Building Trust in Financial Robo-Advisory and Overcoming Ethical Challenges

Alex Zarifis and Manuel Nunes

Abstract Previous research identified five areas where trust could be built for financial robo-advisors. This research identifies the specific methods for building trust for each of these five areas. Building trust with human oversight can be done in three specific ways: (1) oversight of AI model and service it provides, (2) human explanations for advice and processes, and (3) regular human communication outside specific processes. Building trust with transparency and control can be done by: (1) recognise that there are beliefs that do not immediately lead to action, (2) explain actions and demonstrating ability and integrity, and (3) balance automation with control to avoid disempowering users. Building trust with accuracy and usefulness can be done by: (1) accurate and relevant information conveying competence and integrity, (2) usefulness supported by back testing and multiple methods, and (3) enhanced usefulness through personalisation. Building trust with ease of use and support can be done by: (1) achieve ease of use while keeping all the necessary functionality, (2) support time management with nudges, and (3) useful and reliable service with high performance. Building trust with a humanlike interaction can be done by: (1) utilise the latest sentiment analysis and longest contextual memory, (2) the robo-advisor should have a clear persona, and (3) build a relationship and a feeling of friendship and partnership.

Keywords Robo-advisors · Fintech · Insurtech · Trust · Chatbot · Agentic AI

A. Zarifis (✉)
Southampton Business School, University of Southampton, Southampton, UK

Cambridge Centre for Alternative Finance, Cambridge, UK
e-mail: a.zarifis@soton.ac.uk

M. Nunes
University of Southampton, Southampton, UK
e-mail: manuel.nunes@soton.ac.uk

N. Nahidi, A. Zarifis (eds.), *AI, FinTech, and the Future of Robo-Advisory*,
Contributions to Finance and Accounting,
https://doi.org/10.1007/978-3-032-18109-1_5

1 Introduction

Most financial institutions can meet their short-term targets by offering competitive services that are attractive. The best financial institutions go beyond simply meeting short-term targets and build strong relationships. Even in the age of AI and data-driven decisions, evidence of the value of building strong relationships is all around us. Both retail customers and organisations show trust and loyalty to financial institutions they have a deeper emotional connection with. Until now, this usually happened naturally with the face-to-face interaction. Even when we interacted with systems on the internet, until recently these systems did not use AI but human logic coded into them. Now with the prevalence of AI and financial robo-advisory, new ways are needed to build trust (Cheng et al. 2021; Dietzmann et al. 2023). A typical AI implementation is naturally good at many aspects of financial services, but it is not naturally effective at building trust. Typically, some human intervention, such as supervised learning and specific instructions given before every answer the AI gives, are necessary to create an interaction that builds trust. Unfortunately, trying to get the AI to build trust in the way humans build trust doesn't usually work. There is extensive evidence that AI behaving as a human can often backfire or even be creepy, so new methods to build trust are necessary (Mori 2012). This research identifies specific methods to build trust in financial robo-advisors.

If we look for a similar case to the disruption AI is causing, the closest case that has this scale is the internet. The internet is not just a technology, but a series of standards created by Tim Burners-Lee and others over many years to structure the technology so that it delivers a service people trust.

The capabilities of AI increase on a weekly basis but implementing AI in a way that builds trust is still the responsibility of the humans in the organisation. There is a technical and business dimension to this. We often refer to the AI stack, as an AI implementation is not typically one application but a series of systems. We can also consider the other aspects of the financial institution as additional layers. All these technical and business layers to the 'financial robo-advisor stack' must be aligned to a trust-building strategy. AI can offer us many things such as near-infinite scaling but for effective relationships with humans, whether they are employees, customers, regulators, or other stakeholders, there needs to be a human behind the robo-advisor, a puppet master, a ghost in the machine (Fügener et al. 2021).

Previous research has identified four issues that affect trust in robo-advisors for specific financial questions, and five that affect trust in robo-advisors for vague financial questions (Zarifis and Cheng 2024). The four issues are the same for both specific and vague questions, but vague questions have one additional issue. The four issues that affect trust in all financial questions are human oversight, transparency and control, accuracy and usefulness, ease of use, and support. The additional, fifth issue that builds trust in vague questions is humanlike interaction.

While identifying these five areas is helpful, they are not specific methods to build trust but broad issues to focus on to build trust. It is difficult to have specific methods to build trust in financial robo-advisors, as each organisation and their

context are different. Nevertheless, this research attempts to identify some more specific methods to build trust for each of these five issues. Therefore, the research question is:

What are the specific, practical methods to build trust in financial robo-advisors?

The previous research that identified the five trust-building areas had surveyed users of financial robo-advisors. As the user's perspective on these issues has already been recorded, this research implemented focus groups with finance experts to discuss with them the specific ways they would build trust for these five areas.

