

Case Study: Creating a responsible personalized AI finance tool



Why you should use this resource: Teams working on products that use AI or ML need to be aware that even if their products do not contain racial labels in their input data, the technology can still learn to identify and discriminate on the basis of race. This fictional case study is designed for anyone who works on Artificial Intelligence (AI) and Machine Learning (ML) systems: from developers and data labelers, to writers and product managers. It will help you identify places where bias can enter tools that use AI or ML technology and provide some good practices to work towards mitigating that bias. It will also help you identify ways in which equity in AI and ML can contribute to your organization's core mission and values.

Who this is for: Anyone who is working on AI or ML, whether as a developer, data labeler, writer, or product manager. This activity is designed as a facilitated conversation between team members, but it can also be done independently.

What you will have after reviewing this case study:

- A deeper understanding of how racial bias can arise in tools that utilize AI or ML technology
- A set of good practices that you can use to work towards mitigating bias in AI and ML tools

Note: All materials for this case study can be found in the folder [Case Study: Creating a responsible personalized AI finance tool](#). Please feel free to personalize these materials in the spaces labelled for personalization, but be sure to maintain author credits, image credits, and the branding and theme throughout.

Steps

One week before

1. Schedule a meeting to facilitate a Case Study on identifying racial bias in a personalized finance tool (see the following pages). We recommend scheduling 2 hours. You may facilitate this session yourself or bring in an experienced facilitator.
 - Ask participating team members to review the [Guide for having difficult discussions about race & identity in AI/ML research & development](#) and key understandings from our guide, [Towards Responsible Language in AI and ML](#). Emphasize to participants that reviewing the materials is necessary for being able to participate productively in the session.
 - Share the invitation to the event — including the purpose of the conversation — at least one week in advance, if possible.
2. If you are facilitating the session be sure to prepare. Review EGAL's [Guide for having difficult discussions about race & identity in AI/ML research & development](#) and our

guide, [Responsible Language in AI and ML](#), along with the session & facilitator notes in the following pages.

3. Review the [slide deck](#). Make a copy of the slides and add any personalizations. Gather information on other resources available to participants to enhance their ability to design equitable products. For example, is there a team at your organization whose focus is equity and inclusion? Be ready to suggest relevant resources specific to your organization. Make sure the slide deck permissions are configured so that participants will be able to edit the slides during the session; these can be changed by clicking on "Share" in the upper right hand of Google slides.

Day of event

Provide a copy of the [Case Study Worksheet: Creating a responsible personalized AI finance tool](#) (also found on pages 11-12 below) to each participant. If the session will be held in person, print a copy of pages 11-12 for each participant. Print one copy of pages 13-15 (you will distribute these pages to three volunteers for the role play). If the session will be held virtually, provide participants with the link to the [Case Study Worksheet: Creating a responsible personalized AI finance tool](#).

4. Follow the tasks outlined in the session notes and facilitate the discussion.

5. Debrief:

- Take notes for your personal record about the experience for you as a team leader.
- Share with other project managers how it went (without sharing personal information disclosed by team members during the session).
- Remember insights shared during this session when difficult situations arise in the future.

One week after

6. Follow up:

- Check in with participants one week after the case study to find out whether or not they were able to meet the goals they set for themselves and to help problem-solve any barriers they faced. Invite participants to check in with one another as well. You might ask participants, "Which of your goals have you been able to meet? What barriers have you faced to reaching those goals? How can you adapt your goals so that they are realistic and serve your current needs?"
- Repeat these check-ins after one month, three months, and six months.

Case Study Facilitator's Guide: Creating a responsible personalized AI finance tool

What is this case study? The fictional case study below outlines a scenario related to racial bias in Artificial Intelligence (AI). As a product team, you are currently working to incorporate AI into your products in a way that is equitable and responsible. In this fictional case study, you will:

- Learn more about some of the ways that racial bias can emerge from AI systems
- Apply some strategies identified as best practices to mitigate racial biases in AI

Why is this important? AI can widen existing inequities, and one way to mitigate such harmful effects is to better recognize how they arise and understand the long-term impacts on marginalized populations. However, bias and inequitable treatment are not always easy to spot. This fictional case study will help you and your team identify and mitigate places where inequity may emerge so that you can pursue greater equity and positive outcomes in your products.

