also used in the Shin model. The aim here is to measure whether the respective user would use the new technology again, i.e. in this study the AI model with XAI technology.

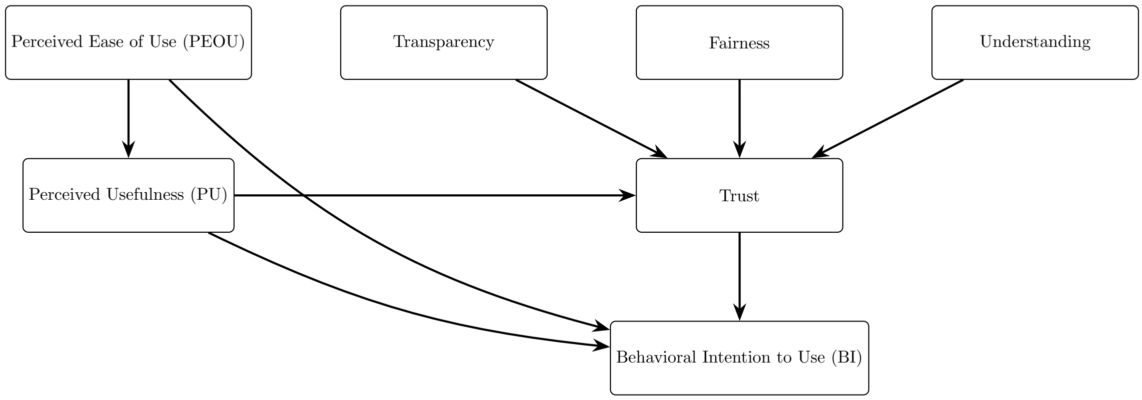


Figure 6: Theoretical research model (TAM + Shin Model)

4.3 Questionnaire Development

In the questionnaire for the following study in this paper, care was taken to ensure that the constructs and items were based on the work of Davis and Shin [15, 47], which had been selected during the previous model selection. The item formulations of Davis' constructs were adapted to the topic of XAI to make them applicable. An attempt was made to create consistency between Davis' original items and the items in the survey, even if this was not possible for every question (e.g., replacing "system" with "explanation").

In Shin's work, the problem was that the exact items used in the study were not clearly accessible, unlike in Davis' work, where the first two tables of his paper show the PU and PEOU items. Nevertheless, attempts were made to address the constructs through targeted items, thereby covering the constructs discussed in Shin's work. Only the construct Understanding and its questions are not based on either of the two models. However, the questions were formulated in a way that aims to cover the definition of Understanding in this research.

In total, seven constructs are covered in the survey, with four questions each, resulting in 28 questions. At the end of the survey, general questions were asked, such as age, gender, educational background, prior knowledge of AI, frequency of AI use in everyday life, etc.

The response format for the items related to the constructs is a 7-point scale, as also used by Davis in his study. This ranges from 1 = "Strongly disagree" to 7 = "Strongly agree." Before the questionnaire was published, it was tested with a small test group. The following table shows the exact constructs along with their respective items from the survey.

Construct	Questions
Perceived Usefulness (PU)	*1.* The explanation helps me better understand the AI's decision. *2.* The explanation helps me to better classify the AI result and draw useful conclusions from it. *3.* The explanation improves my ability to incorporate the decision into my considerations. *4.* The declaration helps me to understand the outcome of my application.
Perceived Ease of Use (PEOU)	*1.* The explanation is easy to understand. *2.* The explanation is linguistically clear and clearly structured. *3.* I was able to understand the explanation without having to read it several times. *4.* The declaration does not contain any unnecessarily complicated terms or formulations.
Trust	*1.* The explanation increases my trust in the AI system. *2.* I would rely on this system in a similar situation. *3.* The declaration conveys reliability in the decision-making process. *4.* The explanation gives me certainty that the decision is correct.
Transparency	*1.* I have the impression that the explanation has revealed to me the basis for the AI's decision. *2.* The explanation tells me how transparent the AI's decision-making process is. *3.* I do not feel that important information about the decision was hidden. *4.* The statement clearly shows on which information the decision is based.
Understanding	*1.* I understand why the AI made this decision. *2.* The explanation shows me how the most important decision factors are related to each other. *3.* The explanation helps me to understand the decision logic step by step. **4.** I could explain this decision to someone else.
Fairness	*1.* The decision seems fair to me based on the explanation. *2.* The statement shows that similar cases are treated equally. *3.* The statement gives the impression that the decision was objective. *4.* The explanation gives me the impression that the decision does not favor or disadvantage anyone.
Behavioral Intention to Use	*1.* I would use an AI system that provides such explanations. *2.* I would recommend such a system to others. *3.* I would be more willing to accept AI decisions if such explanations were always shown. *4.* I prefer AI systems that provide understandable reasons for their decisions.

Table 2: Constructs and corresponding questionnaire items

4.4 Hypotheses

Based on the theoretical model used, consisting of Davis' TAM and Shin's model, this chapter defines the hypotheses that will later be examined through the present study. These hypotheses aim to test the relationships between the constructs that influence the acceptance of XAI explanations.

Transparency → Trust

In AI research, transparency refers to how open and understandable an AI system is to the user [5]. This is precisely one of the main reasons for XAI. Research has shown that transparency is an important factor for trust, as it reduces uncertainty [3]. That is why Shin's model includes the construct of transparency, which influences the construct of trust [47].

Fairness → Trust

Everyone wants to be treated fairly by others, but also by algorithmic decisions such as those made by AI. This is why the concept of fairness is represented in the Shin model and influences trust. People can only trust a system if they feel that it treats them fairly from their point of view [47].

Understanding → Trust

In AI, understanding refers to the extent to which the user truly understands and processes the output, including an XAI explanation, from an AI system. Shin therefore emphasizes the importance of understanding in his discussion. Therefore, in this combined model, understanding is considered a additional construct, which has a positive influence on the construct of trust, which Shin considers [47].

PEOU → PU

The first hypothesis is that PEOU has an influencing effect on PU; both originate from Davis' TAM and are its two main constructs [15]. In the context of XAI, it is likely that an easily understandable explanation will be perceived as more useful, as previous research on the usability of explanations has also demonstrated [37].

PU → Trust

The PU construct, which originates from Davis' research, describes how users believe that a system in this work, the AI system with the XAI explanation, helps their performance [15]. Research such as that by Yudiantara et al. suggests that a higher PU increases trust in a system because users are quicker to trust a system when they perceive it as useful [56].

Trust → Intention to Use (BI)

Trust is considered one of the most important influencing factors in XAI research and other information systems [35]. Trust has a significant influence on the intention to use a system, as users only use systems that they consider trustworthy [35, 47].

PU → Intention to Use (BI)

In Davis' model, PU is described as one of the strongest influencing factors, which has a direct impact on the intention to use. According to Davis, this is because the more useful a user finds a system, the more likely they are to use it [15].

PEOU → Intention to Use (BI)

In Davis' original TAM, he also shows that PEOU has a direct influence on the intention to use [15]. Users are more likely to use a system if it is easy to understand and can be operated without much effort. Later studies and TAMs confirm this [54, 52].

Hypothesis	Statement
H1	Transparency has a positive effect on Trust in AI systems.
H2	Fairness has a positive effect on Trust in AI systems.
H3	Understanding has a positive effect on Trust in AI systems.
H4	Perceived Ease of Use (PEOU) has a positive effect on Perceived Usefulness (PU).
H5	Perceived Usefulness (PU) has a positive effect on Trust in AI systems.
H6	Trust has a positive effect on Behavioral Intention to Use (BI) AI systems.
H7	Perceived Usefulness (PU) has a positive effect on Behavioral Intention to Use (BI) AI systems.
H8	Perceived Ease of Use (PEOU) has a positive effect on Behavioral Intention to Use (BI) AI systems.

Table 3: Overview of the hypotheses

These hypotheses are tested using data collected from the survey using a structural equation model (PLS-SEM).

4.5 Sample & Data Collection

The questionnaire was sent out via various channels and made available to the public. On the one hand, it was forwarded to close contacts via Messengers with the request that they also forward it. An attempt was also made to increase the number of participants via Instagram in the form of a so-called story.

The aim is to get the number of participants above the 40 mark, which should also be maintained after the adjustment in order to be able to carry out meaningful results research.

The participants are not assigned to specific clusters such as age, academic degree or prior knowledge. Even if this data is also collected in the survey, the participants should first be seen as a whole unit without clustering. As already mentioned in other chapters, the survey aims to examine the most important general influencing factors in humans.

4.6 Survey structure

The survey is structured as follows: Before the first construct is introduced, the survey participant is placed in a fictional scenario. In this scenario, it is explained that an AI system is responsible for deciding whether the participant will receive a requested loan or not. The fictional AI then outputs the result that the loan has been denied.

Below this result, the fictional AI provides an explanation for why the loan application was rejected in other words, an XAI explanation. After the participant has read through the explanation, they proceed to the actual survey with the respective constructs and their items. These are to be answered based on the previously described scenario.

Between the different constructs, the AI decision and its explanation are shown again, so that participants can reread them if necessary.

This type of survey is intended to help participants gain a better understanding of XAI explanations, especially since the general public is usually unfamiliar with the term XAI. At the same time, it is meant to make answering the questions easier for participants. Furthermore, this approach aims to make the survey and its results as close as possible to a real-life scenario in order to increase the validity of the findings.

4.7 Data Analysis Methods

The data obtained from the study will be analyzed in accordance with scientific statistical standards. This is why three well-known methods of analysis are used in this work.

Descriptive statistics: This method of analysis is the one that should be familiar to everyone. In this method, the so-called mean, minimum, maximum, distribution, and much more are calculated. Descriptive statistics serve to provide an initial overview of the data and already offer important first insights into the collected data. This method is the fundamental first step in any quantitative analysis and is always useful.

Reliability analysis: The first metric is Cronbach's alpha which measures the internal consistency of a construct, for example, whether the four items assigned to the construct "transparency" actually measure the same thing [13].

Composite Reliability is also examined in the reliability analysis. This value also measures internal consistency but takes into account the individual item loadings, making it more accurate than Cronbach's alpha.

The final metric analyzed is the Average Variance Extracted (AVE), which measures convergent validity meaning it assesses how much of the variance in the items is explained by the construct. The values are evaluated according to the guidelines of Hair et al. [25]. This method is used in the evaluation because it is essential to check whether the items within each construct measure the same concept in order to maintain the validity of the survey.

Correlation analysis (Pearson/Spearman): This final method of analysis is used to measure linear relationships between the respective constructs, such as explainability and trust. It is well suited for formulating initial hypotheses and examining relationships.

Pearson's correlation is used for interval-scaled and normally distributed data and measures, for example, whether Y increases linearly when X increases. In contrast, Spearman's correlation is used for ordinal or non-normally distributed data and measures whether Y increases or decreases as X increases or decreases but not necessarily in a linear way, as with Pearson.

Therefore, both Pearson and Spearman are used in the later analysis, since the Likert scales used in the survey are formally considered ordinal, but are often treated as interval-scaled in practice [41, 12].

These three analysis methods are used to evaluate and analyze the data from the study and the results are discussed in detail in a later chapter.

5 Results

In the following chapter, the study conducted in this work is statistically analyzed and evaluated. The chapter is divided into several subchapters, each focusing on one of the analysis methods discussed in the previous chapter.

5.1 Descriptive statistics

As already mentioned in the previous chapter, descriptive statistics form the foundation of any statistical analysis, providing an initial overview of the study and its results. In addition to collecting general participant data such as age, gender, educational background, etc., the study also included seven constructs, each measured with four items. The constructs and their corresponding items can be found in Table 2 in Chapter 4.3.

In the present study, a total of 49 participants were recorded, and this number remained unchanged after data cleaning. Due to this relatively high number of participants, there was a diverse range of individuals who took part in the survey.

5.1.1 General questions

Age groups

Among other questions, participants were asked about their age. The following response options were available < 20, 20–30, 31–40, 41–50, 51–60, > 60.

In the study, two larger clusters emerged: participants aged 21–30, with a total of 18 people (32.7%), and the group aged 51–60, with 16 people (36.7%). Together, these two groups accounted for over 70% of the total participants.