This research finds that building trust with human oversight can be done in three specific ways: (1) oversight of AI model and service it provides, (2) human explanations for advice and processes, and (3) regular human communication outside specific processes. Building trust with transparency and control can be done by: (1) recognise that there are beliefs that do not immediately lead to action, (2) explain actions and demonstrating ability and integrity, and (3) balance automation with control to avoid disempowering users. Building trust with accuracy and usefulness can be done by: (1) accurate and relevant information conveying competence and integrity, (2) usefulness supported by back testing and multiple methods, and (3) enhanced usefulness through personalisation. Building trust with ease of use and support can be done by: (1) achieve ease of use while keeping all the necessary functionality, (2) support time management with nudges, and (3) useful and reliable service with high performance. Building trust with a humanlike interaction can be done by: (1) utilise the latest sentiment analysis and longest contextual memory, (2) the robo-advisor should have a clear persona, and (3) build a relationship and a feeling of friendship and partnership.

The following section is the theoretical foundation, followed by the methodology section, where the data collection using focus groups and thematic analysis are discussed. The fourth section on the analysis and findings presents the specific methods to implement for each of the five trust-building areas. Finally, the conclusion summarises the main findings.

2 Theoretical Foundation

This section first looks at the challenges to trust in financial robo-advisors before identifying the five issues that build trust in robo-advisory that form the basis for the data collection.

2.1 The Challenges to Trust in Financial Robo-Advisors

Despite the impressive performance of generative AI transforming finance (Chui et al. 2023), receiving financial advice from a robo-advisor still raises several

challenges to trust. From the user's perspective, there are several risks, each of which requires that they have trust in a positive outcome. Where there is risk, trust is needed (Lankton et al. 2015; McKnight and Chervany 2002). The risks from using a financial robo-advisor include that the AI does not understand their real motivation and intentions, technology supporting the robo-advisor does not work correctly, receiving damaging financial advice (Dietzmann et al. 2023), their sensitive personal information leaking, and having a cybersecurity breach (Jain et al. 2023). All these risks are exacerbated by the feeling a user can have that a human on the side of the financial institution is not responsible and anything that happens can be explained away as a technical issue.

There is also the additional potential risk of the robo-advisor making the consumer feel uneasy or even creepy. When we interact with traditional information systems, the interaction is similar to using a tool that has no thought process. When the user interacts with a robo-advisor, it may communicate in an inappropriate way, and this can create an additional feeling of insecurity or even creepiness (Mori 2012; Rajaobelina et al. 2021).

2.2 The Five Issues That Build Trust in Robo-Advisory

Research has found that when interacting with generative AI the nature of the task or questions can influence the perception of risk and trust (Cheng et al. 2021). Previous systems that used static code and not AI were limited to specific processes, such as reporting the price of a share. Because robo-advisors are more versatile and do not operate in a rigid way, they handle a breadth of questions and interact in a quasi-social way (Cheng et al. 2024). While there are many different questions asked, in terms of trust building, an important distinction must be made between specific and vague questions. There is a proven difference in how to build trust in these two different scenarios (Fig. 1).

Recent research has identified four issues that affect trust in robo-advisors for specific financial questions, and five that affect trust in robo-advisors for vague financial questions (Zarifis and Cheng 2024). The four issues are the same for both specific and vague questions, but vague questions have one additional issue. The four issues that affect trust in all financial questions are human oversight, transparency and control, accuracy and usefulness, ease of use, and support. The additional, fifth issue that builds trust in vague questions is humanlike interaction. These five trust-building areas are the starting point for this research, which will explore each of them further. These five areas should be further investigated to identify the specific practical ways to build trust (Fig. 2).