Scenario:

You are on a team developing a tool that helps users manage their finances. You are currently integrating a new chatbot feature into the system that will serve as an automated financial advisor. The chatbot will answer users' questions (which can be entered using text or speech) and provide them with information and suggestions about how they might want to invest their money. The chatbot is underpinned by AI technology so that it can dynamically respond to the needs of individual users, based on a variety of factors, including what they ask the chatbot and any financial and demographic information they have shared with the tool.

In a pilot test with users of diverse geographical and racial backgrounds, users commented that they felt like the chatbot understood them and made them feel comfortable. Some users identified specific words and phrases that the chatbot used that reflected their own use: if a user said "I ain't", the chatbot would use "I ain't" but if a user used "I'm not", it would use "I'm not."

However, you and your teammates develop concerns about whether or not the chatbot is treating all users fairly as it learns from them. For example, you notice that users who use the phrase "I ain't" are more likely to be shown short-term, low-yield investment options while users who use the phrase "I'm not" are more likely to be shown long-term, high yield investment options.

Seeing these results, your team was uneasy about how the chatbot might be subtly influencing different users' financial decisions by giving different users different suggestions. You decided to schedule a conversation about ways in which the chatbot could be perpetuating inequity as a result of the questions it determined were best for each user.

Task:

1. **Introductions & goals** (SUGGESTED TIME: 10 MINUTES)

- Tell participants what your goals are, as the facilitator, and explain the outline of the session.
- If people do not know each other, have them introduce themselves. Ask everyone to think individually for about one minute: What is one fear or concern you have about engaging in this conversation? What is one thing you hope to learn? What is one thing that is bringing you joy at the moment? After one minute, have each person share their fear/concern, what they hope to learn, and what brings them joy, along with their name/pronouns/roles and any other information the facilitator thinks would be useful.

2. **Establish community agreements.** (SUGGESTED TIME: 10 MINUTES)

- Tell participants that you will now work together to co-create a community agreement. A community agreement includes rules of conduct for the training. Community agreements engage participants in the development of these ground rules so that participants take ownership of the rules and are responsible for them—individually and as a group.
- For virtual sessions, use the slide titled “Community agreements for today’s conversation” and take notes based on the conversation. Be sure to share your screen and the link to the slides. For in person settings, use a flipchart and write “Community Agreement” at the top.
- Ask participants: What are some good rules or principles that we should set as a group for the space?
- Have participants share and write them as they share.
- Add some suggestions to the community agreement yourself. Such as:
 - Engage in brave conversations. (This discussion is likely to bring up difficult topics that might be painful for some people. Be thoughtful about what you say and be open to the fact that even comments that have good intentions can have negative impacts. Try to understand others’ perspectives and be open to feedback.)
 - Take space, make space. (Allow everyone to have time to share and react. Be brave and speak up, but also step back and allow others time to contribute.)
 - Share your own experiences and honor others’ experiences.
 - Practice active listening.
 - Honor confidentiality.
 - Expect and accept non-closure.
 - (For virtual settings) Keep video on to the extent that is comfortable or possible for you, and remain focused (refrain from checking email / scrolling online).
- Ask participants: Is there anything you would like to take off this list, or anything else to add to it?
- Tell participants: Please raise your hand (or give a thumbs up in a virtual setting) if you agree to this community agreement. If in a physical space, tape the community agreement to the wall in a place where people can see it.

3. **Distribute the case study worksheet.** (SUGGESTED TIME: 5 MINUTES)

If you are facilitating this session virtually, ask participants to create a copy (File, Create a Copy) or download the document so that they can take notes. Have participants individually read the case study.

4. **Tell participants that we will discuss the scenario.**

(SUGGESTED TIME: 12 MINUTES)

Put participants in groups of 3-4 to discuss for 12 minutes and tell them to discuss the following questions. Remind participants to focus on issues of racial inequity that might arise in the tool. If you are using the slide deck, participants can take notes on slide 7 (group 1) and slide 8 (group 2).