The remaining percentages are distributed as follows:

- < 20 years: 3 participants (6.1%)
- 31–40 years: 5 participants (10.2%)
- 41–50 years: 3 participants (6.1%)
- > 60 years: 4 participants (8.2%)

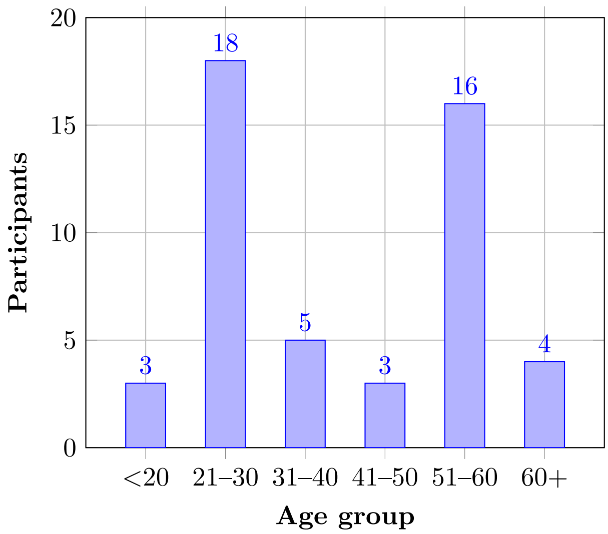


Figure 7: Distribution of participants by age group

Gender

The study also asked participants about their gender. The available options were: male, female, and no response. Among the 49 participants, there was a disproportionately high share of female participants, accounting for 63.3%, which corresponds to 31 individuals.
The second largest group was male, making up 32.7%, which corresponds to 16 participants. Two individuals chose not to disclose their gender, accounting for the remaining 4.1%.

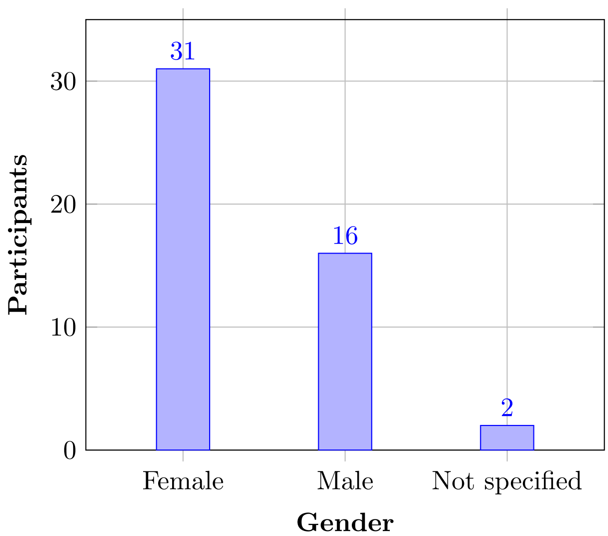


Figure 8: Distribution of participants by gender

Educational qualification

In the study, participants were also asked about their highest level of education in order to potentially draw insights from this data later on. However, unlike age and gender, no dominant groupings emerged in this category. Each educational group remained below 25%.

The largest group held a Master's degree or Diplom, with 11 participants, accounting for 22.4%. This was closely followed by those with a "Fachhochschulreife" (a German qualification similar to the general university entrance qualification), with 10 participants and a share of 20.4%.

The third-largest group consisted of participants with a general university entrance qualification ("Abitur"), totaling 9 participants (18.4%). The Bachelor's degree group followed closely with 8 participants (16.3%), the same number as those who had a "Realschulabschluss" (secondary school diploma).

The last two groups included participants with a "Hauptschulabschluss" (lower secondary school certificate) – one person (2%) – and two participants (4.1%) who indicated that they did not have any formal degree.
The option "Doctorate" was also available, but none of the participants selected it.

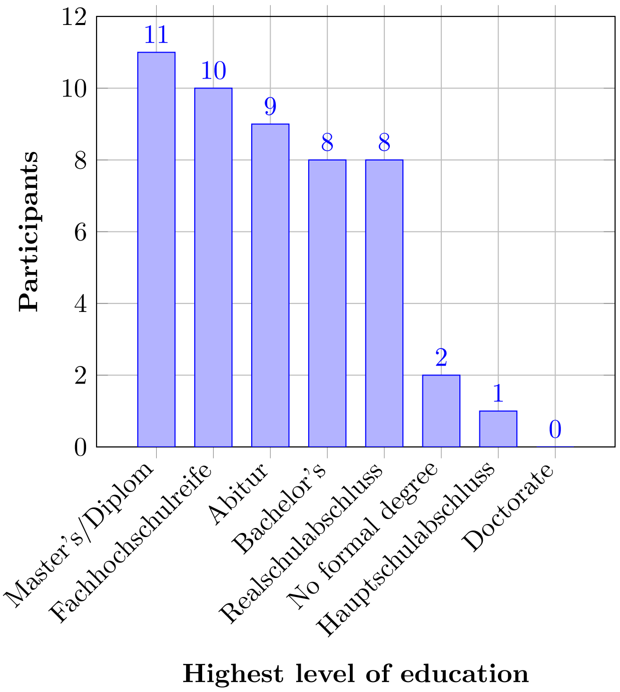


Figure 9: Distribution of participants by highest level of education

Use AI

The second-to-last question in the general section of the survey asked whether participants had ever used an AI system in any form. A total of 75.5% which corresponds to 37 participants answered "yes," indicating that they had used AI at some point in their lives. 22.4% (11 participants) answered "no," while only one person (2%) stated that they did not know whether they had ever used an AI system.

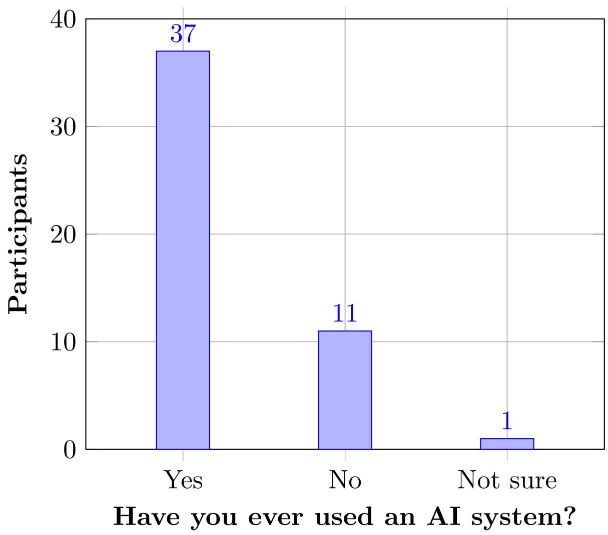


Figure 10: Responses to the question: "Have you ever used an AI system?"

AI use in everyday life

The final question, which was not directly related to the main survey content, asked how often participants use AI in their everyday lives. ChatGPT was mentioned as an example.

As with the question about educational background, none disproportionately large group emerged. The largest group selected the option “More than once a week,” with 13 participants, accounting for 26.5%. This was closely followed by the group who answered “Not at all,” with 12 participants (24.5%). The third-largest group was “Once a week,” chosen by 9 participants (18.4%). The group “Every day” was selected by 8 participants, making up 16.3%. The final option, and also the smallest group, was “Once a month,” chosen by 7 participants, which accounts for the remaining 14.3%.

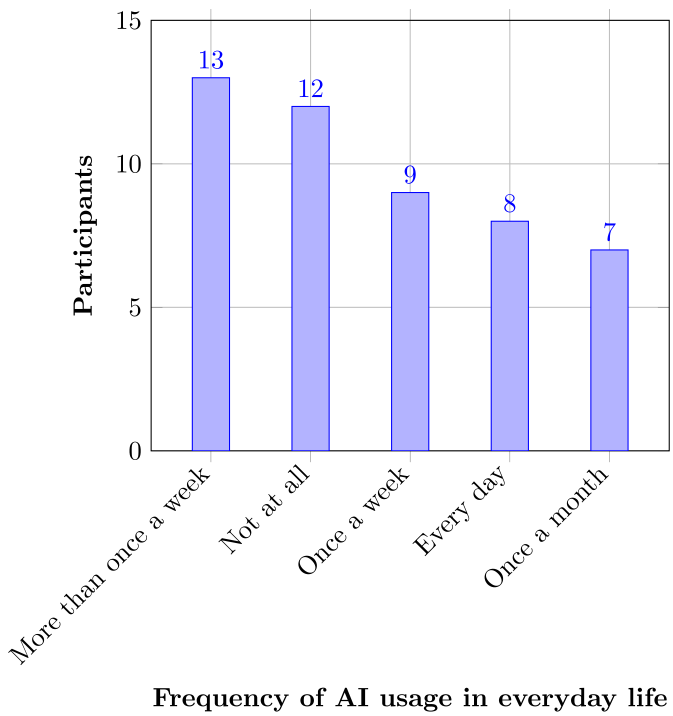


Figure 11: Responses to the question: ”How often do you use AI in your everyday life?”

5.1.2 Perceived usefulness (PU)

The first construct examined in the study is Perceived Usefulness (PU), which originates from the original TAM by Davis. Like all other constructs, it was measured using four items.
Participants were asked to answer the following questions for this construct in relation to the loan case scenario presented in the study, indicating how strongly they agreed or disagreed. As mentioned in a previous chapter, a 7-point Likert scale was used, ranging from 1 to 7.
The following four items should be answered:

- The explanation helps me better understand the AI’s decision
- The explanation helps me to better classify the AI result and draw useful conclusions from it
- The explanation improves my ability to incorporate the decision into my considerations
- The declaration helps me to understand the outcome of my application

The evaluation of the construct and its items produced the following results regarding the mean, standard deviation, minimum, and maximum values.

The **mean** was **5.837**, while the **standard deviation** for this construct was **1.196**. This indicates that participants perceived the usefulness of the AI explanation to be relatively high on average.
However, with a standard deviation of nearly 1.20, it should be noted that there is a certain level

of variation in how participants perceived the usefulness. This can also be clearly seen in the **minimum (1.75)** and **maximum (7)** values, which show that while some participants found the explanation to be extremely useful, others viewed it quite the opposite.

5.1.3 Perceived Ease of Use (PEOU)

The second construct examined, also using four items, was PEOU, which is the second key construct in Davis' model and, according to Davis, has an influence on PU this relationship will be discussed in more detail later.
The following four items for PEOU were presented to participants to answer:

- The explanation is easy to understand
- The explanation is linguistically clear and clearly structured
- I was able to understand the explanation without having to read it several times
- The declaration does not contain any unnecessarily complicated terms or formulations

For the second construct, PEOU, the analysis revealed a **mean** value of **6.24** with a **standard deviation** of **1.15**. This indicates that the understandability of the AI explanation was perceived as very high and, in fact, PEOU achieved the highest average score among all measured constructs.

However, due to the standard deviation (1.15), it must be noted that while the majority of participants agreed, there were still individual cases where participants had a completely different perception. This is clearly reflected as in the previous construct in the **minimum (1.25)** and **maximum (7)** values, with the low minimum particularly emphasizing that not all participants shared the same opinion on this construct.

5.1.4 Trust

The third construct that participants were asked to evaluate in the survey was Trust. This is the first construct derived from Shin's model and is also considered highly important in the XAI literature.
The four items presented to participants for this construct were as follows:

- The explanation increases my trust in the AI system.
- I would rely on this system in a similar situation.
- The declaration conveys reliability in the decision-making process.
- The explanation gives me certainty that the decision is correct.

The descriptive analysis of the Trust construct revealed a **mean** value of **4.15**, which is the lowest mean in the entire study. The **standard deviation** was **1.50**, the highest of all constructs measured.

This indicates that trust in the AI explanation was relatively low overall, and the high standard deviation shows that participants had widely differing opinions on the topic. As with the other constructs, the full range of values was present from a **minimum** of **1.00** to a **maximum** of **7.00** reflecting the diversity of responses and further supporting the interpretation of the mean and standard deviation.

5.1.5 Transparency

Transparency was the fourth construct examined in the study. This construct was also adopted from Shin's model, and transparency is likewise cited in the XAI literature as an important construct.
In the transparency construct, the following four questions were asked in relation to the case scenario:

- I have the impression that the explanation has revealed to me the basis for the AI's decision.
- The explanation tells me how transparent the AI's decision-making process is.

- I do not feel that important information about the decision was hidden.
- The statement clearly shows on which information the decision is based.

The descriptive analysis of the Transparency construct resulted in a **mean** value of **5.09** with a **standard deviation** of **1.36**. This mean value indicates that participants in the study generally rated the items in the transparency construct positively.