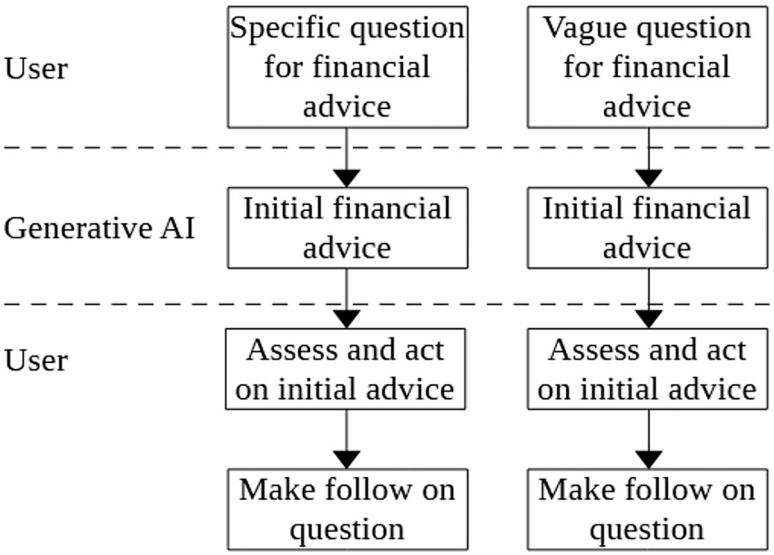

Fig. 1 The process of asking specific and vague financial questions

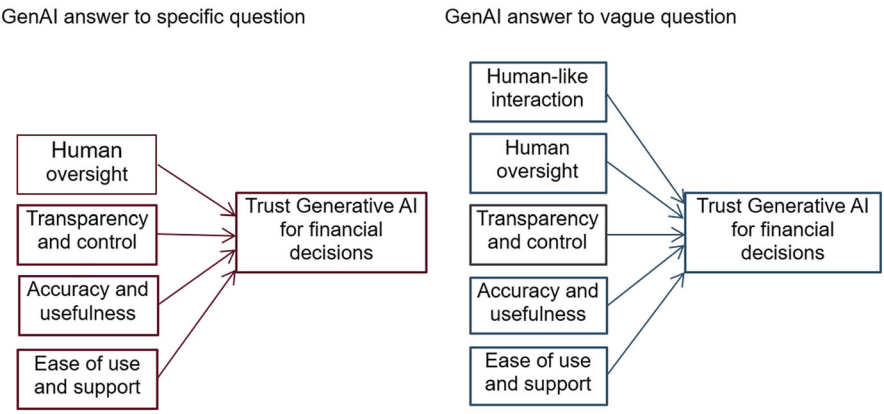

Fig. 2 The trust-building areas for specific and vague financial questions (Based on Zarifis and Cheng 2024)

3 Method

This research attempts to identify specific practical steps to build trust. Previous research identified five areas where trust can be built (Zarifis and Cheng 2024), so this research attempts to find realistic ways practitioners can implement them. For this reason, a qualitative approach with four focus groups was implemented.

3.1 Data Collection

The participants of the focus groups are professionals working in financial institutions. They are the kind of professionals who deal with these issues as part of their role. This ensures that the methods identified are realistic and can be applied. They do not all have the same role, but they are in roles that influence the implementation of the financial robo-advisor. The participants were from IT-related roles (9), financial roles (8), management (5), and marketing (5). They work for traditional banks, born-online banks, and traditional insurers in EU countries.

The original data collection plan was to implement four focus groups and then add more if this was necessary, but after the third focus group, the topics were indeed saturated. The fourth focus group did not add something significantly different, so no additional focus groups were added.

Before the online focus groups, a researcher presented the purpose of the research and the topics that were being explored. The rest of the focus group was divided into the five trust-building areas. So that an honest discussion is made, 'Chatham house rules' were used so no recordings were made and only notes were taken. The notes did not record the identities of the participants, just the main points that were made.

3.2 Data Analysis

The data was analysed using the NVivo software to implement a thematic analysis (Braun and Clarke 2006; Miles and Huberman 1994). As transcripts were not taken, only notes, the thematic analysis was done based on the notes. The high-level themes were the five trust-building areas. Several trust-building themes were identified beneath them.

4 Analysis and Discussion

This section identifies the practical ways to build trust for each of the five areas identified in previous research (Zarifis and Cheng 2024). There are four areas to build trust in both vague and specific questions, and one additional area that only helps for vague questions. There are some common themes such as transparency and having extensive safeguards and checks in place, but each area also has its idiosyncrasies and attempting to oversimplify these issues will not bring the best results. Each of these five areas needs to be handled individually for the best results.

4.1 Human Oversight

Despite the many challenges in keeping a human in the loop, including the problems humans can make, this is still the best way forward. Keeping humans in the loop

ensures a more human character to the process. Despite the capabilities of AI, a real person's judgement and empathy are beneficial. Equally importantly, someone always needs to be responsible. Those using robo-advisors to make financial decisions need to know who is behind the tool they are using. They should also know where the responsibility of those creating and maintaining the chatbot starts and ends, and where their responsibility starts and ends. Human oversight needs to be implemented effectively and then communicated effectively.

Having a team of humans in the loop, in itself, does not guarantee any benefit; it needs to be implemented in a targeted way. Those involved in the human oversight will almost inevitably also be involved in troubleshooting unexpected day-to-day problems, which often take many professionals more time than they plan to spend.

4.1.1 Oversight of AI Model and Service It Provides

The first way to build trust is the oversight of the AI models. The model needs to be regularly checked by humans in quantitative and qualitative ways. Tests can be made for AI model drift, but humans can also regularly go through the process themselves. The 'mystery shopper' has been an effective method to check quality and fix issues that could damage trust before AI emerged, and it is still entirely relevant. AI cannot empathise with the problems the people that use it have. AI can give the appearance that it empathises, but it benefits from supervised learning and human instructions before each response it makes to do this effectively. Additionally, a robo-advisor may be well suited for the regular events it deals with, but less common events such as cyberattacks may need some human intervention.