- The chatbot learns from and uses natural human language. What are ways bias could emerge as a result of learning from and using human language?
- What are other ways that bias might come into play?

5. Discuss as a whole group. (SUGGESTED TIME: 18 MINUTES)

Allow participants to voice their own thoughts and ideas, but be sure that the following points are raised:

- While a personalized chatbot may result in greater user engagement (i.e. number of transactions and positive reviews), it has the potential to reinforce inequities by suggesting different financial products to people with different profiles. This may result in wealthy users being given options that will further increase their wealth while less-wealthy users may see less lucrative options.
- The data accessible to the chatbot contains many variables that could serve as proxies for protected social categories, including race. For example, geographical variables (zip code, school district) may reflect historical biases due to housing discrimination. The amount of wealth and how that wealth is invested may also be the result of historical discrimination that has systematically prevented Black people from acquiring wealth as easily as White people.
- The machine might learn that certain identities tend to be less wealthy or more risk averse in investing. It might then treat an individual in a way that is potentially harmful and untrue for a particular individual. For example, it might assume that a Black person is less wealthy than a White person and offer different (less lucrative) options, or it might assume that a woman is less financially literate than a man.
- In addition to demographic information, the chatbot might use language as a proxy for race. Individuals' language use can reveal information about their identities. For example, if an individual uses features of African American English, the chatbot may infer that they are Black; if an individual uses words that are more frequently used by women, it may infer that they are a woman. Just as we often make assumptions about people's identities based on their voices (think about who you picture on the other end of a telephone call, for example), the chatbot can also make assumptions.
- Even if the intentions are to promote equity, using identity categories can be problematic. For example, if you want to provide support to people with low financial literacy, you might be tempted to assume that people of color have generally been less exposed to financial literacy tools, and provide support to any users identifying as people of color. However, this may further entrench stereotypes. It would be more accurate to use a tool through which users can self-identify as lacking financial literacy skills (check out some recommendations [here](#)). On the other hand, if your goal is to support users of color (for example, offering grants specifically for that demographic), using self-identification of race would be appropriate.
- In designing this product, there is an opportunity to increase equity. The tool has the potential to provide a less-biased experience than what might be possible for users who visit local brick and mortar financial institutions. However, the team must work proactively towards this goal. Focusing on impact is crucial at every step of the way.
- If people are hesitant to discuss race and anti-Blackness directly, point this out: "I noticed that we've been talking about how "specific groups" might end up getting targeted for different services. Can we practice being more specific? Are we

talking about Black women? People who identify as gender non-conforming?”

6. Now that the group has identified how bias can come into play, tell participants that it's time for a role play to figure out how to discuss it as a team and move forward.

7. Ask for volunteers for a role play. You will need 3 volunteers.

(SUGGESTED TIME: 30 MINUTES)

- Tell participants that the role play will be about the conversation regarding which landing page to use. Let them know that the purpose of the role play is to focus on talking about issues of racial bias in the tool and figure out how to move forward as a team.
- Assign the volunteers to one of the roles below. Privately share the following information with the volunteer playing the role. Assign the volunteers to one of the roles below. Privately share the following information with the volunteer playing the role. You may print out these roles (pages 15-17) if you are hosting this session in person. If you are hosting this session virtually, participants can click the links to the roles provided on slide 11 and/or you can copy & paste the following into the chat for the session:

[Project manager](#)

[Team member #1](#)

[Team member #2](#)

- Give the person 1 minute to read their role. While the volunteers are preparing for the role play, have everyone else review the **[Guide for having difficult discussions about race & identity in AI/ML research & development](#)**. Allow role players to act out the scenario for about 10 minutes.
- **Project manager:**
 - You are the project manager for this tool. You are up for a promotion in the coming weeks. You have been pressured to make sure that the company delivers products first-to-market. You recognize there can be ethical concerns with ML products, but are focused on business impacts first and foremost. At the same time, you are concerned about the risk that a biased AI system can pose for the company (and your potential promotion). In the role play you must:
 - » Say the following points:
 - I understand your concerns, but let's remember to maintain positivity and recognize how important this product is, rather than getting stuck focusing on problems.
 - I agree that we need to avoid racial discrimination, but how is what we're seeing here different from class discrimination, or gender discrimination, or anything else? Are we proposing to treat everyone the same, even though they have different backgrounds and needs?
 - Making this product more equitable is important, but isn't our goal to provide a useful tool that's first-to-market?
 - Remind others that this is an important product for the company and push back about making too many changes that would slow its use and uptake among other customers.