However, the standard deviation is relatively high, showing that participants did not share a clear consensus. While most participants tended to rate transparency in the upper range, there were also individual participants who opposed the majority opinion and viewed transparency rather negatively.
As with the other constructs, the range of values is quite evident here, with a **minimum** of **1.25** and a **maximum** of **7.00**. This again demonstrates that opinions ranged from very critical to fully agreeing.

5.1.6 Understanding

The third-to-last construct examined in the study is the Understanding construct. It is defined and mentioned in the XAI literature as an important construct, which is why it is included as a separate construct in this study.
The following questions were selected for the Understanding construct and were to be answered within the scenario of the case example:

- I understand why the AI made this decision.
- The explanation shows me how the most important decision factors are related to each other.
- The explanation helps me to understand the decision logic step by step.
- I could explain this decision to someone else.

The analysis of this construct resulted in a **mean** value of **5.56**, which is clearly above the neutral range. This indicates that, overall, participants considered the XAI explanation in the case example to be understandable.

However, the calculated **standard deviation** was **1.27**, which means that, as with the other constructs examined, the majority responded positively, but there were still some differences in responses regarding the Understanding construct.
As with the previous constructs, the **minimum (1.75)** and **maximum (7)** values show that almost every possible opinion was represented. It should be noted, however, that the minimum here, together with the PU minimum, represents the highest minimum in the study. Nevertheless, there were occasional difficulties in understanding the explanation, even though the overall results suggest a high level of understanding among the participant group.

5.1.7 Fairness

The second-to-last construct examined in the study is the Fairness construct. This is also derived from Shin's model, and it is a construct that repeatedly appears in the XAI literature in discussions about influencing factors. For this reason, it was included and examined in the study as well.
As with the other constructs, four items were used to assess fairness in the XAI explanation presented in the case example:

- The decision seems fair to me based on the explanation.
- The statement shows that similar cases are treated equally.
- The statement gives the impression that the decision was objective.
- The explanation gives me the impression that the decision does not favor or disadvantage anyone.

For the Fairness construct, the calculated **mean** was **5.29**, indicating that, as with other constructs, the mean is above the neutral range and therefore positive. The **standard deviation** for this construct is similar to that of the other constructs, at **1.34**, which also shows that the spread in this construct is moderate.
There is an overall positive tendency, but there are still some isolated critical evaluations of the Fairness construct. The **minimum (1.00)** and **maximum (7)** values clearly underline this, as all possible ranges are represented here.

5.1.8 Behavioral Intention to Use (BI)

The last construct examined and measured was BI, which, in turn, originates from the original TAM by Davis. This construct is essential in Davis' TAM and is also included in many later-developed TAMs, which is why it was examined and analyzed in this study as well.
Here, too, four items were used to cover the construct, which were as follows:

- I would use an AI system that provides such explanations.
- I would recommend such a system to others.
- I would be more willing to accept AI decisions if such explanations were always shown.
- I prefer AI systems that provide understandable reasons for their decisions.

The **mean** value for the final construct, BI, is **4.77**, which is still above the neutral value of 4 and indicates a slightly positive attitude from the participants' perspective. The **standard deviation** for this construct was **1.44**, the second highest value in the study, indicating a wide spread of responses.
The **minimum (1)** and **maximum (7)** values also show that there were participants with the most extreme opinions on both ends of the scale for this construct.

5.1.9 Descriptive Summary

The survey had an unbalanced gender distribution and did not cover the age range evenly, as can be seen from the tables. The majority of participants held a university degree and already had some form of experience with AI systems.

All constructs had a mean value above the neutral midpoint. The two constructs Perceived Ease of Use (6.24) and Perceived Usefulness (5.84) had the highest mean values in the study. On the other hand, the constructs Trust (4.15) and Behavioral Intention to Use (4.77) had the lowest and most moderate mean values, although they still indicated a slightly positive tendency.

For the minimum and maximum values, a certain degree of heterogeneity in participants' opinions can be observed across all constructs.

Construct	**Mean**	**SD**	**Min**	**Max**
Perceived Usefulness (PU)	5.837	1.196	1.75	7.00
Perceived Ease of Use (PEOU)	6.240	1.152	1.25	7.00
Trust	4.153	1.500	1.00	7.00
Transparency	5.087	1.360	1.25	7.00
Understanding	5.556	1.271	1.75	7.00
Fairness	5.286	1.341	1.00	7.00
Behavioral Intention to Use (BI)	4.770	1.443	1.00	7.00

Table 4: Descriptive statistics of the constructs

5.2 Reliability Analysis

The reliability analysis is important in this work to measure and assess internal consistency and validity. It ensures that the items used, depending on the construct being examined, measure the

same concept, thereby providing a solid foundation for future statistical analyses. Furthermore, it checks whether the items that do not originate directly from Shin or Davis also fit appropriately into the respective construct.
As already mentioned in the subchapter Data Analysis Methods, the values for Cronbach's alpha, Composite Reliability, and Average Variance Extracted were calculated and evaluated based on the guidelines of Hair et al. [25].
According to Hair et al., a **Cronbach's alpha** value of 0.70 or higher is considered acceptable in most social science research. Any value below 0.70 is regarded as insufficient [25].
For the **Composite Reliability** value, a range between 0.70 and 0.90 is considered good, while a value above 0.95 may indicate that the items are too redundant with each other [25].
For the final measured and examined value, the **AVE**, Hair et al. state that a value of 0.50 or higher indicates that the construct has good convergent validity, whereas a value below 0.50 means that too little of the variance in the items is explained by the construct [25].

Construct	Cronbach's Alpha	Composite Reliability	AVE
Fairness	0.857	0.902	0.698
Behavioral Intention to Use	0.827	0.884	0.666
Transparency	0.815	0.874	0.635
Understanding	0.895	0.923	0.752
Trust	0.913	0.939	0.793
Perceived Usefulness	0.894	0.927	0.760
Perceived Ease of Use	0.903	0.932	0.774

Table 5: Reliability and Validity Measures for all Constructs

The results clearly show that the Cronbach's alpha values for each individual construct are very good and well above the threshold of 0.70 specified by Hair et al. [25]. The construct with the highest Cronbach's alpha value is Trust, with 0.912, while Transparency has the lowest value at 0.815 which, as mentioned, is still well above the threshold.

For the Composite Reliability value, it can also be seen that the values are above the lower threshold of 0.70 but also exceed the optimal maximum value of 0.90. Only two constructs fall within the optimal range Behavioral Intention to Use and Transparency. All other values are above 0.90, with the Trust construct being close to the 0.95 range, which may indicate that the items are too redundant with each other in this construct. Nevertheless, all constructs remain below the 0.95 mark, which Hair et al. identify as the redundancy threshold [25].

For the AVE value, the results are again very good. All construct values clearly exceed the threshold of 0.50 set by Hair et al. [25]. Trust stands out in particular, with the highest value of 0.793, closely followed by Perceived Ease of Use with a value of 0.774. It can be seen that all constructs fall within a range of 0.635 to 0.793.

5.3 Correlation Analysis

Correlation analysis is an important analysis that examines the relationships between the constructs investigated in the study. It looks at whether the theoretical relationships also exist on an empirical level.

There are two different methods for correlation analysis, both of which are examined in this paper. This is because the Likert scales used in the study are ordinal but are often considered interval-scaled in research. Furthermore, this should show whether the patterns remain consistent. To cover both, Pearson correlation analysis and Spearman correlation analysis were therefore examined. The Pearson method is used for metric normally distributed data, while the Spearman method is used for ordinal and non-normally distributed data.

The results are evaluated according to Cohen et al. [12, 6]. There are three levels of how weak or strong the relationship between the constructs is. According to Cohen et al., the following levels exist:

- 0.10 – 0.29 weak correlation
- 0.30 – 0.49 moderate correlation
- 0.50 – 1.0 strong correlation

5.3.1 Pearson Analysis

In the Pearson analysis, which, as already mentioned, is responsible for interval-scaled normally distributed data, the results show that the construct “perceived comprehensibility” always has the weakest relationship with 5 out of 6 other constructs [41]. For the construct “trust,” the value is even the lowest (0.142) among all results. This is not the case for the relationship between the construct “perceived usefulness,” where the value is 0.421, which, according to Cohen et al., represents a medium correlation [12].

In the Pearson analysis, which, as already mentioned, is responsible for interval-scaled normally distributed data, the results show that the PEOU construct always has the weakest relationship with 5 of the 6 other constructs. For the construct trust, the value is even the lowest (0.142) in all results. This is not the case for the relationship between the construct PU, where the value is 0.479, which according to Cohen et al. represents a medium correlation [12].

The strongest correlation between two constructs in this analysis is between the constructs of intention to use and trust (0.702). The two other strongest correlations are between the constructs of understanding and transparency (0.647) and the constructs of understanding and fairness (0.634).

As already mentioned, the weakest correlations are between PEOU and trust (0.142). This is followed by the relationship between PEOU and intention to use (0.221) and PEOU and fairness (0.313). Despite being the third weakest relationship, the correlation between PEOU and fairness is still considered a moderate correlation according to Cohen et al.[12].
All other relationships between the constructs are in the moderate to strong correlation range.

	PU	PEOU	Trust	Transparency	Understanding	Fairness	Usage
PU	1	0.479	0.399	0.603	0.582	0.384	0.306
PEOU	0.479	1	0.142	0.411	0.364	0.313	0.221
Trust	0.399	0.142	1	0.624	0.502	0.616	0.702
Transparency	0.603	0.411	0.624	1	0.647	0.584	0.526
Understanding	0.582	0.364	0.502	0.647	1	0.634	0.439
Fairness	0.384	0.313	0.616	0.584	0.634	1	0.554
Usage	0.306	0.221	0.702	0.526	0.439	0.554	1

Table 6: Pearson correlation matrix for all constructs

5.3.2 Spearman Analysis

The second correlation analysis is Spearman’s, which was developed to examine ordinal and non-normally distributed data [49]. The Likert scales used, based on Davis’ model, are ordinal and should therefore be analyzed using Spearman’s method.
Nevertheless, in practice, Likert scales are often treated as interval-scaled. To ensure the robustness of this work and its results, both correlation analysis methods are therefore used, as already mentioned.
As with the Pearson analysis, the interpretation of the results in the Spearman analysis is also based on Cohen et al., using the same ranges [12].

In the Spearman analysis, the results showed that the highest correlation between two constructs was between Trust and Behavioral Intention to Use (0.641). Not far behind, and thus the second-highest correlation, was between Fairness and Understanding (0.629). Closely following this, with only a few hundredths less, was the relationship between Transparency and Understanding (0.622).

On the opposite side, PEOU appeared in each of the three weakest correlations. By far the lowest was the correlation between PEOU and Trust (0.149). The second-weakest correlation was between PEOU and Behavioral Intention to Use (0.244), which is about a tenth higher but still considered very weak according to Cohen's classification. The third-weakest relationship was between PEOU and Transparency (0.279). This shows that PEOU, which originates from Davis' TAM, does not correlate well with the constructs from Shin's model. All other relationships between constructs fall between the above-mentioned values and range from moderate to strong correlations.

	PU	PEOU	Trust	Transparency	Understanding	Fairness	Usage
PU	1	0.394	0.374	0.504	0.499	0.312	0.305
PEOU	0.394	1	0.149	0.279	0.450	0.324	0.244
Trust	0.374	0.149	1	0.620	0.478	0.514	0.641
Transparency	0.504	0.279	0.620	1	0.622	0.565	0.492
Understanding	0.499	0.450	0.478	0.622	1	0.629	0.406
Fairness	0.312	0.324	0.514	0.565	0.629	1	0.554
Usage	0.305	0.244	0.641	0.492	0.406	0.554	1

Table 7: Spearman correlation matrix for all constructs

5.4 Hypothesis testing

The hypotheses established in the Methodology chapter and represented in the structural model were tested using PLS-SEM. The path coefficients (β), t-values, and p-values were determined by bootstrapping with 5000 subsamples. Furthermore, the $\mathbf{R^2}$ values were taken into account to analyze the explanatory power of the model. Whether a hypothesis is supported or not is determined according to the standards of Hair et al. [25].
In order for a hypothesis according to Hair et al. to be supported. The p-value, which is calculated from the t-value and indicates the significance, is important [25]. The following distinction is made:

- p $<$ 0.05 $\rightarrow$ hypothesis supported (significant)
- p $>=$ 0.05 $\rightarrow$ hypothesis not supported (not significant)

Furthermore, the β value indicates the direction and strength of the correlation, while the t value is responsible for checking whether the β value is significant. For this, it must not be equal to 0. The following hypotheses were classified as follows based on the framework conditions:

- In the H1 hypothesis, the path coefficient was **positive and significant** (β = 0.393, t = 2.309, p = 0.021).
- In the H2 hypothesis, the path coefficient was **positive and significant** (β = 0.383, t = 2.425, p = 0.015).
- In the H3 hypothesis, the relationship was **not significant** (β = -0.005, t = 0.024, p = 0.981).
- In the H4 hypothesis, the path coefficient was **positive and significant** (β = 0.475, t = 2.438, p = 0.015).
- In the H5 hypothesis, the relationship was **not significant** (β = 0.018, t = 0.105, p = 0.916).
- In the H6 hypothesis, the path coefficient was strongly **positive and significant** (β = 0.692, t = 6.628, p = 0.000).
- In the H7 hypothesis, the relationship was **not significant** (β = -0.043, t = 0.370, p = 0.711).
- In the H8 hypothesis, the relationship was **not significant** (β = 0.154, t = 1.461, p = 0.144).