If these proactive oversight processes are in place this will reduce problems, but reactive oversight will still be needed. It is inevitable that those that use robo-advisors will have both valid and less valid complaints. There will be complaints that are not valid, but the users believe them to be valid. The processing of these complaints should not be done with the priority to achieve efficiency. AI offers scale; these processes should focus on empathy and human interaction. Users must feel there is an organisation full of experts behind the machine they are interacting with, not just a server room and an organisation left on autopilot. The problem escalation path should be clear and as painless as possible. It should be seen as an opportunity to touch bases, not a process that should be stripped down to its most basic and automated.

This layer of human oversight managing the AI model and the service should be overseen by a team of middle managers with a range of expertise, including an understanding of the technical, business, psychological, and sociological dimensions.

4.1.2 Human Explanations for Advice and Processes

In marketing literature, they emphasise the importance of the specific 'touchpoints' a user has with an organisation. The main value of focusing on the interactions the

user has with the organisation is that those resonate more than advertising or back-office operations not visible to the user.

There are opportunities during the process of using financial robo-advisors to get the best of both worlds, the performance of AI and the human touch. As already happens in some implementations, an explanation written by a human can accompany some advice. It does not have to be made for a specific question but for a type of situation, such as an unexpected event. Processes implemented with AI and automation are often effective in providing a good service when the user has their regular needs but often underperform when the user wants to make a big change to what they are doing such a changing their role in an organisation, or when they have a major change in their life. The effectiveness of generative AI is partly down to remembering previous conversations with a user but when there is a sharp change, these previous patterns are a hindrance, not a help.

Clear accountability that is regularly communicated is also important. There are no tricks to building trust; the humans in the financial institution cannot hide behind the robo-advisor. There should be regular updates that include the name of the human who wrote them. Any checks made by a human can also be logged with the person's name. The accountability does not always need to be communicated to the user, in some cases it just needs to be there in case they need it.

4.1.3 Regular Human Communication Outside Specific Processes

It is human nature to not only build an impression of a financial institution, but even give it human attributes such as traditional, aggressive, innovative, caring, friendly, kind, or unreliable. These human attributes we associate with an organisation influence trust. The performance of the organisation inevitably affects these issues, but it is not just about the performance. The communication should convey a coherent message and character both when reacting to requests and outside of specific processes. This is what the best organisations with strong values achieved before AI dominated the interaction with the user, and this is what should still happen today.

The communication must outline how experienced professionals with a successful track record are developing and validating the portfolios and handling exceptions. The more specific this is, the better. Including names, pictures, and where their office is emphasises that these are real people (Table 1).

4.2 Transparency and Control

Transparency and control are two different things, but we cannot have one without the other, so they are discussed together. If there is transparency but no control, what is the benefit of transparency? Similarly, if there is control but no transparency, how can the user utilise the control they have effectively? Financial professionals and

Table 1 How human oversight can build trust in robo-advisors

Trust-building area: Human oversight	
Trust-building method	Specific tips
Oversight of AI model and service it provides	The 'mystery shopper' is still an effective method to check quality and fix issues that could damage trust. The instructions given to AI before each answer it gives must be reviewed and updated regularly. Humans should intervene and tweak the instructions the robo-advisor receives before it replies when there are special events outside the regular routine. Empathetically process complaints. Complaints should be seen as an opportunity to build the relationship, not as a drain on efficiency.
Human explanations for advice and processes	Regularly offer an explanation written by a human to accompany some advice. When there are big changes in the user's life, past performance may not be the best guide so AI must be prepared to adapt. Regularly explain the accountability of humans involved including their names and what measures they take.
Regular human communication outside specific processes	Decide what character the financial institution should have and communicate this consistently. The character of the organisation should be conveyed in terms of human characteristics such as traditional, innovative, or friendly. Regularly communicate how human financial professionals shape the service. Mention their names.

consumers that use robo-advisors are willing to trust robo-advisors under some conditions. They know AI is playing a central role, but they want the process to be transparent. Interacting with a machine for important issues is inherently disempowering, so we start the relationship with a deficit of trust that needs to be dealt with. It needs to be clear what exactly the process is, what data is being used, and what data is being collected. These things need to be done effectively in a way that benefits the user, not a superficial way.