- » Bring up the experiences of women at the organization.
- » Bring up your own identity-based experiences (e.g., “As a White woman...” / “As a Black man...” / “As a Deaf person...”)
- » Tell others that you are concerned about the risk this product poses for the company.
- You will have 10 minutes to talk with the team members and must agree on what you will show on the landing page when you launch the product.
- **Team member #1**
 - You were involved in the pilot study and think that the best way forward is to allow the chatbot to learn from users and provide a personalized experience. You think that equity is important, but you worry that you will lose opportunities for revenue if you prioritize equity over everything else. You would rather launch a product that satisfies at least some clients’ needs and then make incremental improvements after launch. In the role play, you must:
 - » Say points including:
 - But the clients are the ones who are ultimately making choices about what financial products to buy. They can always search for something that wasn’t suggested by the chatbot, and the chatbot will learn from that and update their page in the future.
 - If we don’t personalize chatbot, we’re going to lose clients to competitors who could have even more troublesome products.
 - People are different from one another; what’s biased about responding to those differences with personalized treatment?
 - » Bring up your own identity-based experiences (e.g., “As a White woman...” / “As a Black man...” / “As a Deaf person...”)
 - » Get defensive, such as saying, “I’m not racist, I just think you’re being oversensitive.”
 - You will have 10 minutes to talk with the product manager and the other team member and must agree on what landing page will be shown to clients when the product launches.
- **Team member #2**
 - You were involved in the pilot study and think that the best way forward is to show clients standardized options for financial products rather than letting the AI chatbot make personalized suggestions. You are worried about losing potential clients who don’t feel that their experience is personalized, but you are more worried about reproducing financial inequality. In the role play, you must:
 - » Say points including:
 - I’m sure that no one intends for the chatbot to lead to different financial decisions by race, but that’s what seems to be happening. I think we should really consider the impact of that. How are these choices affecting our clients’ life outcomes?
 - It’s exciting to see that the AI-powered chatbot might help bring in more clients, but I’m wondering if it’s fully aligned with the mission of our organization.
 - I wonder if over time, focusing on equity in our design of a chatbot will have a positive impact on client retention and revenue. There’s research suggesting that equity is a financially smart investment in the long run.
 - One of our goals with this tool was to make a diverse set of financial decisions available and understandable for everyone. I’m wondering if using a personalized AI-powered chatbot makes some decisions less accessible for our clients, and particularly our Black clients.
 - If White clients are being shown more lucrative financial options than

- Black clients, that can have a real-world impact on racial inequity.
- » Stay focused on race: If others shift the conversation away from race, bring the conversation back, using statements like, “I agree that gender/age/class is important, but let’s also think about race.”
- » Bring up your own identity-based experiences (e.g., “As a White woman...” / “As a Black man...” / “As a Deaf person...”)
- » Use “I” statements: “I am wondering...” / “I’m feeling...” / “I’m confused about...”
- » Paraphrase what others have said: “What I hear is...” / “If I’m understanding you, you’re saying...”
- » Ask clarifying questions: “When you said...did you mean...? / “Can you say more...”
- » Empathize: “I can understand why you feel...”
- You will have 10 minutes to talk with the product manager and the other team member and must agree on what landing page will be shown to clients when the product launches.
- Ask each role play participant: How did it feel to be in that particular role?
(SUGGESTED TIME: 3 MINUTES)
- Ask participants: (SUGGESTED TIME: 10 MINUTES)
 - What aspects of the conversation went smoothly? What aspects of the conversation were tense?
 - What good practices or principles did the participants use? (Hint - see the [Guide for having difficult discussions about race & identity in AI/ML research & development](#))
 - What were you surprised to hear during the conversation?
 - What would you have liked to add to the conversation?
 - What did you learn about communicating successfully with your colleagues on the topic of race that you would like to remember when you encounter similar difficult discussions on one of your teams?