Hypothesis	Path	β	t-value	p-value	Supported
H1	Transparency $\rightarrow$ Trust	0.393	2.309	0.021	Yes
H2	Fairness $\rightarrow$ Trust	0.383	2.425	0.015	Yes
H3	Understanding $\rightarrow$ Trust	-0.005	0.024	0.981	No
H4	PEOU $\rightarrow$ PU	0.475	2.438	0.015	Yes
H5	PU $\rightarrow$ Trust	0.018	0.105	0.916	No
H6	Trust $\rightarrow$ Intention to Use	0.692	6.628	0.000	Yes
H7	PU $\rightarrow$ Intention to Use	-0.043	0.370	0.711	No
H8	PEOU $\rightarrow$ Intention to Use	0.154	1.461	0.144	No

Table 8: Results of hypothesis testing with path coefficients, t- and p-values

The so-called coefficient of determination (R^2) shows how a dependent construct (intention to use, trust, and PU) is explained by the independent constructs (PEOU, PU, transparency, trust, fairness, and understanding). PU and trust are both dependent and independent constructs.

The PLS-SEM analysis showed that 50.6% of the variance in intention to use is explained by the predictors, primarily trust, PU, and PEOU.
In the case of the construct trust, 48.5% of the variance is explained by the constructs fairness, transparency, understanding, and PU.
The last dependent construct is PU, which is explained by only 22.6% of the variance of PEOU.

Construct	R^2
Usage Intention	0.506
Trust	0.485
Perceived Usefulness	0.226

Table 9: R^2 values for endogenous constructs

According to Hair et al., none of the three values are in the high range, which starts at 75%. Nevertheless, intention to use and trust are in the middle range, while PU is in the lower range, which goes down to 25% [25].

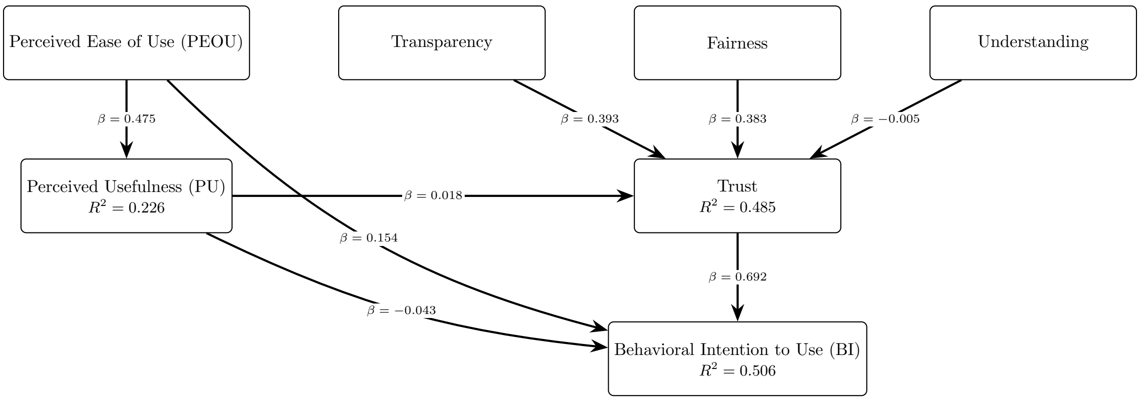


Figure 12: Empirical PLS-SEM results with standardized path coefficients and R^2

6 Discussion

This chapter discusses the results obtained from the study and compares them with existing results and findings from other research. This should lead to possible integrations into later practical applications as well as in theory. Furthermore, it examines whether the research in this work closes a gap in XAI research and/or helps to close one. It also critically reflects on what could have been done better in this research in retrospect and what future research could look like.

6.1 Summary of main results

The results chapter documented values and findings in detail. Now it is time to classify and discuss them. First of all, the descriptive analysis is classified with its values.

As already noted in the results section, the study has two large age groups. These consist of young people aged 21-30, i.e. people who grew up with technologies such as smartphones, the internet and much more. And a group that will soon be retiring and did not grow up with the internet, but rather in a time when telephone booths were still in use. This means that this study has two extreme groups that differ greatly from each other in terms of technology. In terms of gender, there is a disproportionately large group of women. In terms of educational attainment, most participants have at least a high school diploma. In the study, more than three-quarters of participants said they had used an AI system of some kind, which is not bad considering that half of the participants are over 50. Nevertheless, the last question shows that even though most have used an AI system before, a certain proportion did so rather involuntarily, as almost a quarter stated that they do not use AI in their everyday lives at all. Even though the largest group in the question was the “more than once a week” group.

However, the actual aim of this study was to examine the previously defined constructs according to their importance in the case of XAI explanations. The mean value was calculated for each construct, which provides initial information about how strongly the participants find the respective construct in relation to the XAI explanation, which was included in the study as a case study. The mean values of the constructs show that PU and PEOU, i.e., the constructs from Davis' original TAM, perform best, while the construct trust, which will later become very important in the hypothesis testing, has the weakest mean value together with the construct intention to use. This may suggest that, despite classifying the explanation as useful and easy to understand, the participants still have a certain amount of reservations and do not trust the explanation. Nevertheless, all mean values are above average and thus show a slight to strong positive trend, depending on the construct.

Compared to Shin's mean values, the mean values of the constructs in this study (4.1–6.2) are in some cases significantly higher than those reported by Shin (around 3.6–4.6). This may indicate that the participants in this study are more positive about AI explanations than those in Shin's study. There may be various reasons for this, including differences in sample characteristics such as age or gender. Nevertheless, Shin's values are compared with the values from this study, as Shin's work is strongly represented in this study through his model and the constructs are addressed [47]. In terms of reliability values, the values from Shin's study and those from this study are relatively identical. This shows that both studies have robust measurements. When comparing the Cronbach alpha values, it can be seen that the values differ only slightly in the lower range and are even identical in the upper range. Shin's values range from 0.76 to 0.91, while this study's values range from 0.81 to 0.91 [47]. The similarity of the Cronbach alpha values between these two studies indicates clear robustness, even though the contexts in which the constructs were examined differ between the two studies.

Another important analysis that was conducted using the survey data was the correlation analysis. Two types of correlations were calculated: Pearson correlations and Spearman correlations. The reason for this was that the data obtained using the Likert scales in the survey are considered ordinal and not normally distributed, but in practice, Likert scales are often regarded as interval scaled. To avoid problems with the results, both analyses were used to test initial correlations between the constructs in the assembled model. Nevertheless, it must be said that the correlation analysis should be seen more as a prediction of correlations and their strengths between the constructs and does not replace or anticipate the subsequent hypothesis testing. Therefore, the results should be treated with a certain degree of caution, but can still be seen as initial trends in the subsequent hypothesis testing.

The two correlation analyses showed that, regardless of whether Spearman or Pearson was used, the values did not differ greatly [41, 49]. Rather, most model relationships are very similar and sometimes differ by only a few thousandths. The strongest correlation is between trust and intention to use. Here, the Pearson value was 0.702 and the Spearman value was 0.641. This indicates a very strong relationship, even in the possible later hypothesis between these two constructs. This assumption would also be consistent with the work of McKnight et al. and Gefen et al., both of

whom emphasize trust as an important driver for technologies [35, 23]. This assumption would also support Shin, who also sees trust as an important construct in XAI research in his work and model [47].
The relationships between transparency and fairness to the trust construct also show strong values, both for Pearson (0.624 / 0.616) and Spearman (0.620 / 0.514). Only the correlation between the construct of understanding and trust is weaker than the other two relationships. Here, the Pearson value is 0.502 and the Spearman value is 0.478. Nevertheless, this relationship, like the other two, shows a good correlation. Here, too, these results support the general literature that understanding is an important factor [8]. It also shows that there are possible differences between the individual constructs and their importance for the acceptance of XAI explanations. This is also one of the research questions that investigates which factors influence the acceptance of XAI explanations.

However, the model contained in this study does not only include the Shin model [47]. Davis' original model was also incorporated, and here too, correlations were measured between the standard relationships, as in Davis' original TAM, between PEOU and PU, and between these two and the construct Intention to use [15]. This reveals an initial difference that does not entirely correspond with Davis' work. Namely, the correlations between PEOU and intention to use are rather weak (Pearson: 0.221 / Spearman: 0.244). However, the correlation values between PU and Intention to Use are also just above the weak correlation mark according to Cohen et al., where Pearson's value is 0.306 and Spearman's is 0.305 [12]. However, this contradicts Davis' research on the influence of PEOU and PU on the construct "behavioral intention to use," which is defined in this study only as "intention to use" [15]. Only the relationship between PEOU and PU shows stronger values in the moderate to almost strong range. For Pearson, the value here is 0.479 and for Spearman 0.394. These values initially indicate that the pure Davis model may not be very suitable for researching XAI influencing factors. One last relationship that still exists is the relationship between PU and trust, where the Pearson value is 0.399 and the Spearman value is 0.374. This means that the correlation is in the moderate range, which does not give a clear indication of whether linking these two models, as was done in this study, was meaningful or not. One of the research questions was which TAMs are suitable for evaluating XAI decisions. Correlation analysis reveals a slight tendency that the Shin model and its constructs may be more suitable than Davis' original TAM and its constructs.
Until the upcoming hypothesis tests, however, as already mentioned, these correlation analyses should not be given too much weight. Nevertheless, as already mentioned, they can provide initial indications of how the later hypotheses will develop.

Now the hypotheses and their results are tested and verified using structural equation modeling with the SmartPLS software. The evaluation of whether a hypothesis is supported or not is determined by several values. The following values were used to examine the hypotheses: path coefficients, t-values, and p-values from the bootstrapping procedure, which was performed with 5000 subsamples. Furthermore, the R^2 values were used to demonstrate the quality of the model explanations once again. The following table shows all hypotheses again and whether they are supported or not. To see the hypotheses with their respective values, please refer to Table 8, which can be found in the results chapter.

Hypothesis	Result
H1: Transparency → Trust	Supported
H2: Fairness → Trust	Supported
H3: Understanding → Trust	Not supported
H4: PEOU → PU	Supported
H5: PU → Trust	Not supported
H6: Trust → Intention to Use	Supported
H7: PU → Intention to Use	Not supported
H8: PEOU → Intention to Use	Not supported

Table 10: Summary of hypothesis testing results

The first hypothesis (H1), which states that transparency influences trust, was proven to be true after testing. This also reflects the opinion of Anjomshoae et al., who state in their work that transparent decisions made by AI systems increase user trust [3]. This confirmed hypothesis shows once again how important it is for XAI decisions to be transparent and comprehensible to the user in order to increase the user's trust in the AI system.

The second hypothesis (H2), which assumes that fairness influences trust, can also be confirmed. Here, too, the result is consistent with earlier work, such as Shin [47]. Both state in their work that fairness is a prerequisite for trust in technology. In the field of XAI, this means that users consider a fair and unbiased XAI explanation to be important in order to trust an AI system.