4.2.1 Recognise That There Are Beliefs That Do Not Immediately Lead to Action

The fashion of being 'data-driven' based on users' actions often misses people's feelings as they don't always act on them immediately. However, bad feelings building up can lead to negative actions further down the line. Feelings of a lack of transparency and control may not immediately be apparent in the data that come from measuring behaviour. Consumers can often be trapped in a situation where they do not have better options, so they do not change their behaviour for some time. However, as soon as an alternative appears they will change their behaviour. It is common for people to expect linear change in behaviour, but this is not always the

case. There is often a sudden inflection point in behaviour that came from longer-held beliefs. Feelings are not always easy to understand when the interaction is online; we cannot read the persons face for example, but they are powerful and decisive.

It is important to appreciate that AI effectiveness is not enough on its own, as even with an excellent financial robo-advisor there will be both valid and less valid concerns and fears. All these things are dealt with instinctively when humans interact with each other. A professional in insurance or finance would do these without thinking. When we moved to online services these relationships were weakened but people knew that, ultimately, humans were making the decisions as they made the code and the algorithms. This applies to all five areas of trust building, but it is particularly important for transparency and control.

4.2.2 Explaining Actions and Demonstrating Ability and Integrity

The robo-advisor can make an excellent recommendation, but the user may still feel some anxiety and uncertainty. For example, if AI makes a change to a user's investment portfolio, a short explanation of why this was done is necessary. This is an opportunity for the robo-advisor to explain what methods it uses and how capable it is. If the robo-advisor makes a significant change, then it needs to inform the user. The user must not be left in uncertainty. As far as possible, the explanation must be in plain English avoiding specialised terminology and acronyms.

A separate section on the website can be dedicated to more information on the methodologies behind the actions and recommendations. If the user wants more information beyond the messages they receive, they will be able to find it at this methodology hub. These forms of communication not only reduce uncertainty and the perceived risk involved, but they build more realistic expectations.

4.2.3 Balancing Automation with Control to Avoid Disempowering Users

The robo-advisor must make it easy for the user to be in control by providing a range of tools. These are usually simple but give the user a quick way to take action to implement their will. A slider can enable them to adjust the risk level. Switches can be used to make typical investment exclusions such as weapons manufacturers, or organisations that are considered unethical such as cannabis and alcohol producers. There should also be a button that can be pushed that stops each process such as a new investment until it is checked by the user and confirmed. Additional buttons should enable the user to pause or undo a decision. The line between work and leisure time has been blurred and both professionals and retail investors may be making decisions when they are not entirely in their work mindset. This is the current reality so the financial robo-advisor must make it as easy as possible to understand the implications of decisions and correct mistakes where possible. These features create a psychological safety net for users.

Table 2 How transparency and control can build trust in financial robo-advisors

Trust-building area: Transparency and control	
Trust-building method	Specific tips
Recognise that there are beliefs that do not immediately lead to action	Current user behaviour does not tell the whole story. Bad feelings about the lack of transparency and control can build up and can lead to negative actions further down the line. AI's effectiveness is not enough on its own, as even with an excellent financial robo-advisor there will be both valid, and less valid, concerns and fears. The concerns should be understood, and empathy should be shown.
Explaining actions and demonstrating ability and integrity	An explanation in plain English of all the implications of an action should be made. The methods used to reach the decision and all the implications, such as changes to fees, must be explained. A section on the website can be dedicated to more information on the methodologies behind the recommendations. This not only reduces uncertainty, but it builds more realistic expectations.
Balance automation with control to avoid disempowering users	It must be made easy for the user to be in control by providing a range of tools. The tools give the user a quick way to implement their will. These features create a psychological safety net for users. The user can be encouraged to develop their knowledge where necessary. If they are doing something they have not done before, the robo-advisor can check if they are familiar with the implications.

The robo-advisor can also guide the user to develop their knowledge where necessary. If they are doing something they have not done before, such as buying precious metals, the robo-advisor can check if they are familiar with the implications (Table 2).

4.3 Accuracy and Usefulness

It is not a surprise that the accuracy and usefulness of the information build trust in the robo-advisor. The role accuracy and usefulness play in this context, however, is worth exploring further. Understanding their role in this specific context will make it easier to develop a comprehensive approach to building trust.

4.3.1 Accurate and Relevant Information Conveys Competence and Integrity

The robo-advisor must provide precise, reliable, and relevant information. There are many things an organisation can do to convey their competence, but if the accuracy is not there, all the other measures will be in vain. If there is accuracy, then other

measures, such as communicating the value of the advice, can reinforce the feeling of competence. This is an obvious point but it needs to be made as AI often struggles with accuracy, especially if it is bringing a lot of information together such as the user's risk tolerance, fees, and so on. If there is a lack of accuracy, it may be perceived as a lack of competence, or a lack of integrity, or both. A lack of accuracy in any aspect may be perceived as dishonesty and may damage trust. With the risk involved in financial decisions people are particularly sensitive to the possibility of being misled.