8. Review the summary of good practices & guiding principles on slide 13 and any resources specific to your organization on slide 14.

(SUGGESTED TIME: 5 MINUTES)

9. Tell participants we’ll end by setting some “SMART” goals to take one or more things they’ve learned about talking about race from the session that they can apply in their daily practices.

(SUGGESTED TIME: 5 MINUTES)

Have participants read the suggestions for setting “SMART” goals and have them write down their goal(s).

SMART goals are Specific, Measurable, Achievable, Realistic, and Timely. For example, you might set a goal like:

- “This week, I will talk to my coworker about this issue that’s been on my mind. I’ll use “I” statements and lead with how I understand why they said what they said, but that I’m still feeling the negative impact.”
- “I will take 15 minutes at the next group meeting to engage in a brainstorm about possible equity issues in our product design. If people push back, I’ll justify the

importance of equity using this case study as an example.”

10. Go around the group inviting each participant to share back at least one goal with the group. (SUGGESTED TIME: 10 MINUTES)

Encourage everyone to share their goal or a learning from the session plus one word that represents how they are feeling. Ask participants to share and select who will share after them.

11. Thank participants for their time and encourage them to check in with one another on their goals. (SUGGESTED TIME: 5 MINUTES)

Facilitator Notes:

It can be difficult to talk about issues of racial bias that might occur in products that you and your team have worked hard to develop. Here are some tips to remember (and to share with your teammates) to keep the conversation going (these can also be found in our [Guide for having difficult discussions about race & identity in AI/ML research & development](#)):

- Separate intent from impact: Try to assume that individuals have good intentions, but don't allow that to diminish any negative impacts. Let them know that you're concerned about impact but not judging them: "I assume it wasn't your intent for the product to discriminate, but it's having a negative impact on Black users, so let's figure out how to fix this." Try using "I" statements to keep the conversation on track: "I have a different perspective... / I appreciate how thoughtful you were in the design phase, but I think this issue can still be addressed..." If you realize that one of your actions had a negative impact — even if it was unintentional — recognize the impact: "I'm sorry. I see how that choice resulted in a negative impact for our users. Let's see how we can address it."
- Listen to different perspectives and types of data: Both quantitative and qualitative data are valuable for understanding ways in which bias can emerge in AI. In this scenario, the qualitative (e.g. comments from users) and quantitative (e.g. data on which financial purchases were made) data from the pilot study might tell different stories. During the conversation, encourage all participants to listen to each other before jumping in with their own thoughts.
- Stay with the issue of racism: Because talking about race can be uncomfortable, people may try to shift the conversation to issues they may be more comfortable with, such as gender or socio-economic status. While issues related to equity in these areas are also important, so are issues of race. Be precise and stick with race — or intersections between race and other categories — during conversations about race. For example, imagine that a White woman speaks up during a conversation about anti-blackness to note that women also face barriers. While marginalization of women is an important issue to address, in this situation it might be serving to shift the conversation from anti-blackness (which is outside of some people's conversational comfort zone) to gender (which is a more comfortable topic for some). In this case, we want to make sure that the conversation about race isn't avoided. You might redirect the conversation by saying:
 - "I agree that gender is another important factor. Can you help me understand why gender is standing out to you in a different way from race? Are there ways in which the effects gender and race intersect or are distinct in this situation?"
 - "All of these issues of inequity are important, so let's make sure we address them all. Right now, I'd like us to continue this discussion of race."