The first hypothesis that is not supported is the third hypothesis (H3). The hypothesis was that understanding, like the two previous hypotheses, also influences trust. According to this study, however, this contradicts Shin's assumption that people want to understand how the results are generated and that this has an influence on trust [47]. This is one of the reasons why the hypothesis could not be supported in this study. It could be that the case study used in this study (credit approval) was not particularly suitable. Does this mean that participants in this study tend to see fairness and transparency as more important factors influencing trust than the comprehensibility of the explanation? This could mean that when it comes to important decisions such as granting loans, comprehensibility does not play as big a role in increasing trust as other factors.

The fourth hypothesis examined was derived from the Davis model and investigated whether PEOU has a significant effect on PU [15]. This is supported by the study and its evaluation. As already mentioned, this hypothesis is in line with Davis' work and TAM, which states that usability has an influence on usefulness [15]. In the context of XAI, this means that simple interaction with the AI system and its explanations increases the usefulness of AI systems for the user. This also means that user-friendly explanations increase the perceived usefulness of AI systems.

The fifth hypothesis (H5) that was examined is the hypothesis that PU has an influence on trust. This is the relationship between Davis' TAM and the Shin model. However, it must be said that this hypothesis is not supported. This deviates from Davis' research, but also from other research which states that usefulness increases trust among users [56]. In the context of XAI, this means that even if an XAI explanation is perceived as useful, it does not increase the user's trust in the AI system, and fairness and transparency are the most important factors influencing trust.

The sixth hypothesis (H6) was supported by this study. This hypothesis assumed that trust influences the intention to use. This assumption is confirmed and also reflects earlier work, such as that of McKnight et al. and Gefen et al., both of whom assume in their work that trust is one of the most important factors for technology acceptance [35, 23]. From an XAI perspective, it is confirmed that trust is an important influencing factor in persuading users to accept an AI-generated explanation.

The penultimate hypothesis (H7) stated that PU significantly increases the intention to use. However, this hypothesis is not supported by the study and its results. This result is very surprising, as Davis's work, from which this relationship originates, states that TAM has a significant correlation here [15]. This lack of a strong correlation may indicate that in the field of XAI, the influencing factor of trust plays a much greater and more essential role in the intention to use than perceived usefulness.

The last hypothesis (H8) that was tested is that PEOU has a strong influence on intention to use. However, this hypothesis is also not supported by this research, which again contradicts Davis' TAM, in which usability has a direct influence on the intention to use [15]. This lack of support can possibly be explained by the fact that users already assume a certain basic level of usability when it comes to XAI, and therefore further improvements in this influencing factor do not have a strong effect on the intention to use.

At the end of the hypothesis testing, it turns out that only four of the eight hypotheses are supported. It can also be seen that the constructs of the classic TAM perform worst when it comes to the reference to the intention to use in the XAI area, which indicates that these relationships are not really applicable there. It should also be emphasized that the influencing factor of trust is the most important for the intention to use, but that trust also depends on other factors such as fairness and transparency.

6.2 Theoretical implications

The study and its findings have revealed several theoretical implications in this work. The first point is that the results of this study show that Davis's pure TAM cannot be applied to XAI and that several additional factors are needed for it to be applicable to XAI. While PU and PEOU are direct drivers of behavioral intention in Davis' model, this study showed that these factors are weak and that other factors have a much greater influence on usage intention. Constructs such as trust, fairness, and transparency play a much more important role in XAI. Understanding also plays a greater role than the constructs of the original TAM.

The second implication emphasizes that, with XAI, the factor of trust has a very large influence on users' intention to use it. Trust proved to be the strongest factor for the intention to use, and this relationship was also the strongest in the overall model. Although trust itself is primarily influenced by the constructs of fairness and transparency. This is also supported by Shin's findings, in which fairness and transparency also have a significant influence on trust.

Thirdly, the lack of a significant influence of understanding on trust calls existing theoretical assumptions into question. Shin had assumed that understandability might also influence trust, but the results of this study suggest that understandability alone does not influence trust in credit decisions [47]. It seems that users tend to prefer the factors of fairness and transparency over pure clarity in explanations when it comes to credit decisions, as in this study.

It can therefore be said that the findings from this study show that TAM needs to be expanded with various constructs in order to be applicable to XAI explanations. By integrating XAI-specific constructs such as trust, transparency, and fairness, more accurate explanations can be made regarding user acceptance of XAI explanations than with TAM alone.

6.3 Practical implications

From a practical perspective, this study is relevant for sensitive areas such as credit decisions in the banking sector when implementing explainable AI systems. But above all, to examine the general acceptance of AI and XAI. The results of the study show that trust is a decisive factor in whether users accept an AI-generated explanation. Therefore, when implementing XAI, developers and companies should prioritize trust mechanisms among the factors of an XAI explanation in their system.

To increase the influence factor of trust, two additional factors (transparency and fairness) are required, which have a direct influence on trust and significantly affect this factor. This means that explanations must not only be understandable, but also perceived as unbiased and transparent in the decision-making process. In the example of lending, providing the decision-making criteria that influence the decision and proof of non-discrimination could be used to increase user trust in the AI system. For example, in medicine, a hospital might transparently show the patient the results and which ones were decisive. Such ideas could increase fairness and the transparency in practice.

The results also suggest that comprehensibility alone is not sufficient to increase user trust. Even if users find the explanation easy to understand, this alone is not enough to reduce skepticism. The explanation must include the factors of fairness and transparency in order to reduce skepticism.

Due to the subordinate role of the original TAM constructs (PU and PEOU), companies should ensure during development that the main focus is not on user-friendliness, but rather that the AI system is developed in such a way that direct emphasis is placed on the factors of fairness and transparency. This will increase user confidence and ultimately increase the intention to use the system.

The results underscore that the success of an XAI system depends not only on technical performance, but rather on how the system conveys transparency, fairness, and, above all, trust to the user.

6.4 Scientific contributions

This work and its study make several scientific contributions to XAI research from the user perspective and TAM. First, it integrates Davis' original TAM with new constructs and relationships from Shin's model. This yields new insights into how classic acceptance constructs in combination with XAI factors jointly influence user acceptance.
Furthermore, the work provides insights from a sensitive area, namely the financial sector in the field of credit decisions. The case study in the study also increases the practical usefulness for both science and industry.
Ultimately, that study makes an important contribution through the application of comprehensive analysis strategies ranging from descriptive statistics, reliability analysis, and correlation analysis to structural equation modeling using SmartPLS. This increases the robustness and value of this study and thus provides a good basis for future research investigating the user acceptance of explainable AI in conjunction with TAM.

6.5 Limitations

Like this study, this work also has certain limitations, which will now be explained in more detail and should be taken into account in the results. On the one hand, the sample size in the study was limited and also showed inequalities in terms of gender and age coverage. This means that the results are limited in their generalizability and that future studies are needed to address the topic under investigation and resolve the limitations mentioned.

Another important point is that the case study is limited to a specific domain (loan decision-making). Although this case study covers an important area, the results may not be applicable to other domains such as healthcare or law. Therefore, future research should test the proposed combined model in other highly sensitive areas to validate its general applicability.

A third point is that the study relies on Likert scales, as was also the case with Davis and Shin. While this approach is not uncommon, future studies should also incorporate experiments or behavioral measurements, such as the actual observation of user interactions with AI systems providing XAI explanations.

In addition, the study was conducted within a two-week period, which means that no long-term conclusions can be drawn, as would be possible with longitudinal studies. Such long-term studies could explore how trust and acceptance of XAI develop over time when participants interact with such systems.

Despite these limitations, the study opens up avenues for further research by highlighting the influencing factors of trust, fairness, and transparency in the field of XAI and pointing out that certain TAMs, such as Davis' model, may need to be adapted in order to be applicable to XAI [15].

7 Conclusion and Outlook

This thesis aimed to examine the acceptance of XAI-Explanation from the user perspective, based on a TAM model. The focus of the work was to explore how acceptance of AI systems using XAI can be increased and which influencing factors are most important in this context.

A study was conducted that evaluated a combined model: Davis's original TAM and Shin's model, developed specifically for XAI. The study included a fictitious AI decision with an XAI explanation, a case study from the area of loan applications. The following research questions guided the investigation:

1. Which technology acceptance models are suitable for evaluating AI-generated explanations?
2. According to TAM, which factors influence the acceptance of AI-generated explanations?
3. How influential are the respective factors for the acceptance of XAI explanations?

7.1 Summary of key findings

The results of the study yielded several insights. First, it became evident that Davis' original TAM is not sufficient to capture the acceptance of XAI and its influencing factors. In Davis' TAM, the constructs PU and PEOU are the two main factors influencing the intention to use [15]. However, according to this study, these factors do not have a significant impact on usage intention in the context of XAI.

Other influencing factors prove to be more decisive, namely those from Shin's model, which are also emphasized in the XAI literature [47]. In particular, factors such as trust, transparency, and fairness proved to be extremely important from the XAI user perspective. Trust, in particular, showed the strongest correlation with usage intention. Transparency and fairness, in turn, were found to be important influencing factors for trust, while understanding did not show a significant effect.

This, however, contradicts previous assumptions in the field, where it is often stated that understandability is an important prerequisite for trust. The findings of this study suggest that, in the context of loan decisions, transparency and fairness are perceived as more important and carry greater weight than understanding in an XAI explanation. The model analysis in Chapter 3.2 also revealed that no TAM was found that contains all the important influencing factors mentioned in the literature regarding XAI.

7.2 Theoretical and practical contributions

From a theoretical perspective, it becomes clear that classical acceptance models need to be extended with additional XAI-specific constructs in order to be applicable. Furthermore, trust should function as a mediating variable between explanation attributes (e.g., transparency and fairness) and usage intention, which represents the ultimate goal of AI systems.

From a practical perspective, companies should ensure that their explanations are designed to incorporate a high degree of fairness and transparency. These factors influence trust much more strongly than usefulness does, and trust, in turn, has the greatest impact on usage intention.

7.3 Restrictions

Like any other empirical study, the conducted study also has limitations that should be acknowledged. First, the sample size was relatively small, with only 49 participants, and there were imbalances within the sample in terms of gender and age groups. Furthermore, the study examined only a single case example (loan decision), which limits the transferability of the findings to other sensitive domains.

In addition, the results are based on self-reports using Likert scales, which may lead to biases due to subjective perception. Finally, the study represents a snapshot in time rather than a longitudinal study.

7.4 Outlook

Future research aiming to apply the combined model should project it onto different scenarios, such as in the medical domain where research on XAI already exists or in the legal domain. These studies should involve larger sample sizes and also be designed longitudinally. This would make it possible to gain deeper insights into how the acceptance and the Influence Factors develop and change over time. Furthermore, future studies could classify participants into different groups to perhaps find differences between the respective groups.
It would also be useful to further refine constructs such as Understanding in order to determine more precisely under which conditions trust is actually fostered. But perhaps also to incorporate other constructs into the model that were not considered in this study. This could allow future research to develop a more differentiated picture of XAI acceptance and its influencing factors.