4.3.2 Usefulness Supported by Back Testing and Multiple Methods

Generative AI has impressive capabilities, but the accuracy of the most popular generative AI most people use is achieved by implementing unsupervised and supervised training, along with several special commands. For financial questions, additional layers must be added on top of these typical layers. For example, AI can identify patterns in a wide range of data that influence risk. An example is that some insurers track alternative behavioural data such as a user typing speed, as slow typing has been linked to higher risk. These strengths of AI must be combined with traditional methods and back testing to previous scenarios. Finance and insurance are particularly challenging for generative AI, whose characteristics bring both strengths and weaknesses in this area. There are many unexpected sudden twists in these areas that are hard to predict by AI if it just plays to its natural strengths. The multiple AI models and methods will enhance the usefulness. The saying 'two minds are better than one' could be adapted here to 'multiple AI models and methodologies are better than one'.

4.3.3 Enhanced Usefulness Through Personalisation

While AI has both strengths and weaknesses in making financial predictions, personalising a service is something AI is naturally very strong at. The obvious strength of AI here is bringing data on the user's behaviour together to understand their specific needs. Less obvious is that AI can encourage more interaction from the user, so they keep sharing their needs as they evolve. In addition to measuring behaviour, surveys, and other opportunities to encourage feedback can be used. This dialogue can be elicited when there is unexpected performance or unexpected events. This should not come across as nudges but as opportunities for the user to have their say (Table 3).

4.4 Ease of Use and Support

When there is a big disruption with new technologies and new capabilities emerging, it is tempting to assume that getting the basics right, such as ease of use, is

Table 3 How accuracy and usefulness can build trust in financial robo-advisors

Trust-building area: Accuracy and usefulness	
Trust-building method	Specific tips
Accurate and relevant information conveys competence and integrity	AI often struggles with accuracy. The robo-advisor must provide precise, reliable, and relevant information. If there is a lack of accuracy, it may be perceived as a lack of competence, or a lack of integrity, or both.
Usefulness supported by back testing and multiple methods	Finance and insurance are a particularly challenging for generative AI, whose characteristics bring both strengths and weaknesses in this area. Multiple AI models and methods will enhance the usefulness.
Enhanced usefulness through personalisation	AI brings data on the user's behaviour together to understand their specific needs. AI can encourage more interaction from the user, so they keep sharing their needs as they evolve.

guaranteed. Unfortunately, this is not always the case. There are some issues that have an effect in one context of using a technology but do not have a decisive effect in others. For example, the role of institutions in building trust is a significant factor in some contexts but not others. Ease of use, however, almost always has a significant influence on users' trust and intention to use a technology.

4.4.1 Achieve Ease of Use While Keeping All the Necessary Functionality

As with other trust-building areas discussed here, in principle they are straightforward but when it comes to the implementation things can get tricky. For example, to build trust in the other areas discussed here additional functionality is needed to provide more information to the user and elicit more feedback from them. All this functionality adds complexity which could reduce the ease of use. That is why many financial organisations, especially the born-online banks prioritise simplicity. However, the best way to build trust is to have all the functionality necessary and still manage to keep it simple to use.

By managing to keep everything intuitive, anxiety is reduced and the robo-advisor appears more competent. Users are not always in a critical mood, but when dealing with a lot of information and important financial decisions, it creates anxiety. If the robo-advisor is hard to use, this creates negative emotions and makes the user question whether the aspects of the service they cannot see, the back-office functions, are working correctly. The friction must be reduced as far as possible, and the confidence should be increased. Processes such as onboarding and security checks used to take a lot of time and effort, but the best financial organisations are finding ways to reduce the time and effort needed.

4.4.2 Support Time Management with Nudges

The robo-advisor can give updates and remind the user when they need to check something or if there is something they might be interested in. Effective proactive communication from the robo-advisor benefits trust in several ways but here the purpose is to reduce the cognitive load. The nudges can be based on what usually works for most users but over time it can be personalised to the individual. Milestones can also be celebrated, as this is an additional way to enhance the user's contextual awareness.

4.4.3 Useful and Reliable Service with High Performance

The more useful the robo-advisor is, the more trust it will build and sustain. This does not mean that an extremely useful robo-advisor does not need to build trust in other ways. Trust must be built in all the areas identified, as being weak in some of them will undo the benefit from the others. The usefulness can be enhanced by showing a deeper understanding of the implications of the financial actions on a person's life. Each action can be annotated with a simple explanation. Ideally, the user will feel that the robo-advisor has been sent on a clear mission and will continue to pursue the agreed goals without the need for oversight or micromanaging (Table 4).