If participants need additional information or support to identify ways in which bias can emerge

through the use of language in AI, share some of the information below (which can also be found in our guides, [Responsible Language in AI and ML](#) and [Mitigating Bias in Artificial Intelligence Playbook](#)):

- Input data may be biased: If Black people have historically been encouraged by financial institutions to buy certain financial products (or have been discouraged from buying others), using previous financial history to train the model will result in similar biases against Black users. If the model is trained on human data, it is likely to replicate the same biases that are present in the input data.
- Social factors can be revealed through language: In addition to overt mentions of social categories like race and gender (such as sharing self-identification of race or gender with the chatbot), other words may also subtly indicate information about the user. For example, men and women may use different verbs when describing their activities (i.e. men are more likely to say things like, “I executed this transfer” while women are more likely to say, “I helped complete this transfer”); this difference can then be learned by the AI system as a proxy to determine gender. Features of regional or social language varieties (such as African American English, Chicanx English, or Southern American English) can also be learned by the chatbot and used to group people into social categories. Just because overt mentions of social categories have been removed from the data doesn’t mean all bias has been eliminated.
- Changing language helps, but it’s not enough: Making the language used in communications like the user interface equitable is an important step, but it’s not sufficient. It’s also important to address racial inequity and injustice perpetuated beyond language. For example, working to remove biases from the product landing page is a good start, but if the AI chatbot is learning to treat users differently based on historical barriers between racial groups, the service will still remain inequitable.

This case study is an accompanying resource to the guide, [Responsible Language in AI & ML](#). It was authored by Julia Nee, Genevieve Smith, Ishita Rustagi, and Alicia Sheares with the Center for Equity, Gender & Leadership (EGAL) at the UC Berkeley’s Haas School of Business. It benefited from invaluable feedback and contributions from practitioners at leading tech companies and in academia, including Clementine Collett, Antoni Lewa, Dr. Kellie McElhaney, and Dominique Wimmer. We respectfully acknowledge that this work has been developed at UC Berkeley, which sits on unceded Ohlone land.



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

Case Study Worksheet: Creating a responsible personalized AI finance tool — worksheet

What is this Case Study? The fictional case study below outlines a fictional scenario related to racial bias in Artificial Intelligence (AI). As a product team, you are currently working to incorporate AI into your products in a way that is equitable and responsible. In this fictional case study, you will:

- Learn more about some of the ways that racial bias can emerge from AI systems
- Apply some strategies identified as best practices to mitigate racial biases in AI

Why is this important? AI can widen existing inequities, and one way to mitigate such harmful effects is to better recognize how they arise and understand the long-term impacts on marginalized populations. However, bias and inequitable treatment are not always easy to spot. This fictional case study will help you and your team identify and address places where inequity may emerge so that you can pursue greater equity and positive outcomes in your products.

Scenario:

You are on a team developing a tool that helps users manage their finances. You are currently integrating a new chatbot feature into the system that will serve as an automated financial advisor. The chatbot will answer users' questions (which can be entered using text or speech) and provide them with information and suggestions about how they might want to invest their money. The chatbot is underpinned by AI technology so that it can dynamically respond to the needs of individual users, based on a variety of factors, including what they ask the chatbot and any financial and demographic information they have shared with the tool.

In a pilot test with users of diverse geographical and racial backgrounds, users commented that they felt like the chatbot understood them and made them feel comfortable. Some users identified specific words and phrases that the chatbot used that reflected their own use: if a user said "I ain't", the chatbot would use "I ain't" but if a user used "I'm not", it would use "I'm not."

However, you and your teammates develop concerns about whether or not the chatbot is treating all users fairly as it learns from them. For example, you notice that users who use the phrase "I ain't" are more likely to be shown short-term, low-yield investment options while users who use the phrase "I'm not" are more likely to be shown long-term, high yield investment options.

Seeing these results, your team was uneasy about how the chatbot might be subtly influencing different users' financial decisions by giving different users different suggestions. You decided to schedule a conversation about ways in which the chatbot could be perpetuating inequity as a result of the questions it determined were best for each user.

Discussion Questions:

1. The chatbot learns from and uses natural human language. What are ways bias could emerge as a result of learning from and using human language?
2. What are other ways that bias might come into play?

SMART goals are Specific, Measurable, Achievable, Realistic, and Timely. For example, you might set a goal like:

- "This week, I will talk to my coworker about this issue that's been on my mind. I'll use "I" statements and lead with how I understand why they said what they said, but that I'm still feeling the negative impact."
- "I will take 15 minutes at the next group meeting to engage in a brainstorm about possible equity issues in our product design. If people push back, I'll justify the importance of equity using this case study as an example."