References

[1] Adadi, A., Berrada, M.: Peeking inside the black-box: A survey on explainable artificial intelligence (xai). IEEE Access **6**, 52138–52160 (2018). https://doi.org/https://doi.org/10.1109/ACCESS.2018.2870052

[2] Afroogh, S., Akbari, A., Malone, E., Kargar, M., Alambeigi, H.: Trust in ai: progress, challenges, and future directions. Humanities and Social Sciences Communications **11**(1) (Nov 2024). https://doi.org/https://doi.org/10.1057/s41599-024-04044-8

[3] Anjomshoae, S., Najjar, A., Calvaresi, D., Främling, K.: Explainable agents and robots: Results from a systematic literature review. In: 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019. pp. 1078–1088. International Foundation for Autonomous Agents and Multiagent Systems (2019)

[4] Baroni, I., Re Calegari, G., Scandolari, D., Celino, I.: Ai-tam: a model to investigate user acceptance and collaborative intention inhuman-in-the-loop ai applications. Human Computation **9**(1), 1–21 (May 2022). https://doi.org/https://doi.org/10.15346/hc.v9i1.134

[5] Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F.: Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. Information Fusion **58**, 82–115 (Jun 2020). https://doi.org/https://doi.org/10.1016/j.inffus.2019.12.012

[6] Brydges, C.R.: Effect size guidelines, sample size calculations, and statistical power in gerontology. Innovation in Aging **3**(4) (Aug 2019). https://doi.org/https://doi.org/10.1093/geroni/igz036

[7] Burrell, J.: How the machine 'thinks': Understanding opacity in machine learning algorithms. Big Data amp; Society **3**(1) (Jan 2016). https://doi.org/https://doi.org/10.1177/2053951715622512

[8] Buschmeier, H., Buhl, H.M., Kern, F., Grimminger, A., Beierling, H., Fisher, J., Groß, A., Horwath, I., Klowait, N., Lazarov, S., Lenke, M., Lohmer, V., Rohlfing, K., Scharlau, I., Singh, A., Terfloth, L., Vollmer, A.L., Wang, Y., Wilmes, A., Wrede, B.: Forms of understanding for xai-explanations (2023). https://doi.org/https://doi.org/10.48550/arXiv.2311.08760

[9] Chaiken, S.: Heuristic versus systematic information processing and the use of source versus message cues in persuasion. Journal of Personality and Social Psychology **39**(5), 752–766 (Nov 1980). https://doi.org/https://psycnet.apa.org/doi/10.1037/0022-3514.39.5.752

[10] Chau, P.Y.: An empirical assessment of a modified technology acceptance model. Journal of Management Information Systems **13**(2), 185–204 (Sep 1996). https://doi.org/https://doi.org/10.1080/07421222.1996.11518128

[11] Chuttur, M.: Overview of the technology acceptance model: Origins, developments and future directions (2009)

[12] Cohen: Applied multiple regression/correlation analysis for the behavioral sciences (Jun 2013). https://doi.org/https://doi.org/10.4324/9780203774441

[13] Cronbach, L.J.: Coefficient alpha and the internal structure of tests. Psychometrika **16**(3), 297–334 (Sep 1951). https://doi.org/https://doi.org/10.1007/BF02310555

[14] Dahlberg, T., M.N.O.J..Z.A.: Trust enhanced technology acceptance model – consumer acceptance of mobile payment solutions: Tentative evidence. Presented at the Stockholm Mobility Roundtable (2003)

[15] Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly **13**(3), 319 (Sep 1989). https://doi.org/https://doi.org/10.2307/249008

[16] Dixit, R.V., Prakash, G.: Intentions to use social networking sites (sns) using technology acceptance model (tam): An empirical study. Paradigm: A Management Research Journal **22**(1), 65–79 (Apr 2018). https://doi.org/https://doi.org/10.1177/0971890718758201

[17] Dodge, J., Liao, Q.V., Zhang, Y., Bellamy, R.K.E., Dugan, C.: Explaining models: an empirical study of how explanations impact fairness judgment pp. 275–285 (Mar 2019). https://doi.org/https://doi.org/10.1145/3301275.3302310

[18] Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., Qian, B., Wen, Z., Shah, T., Morgan, G., Ranjan, R.: Explainable ai (xai): Core ideas, techniques, and solutions. ACM Computing Surveys **55**(9), 1–33 (Jan 2023). https://doi.org/https://doi.org/10.1145/3561048

[19] Ehsan, U., Tambwekar, P., Chan, L., Harrison, B., Riedl, M.O.: Automated rationale generation: a technique for explainable ai and its effects on human perceptions (Mar 2019). https://doi.org/https://doi.org/10.1145/3301275.3302316

[20] von Eschenbach, W.J.: Transparency and the black box problem: Why we do not trust ai. Philosophy amp; Technology **34**(4), 1607–1622 (Sep 2021). https://doi.org/https://doi.org/10.1007/s13347-021-00477-0

[21] Fan, F., Xiong, J., Li, M., Wang, G.: On interpretability of artificial neural networks: A survey (2020). https://doi.org/10.48550/ARXIV.2001.02522

[22] Fishbein, M., Ajzen, I.: Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Addison-Wesley, Reading, MA (1975)

[23] Gefen, Karahanna, Straub: Trust and tam in online shopping: An integrated model. MIS Quarterly **27**(1), 51 (2003). https://doi.org/https://doi.org/10.2307/30036519

[24] Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., Yang, G.Z.: Xai—explainable artificial intelligence. Science Robotics **4**(37) (Dec 2019). https://doi.org/https://doi.org/10.1126/scirobotics.aay7120

[25] Hair, J.F., Hult, G.T.M., Ringle, C.M., Sarstedt, M., Danks, N.P., Ray, S.: Partial least squares structural equation modeling (pls-sem) using r: A workbook. Classroom Companion: Business (2021). https://doi.org/https://doi.org/10.1007/978-3-030-80519-7

[26] Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., Hussain, A.: Interpreting black-box models: A review on explainable artificial intelligence. Cognitive Computation **16**(1), 45–74 (Aug 2023). https://doi.org/10.1007/s12559-023-10179-8

[27] Herrera, F.: Reflections and attentiveness on explainable artificial intelligence (xai). the journey ahead from criticisms to human–ai collaboration. Information Fusion **121**, 103133 (Sep 2025). https://doi.org/https://doi.org/10.1016/j.inffus.2025.103133

[28] Hoffman, R.R., Miller, T., Klein, G., Mueller, S.T., Clancey, W.J.: Increasing the value of xai for users: A psychological perspective. KI - Künstliche Intelligenz **37**(2–4), 237–247 (Jul 2023). https://doi.org/https://link.springer.com/article/10.1007/s13218-023-00806-9

[29] Hoffman, R.R., Mueller, S.T., Klein, G., Litman, J.: Metrics for explainable ai: Challenges and prospects (2018). https://doi.org/https://doi.org/10.48550/arXiv.1812.04608

[30] Hossain, S.A., Bao, Y., Hasan, N., Islam, M.F.: Perception and prediction of intention to use online banking systems: An empirical study using extended tam. International Journal of Research in Business and Social Science **9**(1), 112–126 (2020)

[31] Joo, J., Sang, Y.: Exploring koreans' smartphone usage: An integrated model of the technology acceptance model and uses and gratifications theory. Computers in Human Behavior **29**(6), 2512–2518 (Nov 2013). https://doi.org/https://doi.org/10.1016/j.chb.2013.06.002

[32] Kalorii, D.: Measuring the impact of explainability on decision-making in xai systems (2025). https://doi.org/https://dx.doi.org/10.2139/ssrn.5392341

[33] Lundberg, S.M., Lee, S.I.: A unified approach to interpreting model predictions. In: Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (eds.) Advances in Neural Information Processing Systems. vol. 30. Curran Associates, Inc. (2017), https://proceedings.neurips.cc/paper$_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-$ $Paper.pdf$

[34] Marey, A., Arjmand, P., Alerab, A.D.S., Eslami, M.J., Saad, A.M., Sanchez, N., Umair, M.: Explainability, transparency and black box challenges of ai in radiology: impact on patient care in cardiovascular radiology. Egyptian Journal of Radiology and Nuclear Medicine **55**(1) (Sep 2024). https://doi.org/https://doi.org/10.1186/s43055-024-01356-2

[35] Mcknight, D.H., Carter, M., Thatcher, J.B., Clay, P.F.: Trust in a specific technology: An investigation of its components and measures. ACM Transactions on Management Information Systems **2**(2), 1–25 (Jun 2011). https://doi.org/https://doi.org/10.1145/1985347.1985353

[36] Mian, S.M., Khan, M.S., Shawez, M., Kaur, A.: Artificial intelligence (ai), machine learning (ml) & deep learning (dl): A comprehensive overview on techniques, applications and research directions pp. 1404–1409 (Jul 2024). https://doi.org/https://doi.org/10.1109/ICSCSS60660.2024.10625198

[37] Miller, T.: Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence **267**, 1–38 (Feb 2019). https://doi.org/https://doi.org/10.1016/j.artint.2018.07.007

[38] Nallakaruppan, M.K., Chaturvedi, H., Grover, V., Balusamy, B., Jaraut, P., Bahadur, J., Meena, V.P., Hameed, I.A.: Credit risk assessment and financial decision support using explainable artificial intelligence. Risks **12**(10), 164 (Oct 2024). https://doi.org/https://doi.org/10.3390/risks12100164

[39] Panagoulias, D.P., Virvou, M., Tsihrintzis, G.A.: A novel framework for artificial intelligence explainability via the technology acceptance model and rapid estimate of adult literacy in medicine using machine learning. Expert Systems with Applications **248**, 123375 (Aug 2024). https://doi.org/https://doi.org/10.1016/j.eswa.2024.123375

[40] Patidar, N., Mishra, S., Jain, R., Prajapati, D., Solanki, A., Suthar, R., Patel, K., Patel, H.: Transparency in ai decision making: A survey of explainable ai methods and applications. Advances of Robotic Technology **2**(1) (2024). https://doi.org/10.23880/art-16000110

[41] PEARSON, K.: Notes on the history of correlation. Biometrika **13**(1), 25–45 (1920). https://doi.org/https://doi.org/10.2307/2331722

[42] Rahmani, A.M., Yousefpoor, E., Yousefpoor, M.S., Mehmood, Z., Haider, A., Hosseinzadeh, M., Ali Naqvi, R.: Machine learning (ml) in medicine: Review, applications, and challenges. Mathematics **9**(22), 2970 (Nov 2021). https://doi.org/https://doi.org/10.3390/math9222970

[43] Ribeiro, M.T., Singh, S., Guestrin, C.: "why should i trust you?": Explaining the predictions of any classifier pp. 1135–1144 (Aug 2016). https://doi.org/https://doi.org/10.1145/2939672.2939778

[44] Rudin, C.: Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead (2018). https://doi.org/https://doi.org/10.48550/arXiv.1811.10154

[45] Sadok, H., Sakka, F., El Maknouzi, M.E.H.: Artificial intelligence and bank credit analysis: A review. Cogent Economics amp; Finance **10**(1) (Jan 2022). https://doi.org/https://doi.org/10.1080/23322039.2021.2023262

[46] Schmidt, A.: Longitudinal assessment of user acceptance of explainable artificial intelligence (xai): Analyzing contextual influences and evolution over time (2025). https://doi.org/https://dx.doi.org/10.2139/ssrn.5372707

[47] Shin, D.: The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable ai. International Journal of Human-Computer Studies **146**, 102551 (Feb 2021). https://doi.org/https://doi.org/10.1016/j.ijhcs.2020.102551

[48] Soori, M., Jough, F.K.G., Dastres, R., Arezoo, B.: Ai-based decision support systems in industry 4.0, a review. Journal of Economy and Technology (Aug 2024). https://doi.org/https://doi.org/10.1016/j.ject.2024.08.005

[49] Spearman, C.: The proof and measurement of association between two things. The American Journal of Psychology **15**(1), 72 (Jan 1904). https://doi.org/https://doi.org/10.2307/1412159

[50] Tricot, R.: Venture capital investments in artificial intelligence: Analysing trends in vc in ai companies from 2012 through 2020 (2021)

[51] Velarde, G.: Artificial intelligence trends and future scenarios: Relations between statistics and opinions pp. 64–70 (Dec 2021). https://doi.org/https://doi.org/10.1109/CogMI52975.2021.00017

[52] Venkatesh, Morris, Davis, Davis: User acceptance of information technology: Toward a unified view. MIS Quarterly **27**(3), 425 (2003). https://doi.org/10.2307/30036540

[53] Venkatesh, V., Bala, H.: Technology acceptance model 3 and a research agenda on interventions. Decision Sciences **39**(2), 273–315 (May 2008). https://doi.org/https://doi.org/10.1111/j.1540-5915.2008.00192.x

[54] Venkatesh, V., Davis, F.D.: A theoretical extension of the technology acceptance model: Four longitudinal field studies. Management Science **46**(2), 186–204 (Feb 2000). https://doi.org/https://doi.org/10.1287/mnsc.46.2.186.11926

[55] Westphal, M., Vössing, M., Satzger, G., Yom-Tov, G.B., Rafaeli, A.: Decision control and explanations in human-ai collaboration: Improving user perceptions and compliance. Computers in Human Behavior **144**, 107714 (Jul 2023). https://doi.org/https://doi.org/10.1016/j.chb.2023.107714