Table 4 How ease of use and support can build trust in financial robo-advisors

Trust-building area: Ease of use and support	
Trust-building method	Specific tips
Achieve ease of use while keeping all the necessary functionality	Achieving ease of use by providing a stripped-down basic service is easier to achieve but this means missing other opportunities to build trust. All the trust-building areas must be implemented while still achieving ease of use. By keeping everything intuitive anxiety is reduced and the robo-advisor seems more competent. The friction must be reduced as far as possible, and the confidence should be increased.
Support time management with nudges	The robo-advisor can give updates and remind the user when they need to check something or if there is something they might be interested in.
Useful and reliable service with high performance	The usefulness can be enhanced by showing a deeper understanding of the implications of the financial actions on a person's life. The user should feel that the robo-advisor has been sent on a clear mission and will continue to pursue the agreed goals without the need for oversight or micromanaging.

4.5 Humanlike Interaction

An attempt at humanness that is out of place, or in any other way ineffective, can backfire and even feel creepy. However, with the capabilities of robo-advisors today and the effort that is put into training them and fine-tuning their behaviour, it is feasible to reliably achieve a positive result from humanlike behaviour. Previous research has found that humanlike behaviour should be limited to the interactions where the user asks broader vague questions. A more human, quasi-social interaction does not build trust when a specific question is asked and might seem out of place.

4.5.1 Utilise the Latest Sentiment Analysis and Longest Contextual Memory

The advisor must be updated constantly so it is at the cutting edge of understanding context, sentiment, and intent. This is the 'arms race' between the AI different financial institutions use. The previous conversations must be remembered and linked to current discussions, so the context is clear. Whether voice or text is used to communicate, the sentiment analysis must be able to identify feelings such as frustration and anxiety. The intent should be understood as quickly as possible so the user can use shorter simpler commands and not need extensive prompt engineering, or inputting a whole specification, to get the result they need. This technical prowess is necessary to implement the other trust-building methods in this area.

4.5.2 The Robo-Advisor Should Have a Clear Persona

The robo-advisor must convey an identity and personality. As with many things, the AI is presenting this without it being real, so it is a persona and not a real personality. Each financial institution has a certain image, but its employees still have room to show their different personalities. In a similar way, the robo-advisor must have a persona that fits the institutions image but there is still room for customisation to each user. For example, it is well documented that younger traders often make more regular trades than older traders. The robo-advisor can use data on demographics to customise their persona up to a certain point, and then do the final customisation based on the individual's behaviour. Another option is for humans to prepare several personas for the robo-advisor to choose from based on what the specific user would probably appreciate more.

4.5.3 Build a Relationship and a Feeling of Friendship and Partnership

Inclusive language can be used to give the user the sense that the robo-advisor is on its side. An example is 'Let's review this week's performance'. This inclusive

Table 5 How humanlike interaction can build trust in financial robo-advisors

Trust-building area: Humanlike interaction	
Trust-building method	Specific tips
Utilise the latest sentiment analysis and longest contextual memory	The previous conversations must be remembered and linked to current discussions, so the context is clear. Whether voice or text is used to communicate, the sentiment analysis must be able to identify feelings such as frustration and anxiety. The intent should be understood as quickly as possible so the user can use shorter simpler commands and not need prompt engineering.
The robo-advisor should have a clear persona	The robo-advisor can use data on demographics to customise up to a certain point, and then do the final customisation based on the individual's behaviour. Humans can prepare several personas that the robo-advisor chooses from, based on what the specific user would probably appreciate more.
Build a relationship and a feeling of friendship and partnership	Inclusive language can be used to give the user the sense that the robo-advisor is on its side. An example is 'Let's review this week's performance'.

language makes the user feel like they are partners with the robo-advisor and it is not just a tool. The user will not perceive this as fake but as empathetic. As long as this is not taken too far this is effective (Table 5).

5 Conclusion

Previous research has identified four issues that affect trust in robo-advisors for specific financial questions, and five that affect trust in robo-advisors for vague financial questions (Zarifis and Cheng 2024). The four issues that affect trust in all financial questions are human oversight, transparency and control, accuracy and usefulness, ease of use, and support. The additional, fifth issue that builds trust in vague questions is humanlike interaction. Knowing these five areas is helpful, but in order to take action, specific methods to build trust for each of them are necessary. Therefore, this research identified the specific methods to build trust in robo-advisors for each of the five areas.

This research finds that building trust with human oversight can be done in three specific ways: (1) oversight of AI model and service it provides, (2) human explanations for advice and processes, and (3) regular human communication outside specific processes. Building trust with transparency and control can be done by: (1) recognise that there are beliefs that do not immediately lead to action, (2) explain actions and demonstrating ability and integrity, and (3) balance automation with control to avoid disempowering users. Building trust with accuracy and usefulness can be done by: (1) accurate and relevant information conveying

competence and integrity, (2) usefulness supported by back testing and multiple methods, and (3) enhanced usefulness through personalisation. Building trust with ease of use and support can be done by: (1) achieve ease of use while keeping all the necessary functionality, (2) support time management with nudges, and (3) useful and reliable service with high performance. Building trust with a humanlike interaction can be done by: (1) utilise the latest sentiment analysis and longest contextual memory, (2) the robo-advisor should have a clear persona, and (3) build a relationship and a feeling of friendship and partnership.