What is at least one thing you've learned about talking about race from this case study that you plan to incorporate into your work?

Project Manager



You are the project manager for this tool. You are up for a promotion in the coming weeks. You have been pressured to make sure that the company delivers products first-to-market. You recognize there can be ethical concerns with ML products, but are focused on business impacts first and foremost. At the same time, you are concerned about the risk that a biased AI system can pose for the company (and your potential promotion). In the role play you must:

- Say the following points:
 - I understand your concerns, but let's remember to maintain positivity and recognize how important this product is, rather than getting stuck focusing on problems.
 - I agree that we need to avoid racial discrimination, but how is what we're seeing here different from class discrimination, or gender discrimination, or anything else? Are we proposing to treat everyone the same, even though they have different backgrounds and needs?
 - Making this product more equitable is important, but isn't our goal to provide a useful tool that's first-to-market?
- Remind others that this is an important product for the company and push back about making too many changes that would slow its use and uptake among other customers.
- Bring up the experiences of women at the organization.
- Bring up your own identity-based experiences (e.g., "As a White woman..." / "As a Black man..." / "As a Deaf person...")
- Tell others that you are concerned about the risk this product poses for the company.

You will have 10 minutes to talk with the team members and must agree on what you will show on the landing page when you launch the product.

Team member #1: Advocate of AI-powered chatbot



You were involved in the pilot study and think that the best way forward is to allow the chatbot to learn from users and provide a personalized experience. You think that equity is important, but you worry that you will lose opportunities for revenue if you prioritize equity over everything else. You would rather launch a product that satisfies at least some clients' needs and then make incremental improvements after launch. In the role play, you must:

- Say points including:
 - But the clients are the ones who are ultimately making choices about what financial products to buy. They can always search for something that wasn't suggested by the chatbot, and the chatbot will learn from that and update their page in the future.
 - If we don't personalize chatbot, we're going to lose clients to competitors who could have even more troublesome products.
 - People are different from one another; what's biased about responding to those differences with personalized treatment?
- Bring up your own identity-based experiences (e.g., "As a White woman..." / "As a Black man..." / "As a Deaf person...")
- Get defensive, such as saying, "I'm not racist, I just think you're being oversensitive."

You will have 10 minutes to talk with the product manager and the other team member and must agree on what landing page will be shown to clients when the product launches.



Team member #2: Advocate of standardized landing page

You were involved in the pilot study and think that the best way forward is to show clients standardized options for financial products rather than letting the AI chatbot make personalized suggestions. You are worried about losing potential clients who don't feel that their experience is personalized, but you are more worried about reproducing financial inequality. In the role play, you must:

- Say points including:
 - I'm sure that no one intends for the chatbot to lead to different financial decisions by race, but that's what seems to be happening. I think we should really consider the impact of that. How are these choices affecting our clients' life outcomes?
 - It's exciting to see that the AI-powered chatbot might help bring in more clients, but I'm wondering if it's fully aligned with the mission of our organization.
 - I wonder if over time, focusing on equity in our design of a chatbot will have a positive impact on client retention and revenue. There's research suggesting that equity is a financially smart investment in the long run.
 - One of our goals with this tool was to make a diverse set of financial decisions available and understandable for everyone. I'm wondering if using a personalized AI-powered chatbot makes some decisions less accessible for our clients, and particularly our Black clients.
 - If White clients are being shown more lucrative financial options than Black clients, that can have a real-world impact on racial inequity.
- Stay focused on race: If others shift the conversation away from race, bring the conversation back, using statements like, "I agree that gender/age/class is important, but let's also think about race."
- Bring up your own identity-based experiences (e.g., "As a White woman..." / "As a Black man..." / "As a Deaf person...")
- Use "I" statements: "I am wondering..." / "I'm feeling..." / "I'm confused about..."
- Paraphrase what others have said: "What I hear is..." / "If I'm understanding you, you're saying..."
- Ask clarifying questions: "When you said...did you mean...?" / "Can you say more..."
- Empathize: "I can understand why you feel...."

You will have 10 minutes to talk with the product manager and the other team member and must agree on what landing page will be shown to clients when the product launches.