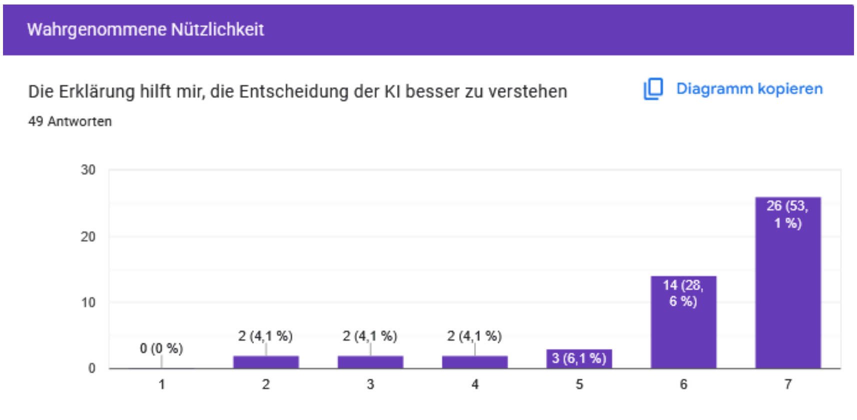
[56] Yudiantara, P.O., Widagda, I.G.N.J.A.: role of trust in mediating the effect of perceived usefulness and perceived ease of use on decisions to use the linkaja digital wallet: Study on linkaja application users in denpasar city. International journal of health sciences pp. 6310–6327 (Jul 2022). https://doi.org/https://doi.org/10.53730/ijhs.v6nS4.11176

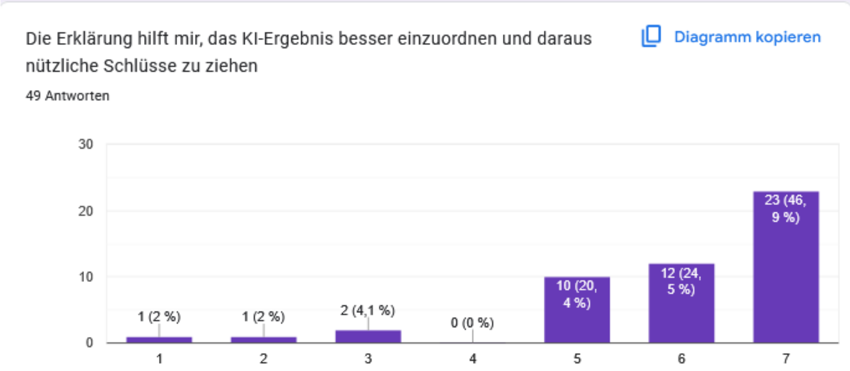
[57] Zednik, C.: Solving the black box problem: A normative framework for explainable artificial intelligence. Philosophy amp; Technology **34**(2), 265–288 (Dec 2019). https://doi.org/https://link.springer.com/article/10.1007/s13347-019-00382-7

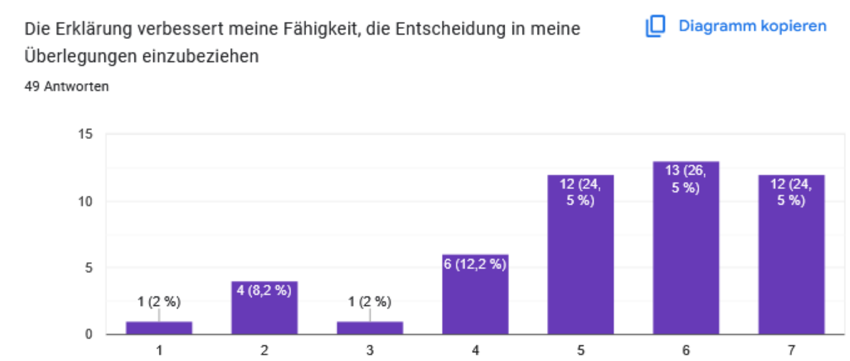
[58] Zhai, X., Chu, X., Chai, C.S., Jong, M.S.Y., Istenic, A., Spector, M., Liu, J.B., Yuan, J., Li, Y.: A review of artificial intelligence (ai) in education from 2010 to 2020. Complexity **2021**(1) (Jan 2021). https://doi.org/https://doi.org/10.1155/2021/8812542

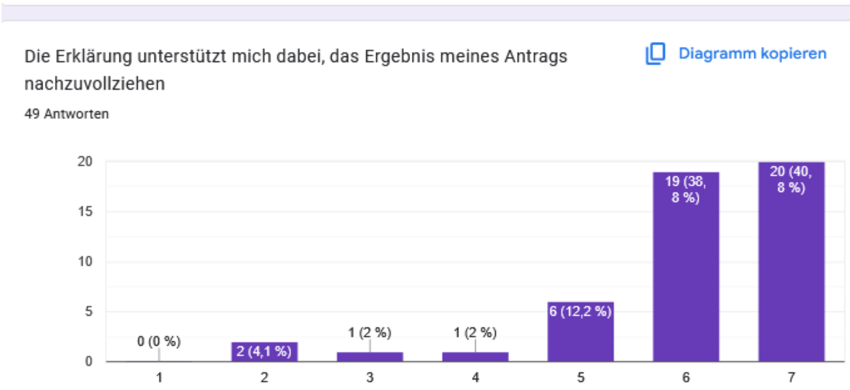
8 Attachment

8.1 Survey results







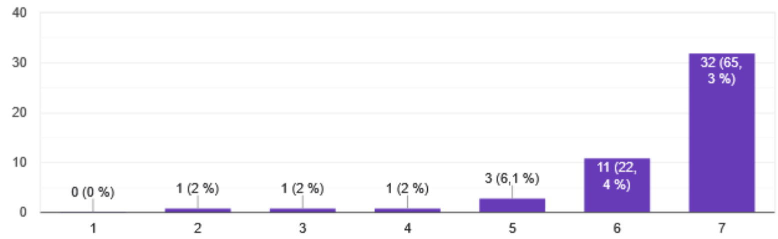


Wahrgenommene Verständlichkeit

Die Erklärung ist leicht verständlich

Diagramm kopieren

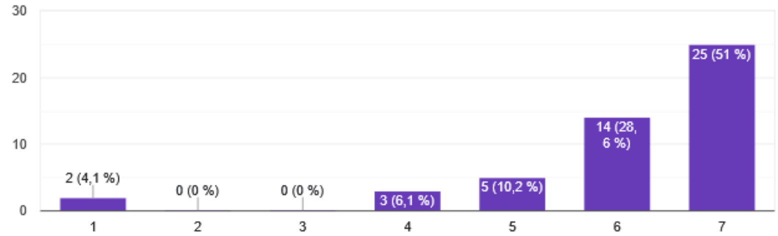
49 Antworten



Die Erklärung ist sprachlich klar und übersichtlich strukturiert

Diagramm kopieren

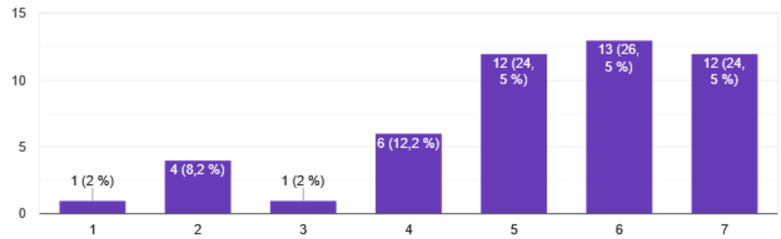
49 Antworten



Die Erklärung verbessert meine Fähigkeit, die Entscheidung in meine Überlegungen einzubeziehen

Diagramm kopieren

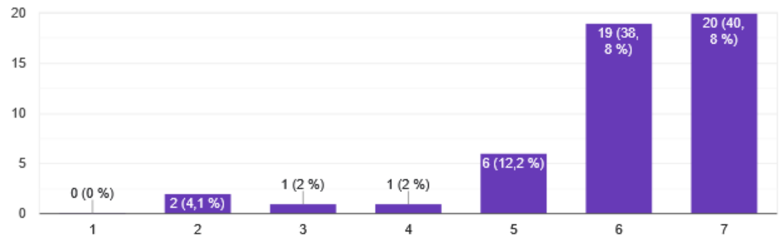
49 Antworten



Die Erklärung unterstützt mich dabei, das Ergebnis meines Antrags nachzuvollziehen

Diagramm kopieren

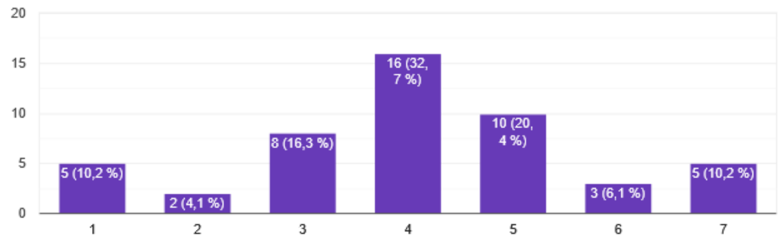
49 Antworten



Vertrauen

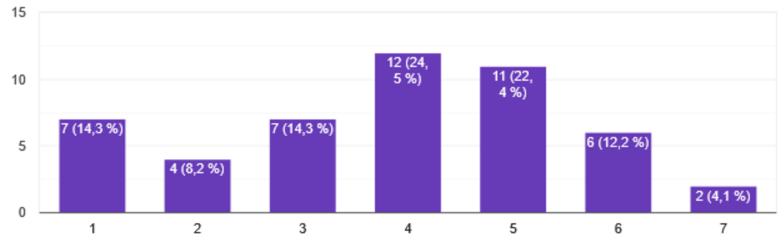
Die Erklärung erhöht mein Vertrauen in das KI-System Diagramm kopieren

49 Antworten



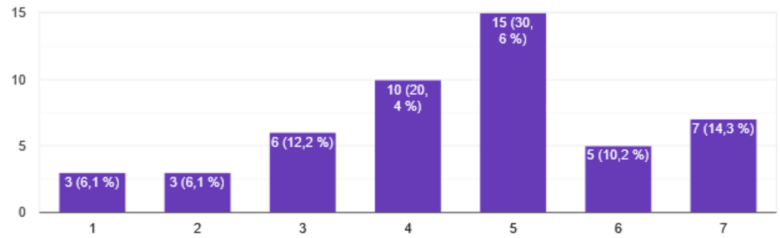
Ich würde mich in einer ähnlichen Situation auf dieses System verlassen Diagramm kopieren

49 Antworten



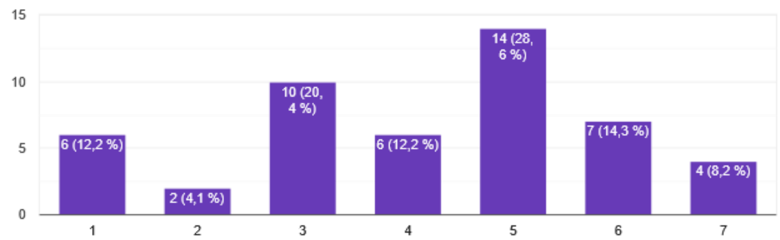
Die Erklärung vermittelt Zuverlässigkeit im Entscheidungsprozess Diagramm kopieren

49 Antworten



Die Erklärung gibt mir Sicherheit, dass die Entscheidung korrekt ist Diagramm kopieren

49 Antworten

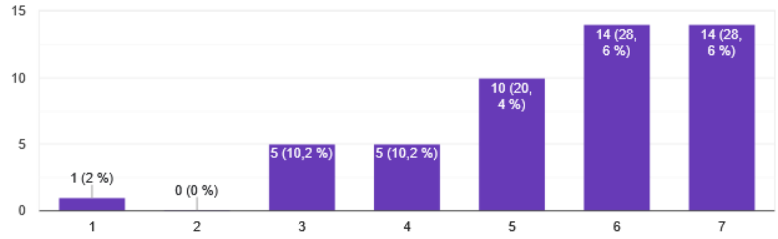


Transparenz

Ich habe den Eindruck, dass mir durch die Erklärung die Entscheidungsgrundlage der KI offengelegt wurde

Diagramm kopieren

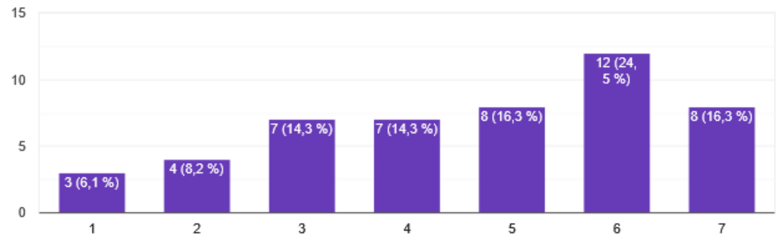
49 Antworten



Die Erklärung vermittelt mir, wie transparent der Entscheidungsprozess der KI ist