Yet these trust-building methods do not operate in isolation. Algorithmic bias and ethical trade-offs represent systemic risks that can undermine client confidence if left unaddressed. Bias can enter robo-advisory systems through data, features, model design, and deployment choices, producing suitability failures, excess trading costs, and inappropriate drawdown exposure. Practitioners must also navigate genuine tensions between personalisation and fairness, transparency and proprietary interest, automation and human oversight, and predictive accuracy and interpretability. The governance architecture of model risk management, suitability assurance, fairness monitoring, and documentation provides the institutional foundation without which the practitioner strategies identified in Sect. 4 cannot be sustained at scale.

This analysis carries implications for investors, firms, and regulators alike. Investors benefit from understanding that robo-advisory trust is multidimensional and that the presence of sophisticated technology does not guarantee aligned interests. Questions about oversight, transparency, and fairness monitoring are reasonable to ask and material to the advisory relationship. Firms benefit from recognising that trust-building is not merely a marketing exercise but requires genuine investment in governance infrastructure, bias testing, and explainability. The competitive advantage of robo-advisory, scalable personalised advice at low cost, can only be sustained if clients believe the system operates in their interest. Regulators benefit from the framework's emphasis on outcomes rather than prescriptive rules. Suitability, fairness, and transparency can be assessed through the lens of client experience and distributional analysis, without requiring detailed algorithmic audits that may be impractical at scale.

The chapter began by observing that trust, built naturally through face-to-face interaction in traditional advisory relationships, must be deliberately constructed in the age of AI-driven financial services. The trust-building areas identified in Sect. 2 and operationalised in Sect. 4 provide the conceptual and practical architecture for this construction. They offer a comprehensive approach to building financial robo-advisory systems that are not merely efficient and scalable, but genuinely worthy of client confidence.

We thank Dr Peng Xu for proofreading the paper.

References

Braun V, Clarke V (2006) Using thematic analysis in psychology. Qual Res Psychol 3(2):77–101. https://doi.org/10.1191/1478088706qp063oa

Cheng X, Bao Y, Zarifis A, Gong W, Mou J (2021) Exploring consumers' response to text-based chatbots in e-commerce: the moderating role of task complexity and chatbot disclosure. Inter Res 21. https://doi.org/10.1108/INTR-08-2020-0460

Cheng X, Yan Y, Zarifis A (2024) Understanding users' response to Chatbots from the perspective of interactive ritual Chain. In: Americas Conference on Information Systems (AMCIS), 1–5

Chui M, Hazan E, Roberts R, Singla A, Smaje K, Sukharevsky A, Yee L, Zemmel R (2023) The economic potential of generative AI (Issue June)

Dietzmann C, Jaeggi T, Alt R (2023) Implications of AI-based robo-advisory for private banking investment advisory. J Electr Bus Dig Econ 2(1):3–23. https://doi.org/10.1108/jebde-09-2022-0037

Fügener A, Grahl J, Gupta A, Ketter W (2021) Will humans-in-the-loop become Borgs? Merits and pitfalls of working with AI. MIS Quart 45(3):1527–1556. https://doi.org/10.25300/MISQ/2021/16553

Jain R, Kumar S, Sood K, Grima S, Rupeika-Apoga R (2023) A systematic literature review of the risk landscape in Fintech. Risks 11(2). https://doi.org/10.3390/risks11020036

Lankton N, McKnight H, Tripp J (2015) Technology, humanness, and trust: rethinking trust in technology. J Assoc Inf Technol 16(10):880–918

McKnight H, Chervany NL (2002) What trust means in E-Commerce customer relationships: an interdisciplinary conceptual typology. Int J Electr Commer 6(2):35–59

Miles MB, Huberman AM (1994) Qualitative data analysis. Sage Publications

Mori M (2012) The Uncanny Valley: the original Essay by Masahiro Mori. IEEE Robot Automat Magaz 12(Figure 1):1–6. https://spectrum.ieee.org/automaton/robotics/humanoids/the-uncanny-valley

Rajaobelina L, Prom Tep S, Arcand M, Ricard L (2021) Creepiness: its antecedents and impact on loyalty when interacting with a Chatbot. Psychol Mark 38(12):2339–2356. https://doi.org/10.1002/mar.21548

Zarifis A, Cheng X (2024) How to build trust in answers given by Generative AI for specific and vague financial questions. J Electr Bus Dig Econ 3(3):236–250. https://doi.org/10.1108/JEBDE-11-2023-0028