Diagramm kopieren

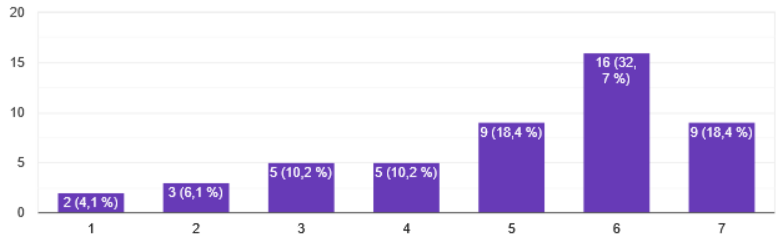
49 Antworten



Ich habe **nicht** das Gefühl, dass wichtige Informationen zur Entscheidung verborgen wurden

Diagramm kopieren

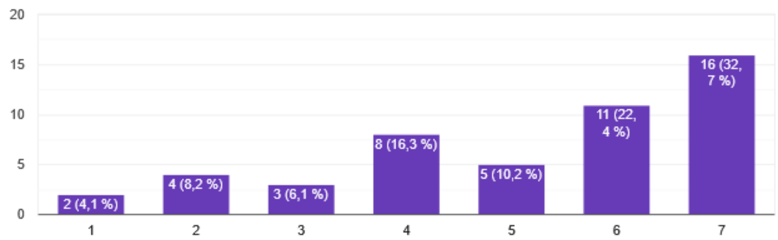
49 Antworten



Die Erklärung zeigt offen, auf welchen Informationen die Entscheidung basiert

Diagramm kopieren

49 Antworten

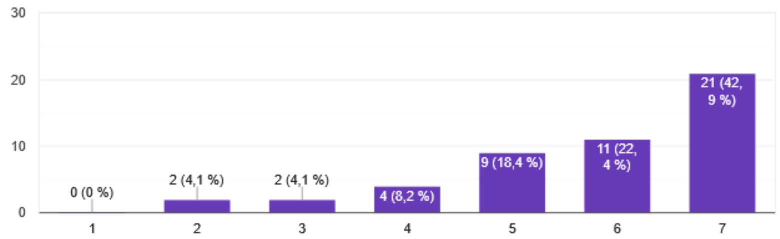


Verständnis

Ich verstehe, warum die KI diese Entscheidung getroffen hat

Diagramm kopieren

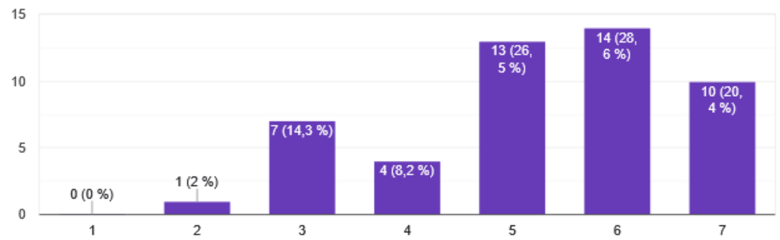
49 Antworten



Die Erklärung zeigt mir, wie die wichtigsten Entscheidungsfaktoren miteinander zusammenhängen

Diagramm kopieren

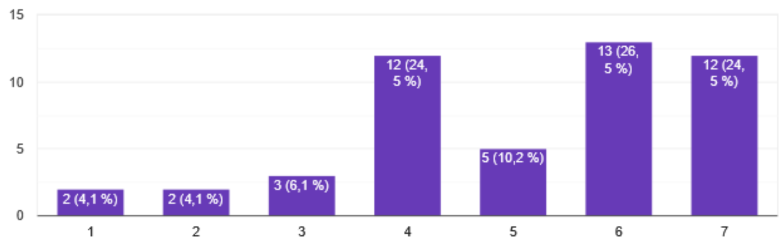
49 Antworten



Die Erklärung hilft mir, die Entscheidungslogik Schritt für Schritt nachzuvollziehen

Diagramm kopieren

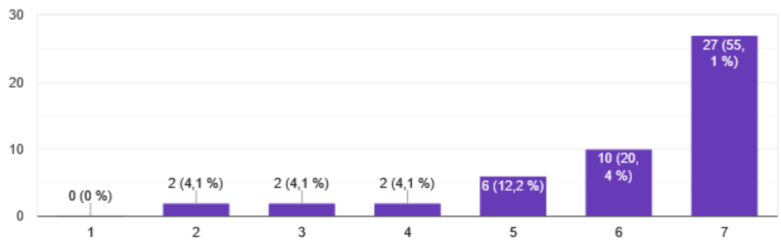
49 Antworten



Ich könnte diese Entscheidung jemand anderem erklären

Diagramm kopieren

49 Antworten

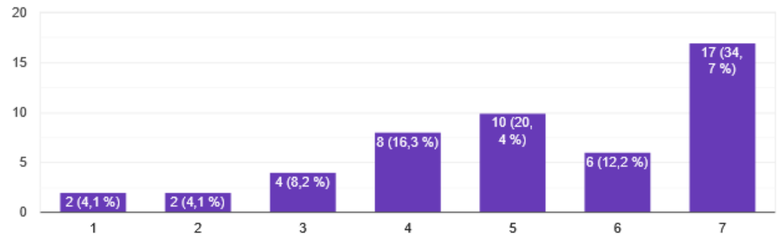


Fairness

Die Entscheidung erscheint mir auf Basis der Erklärung fair

Diagramm kopieren

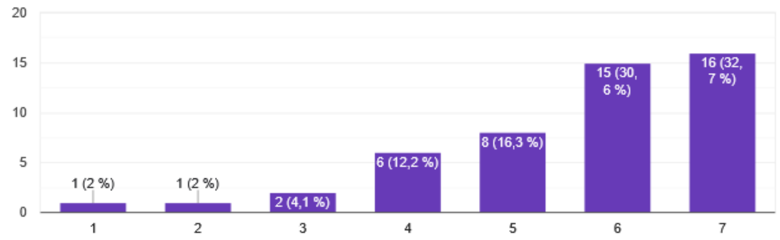
49 Antworten



Die Erklärung zeigt, dass ähnliche Fälle gleich behandelt werden

Diagramm kopieren

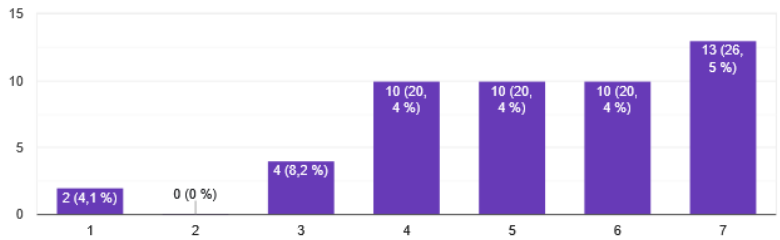
49 Antworten



Die Erklärung vermittelt den Eindruck, dass die Entscheidung objektiv war

Diagramm kopieren

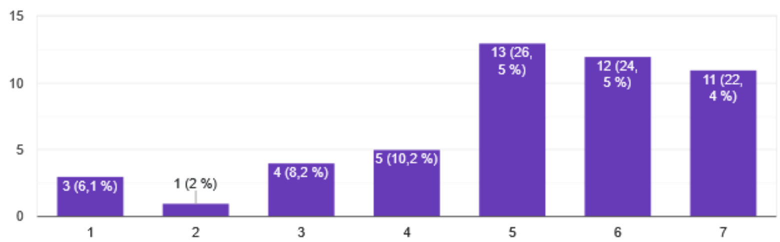
49 Antworten



Die Erklärung vermittelt mir den Eindruck, dass die Entscheidung niemanden bevorzugt oder benachteiligt

Diagramm kopieren

49 Antworten

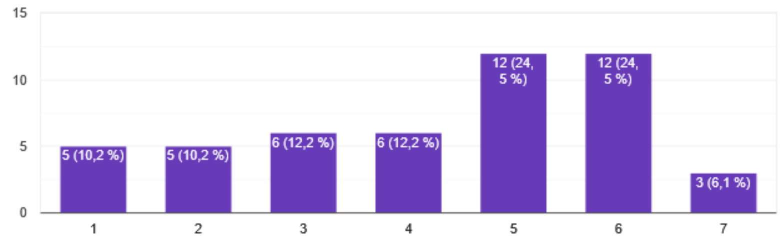


Nutzungsabsicht

Ich würde ein KI-System nutzen, das solche Erklärungen gibt

Diagramm kopieren

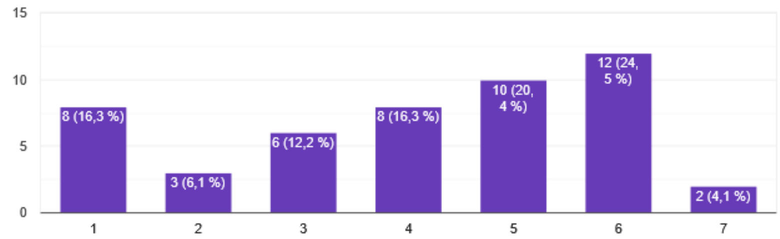
49 Antworten



Ich würde ein solches System anderen empfehlen

Diagramm kopieren

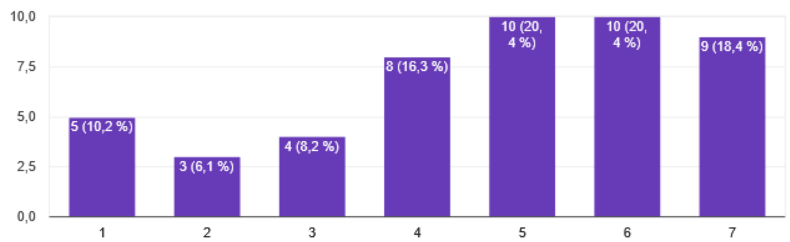
49 Antworten



Ich wäre eher bereit, Entscheidungen der KI zu akzeptieren, wenn solche Erklärungen immer gezeigt würden

Diagramm kopieren

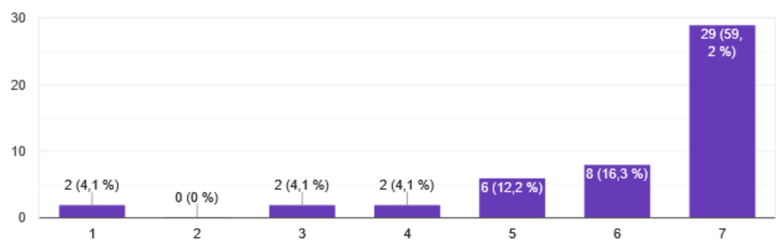
49 Antworten



Ich bevorzuge KI-Systeme, die ihre Entscheidungen verständlich begründen

Diagramm kopieren

49 Antworten

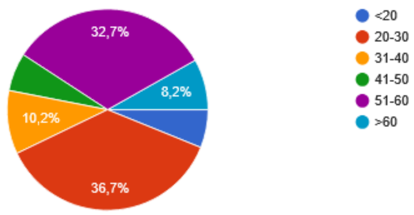


Allgemeine Fragen

Alter des Befragten

Diagramm kopieren

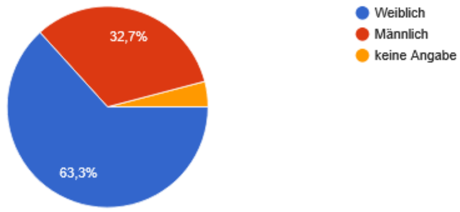
49 Antworten



Geschlecht

Diagramm kopieren

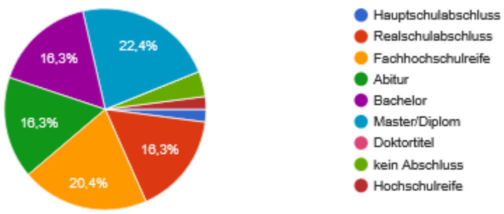
49 Antworten



Höchster Bildungsabschluss

Diagramm kopieren

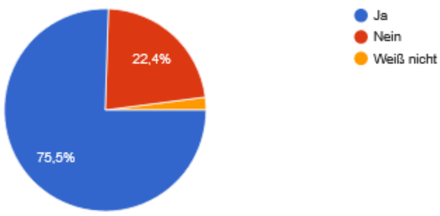
49 Antworten



Hast du schon einmal ein KI-System verwendet?

Diagramm kopieren

49 Antworten



Wie oft wird KI im Alltag bei dir benutzt? (bspw. ChatGPT)

Diagramm kopieren

49 Antworten