Introduction

Welcome, facilitator!

To plan and run a People + AI Guidebook workshop you will need:

- These slides
- The Facilitator Guide

The workshop is made to run remotely. It can be entirely conducted from these slides.

Throughout, you'll see **green facilitator notes.** They provide information for you to plan and set up the workshop. Think of these as stickers you can delete when you're done with them.

Facilitator note

This is a facilitator note about facilitator notes.

You can delete this whole slide when you're done with it.

Delete this note

Table of contents

Facilitator note

This table of contents is to help facilitators plan with planning.

You can delete this whole slide when you're done with it.

Delete this note

Your workshop name

A People + AI Guidebook workshop

Google

Peop

Facilitator note

This is your title slide. Fill in the name of your workshop.

Delete this note

Our challenge

Challenge statement goes here

Google

Peop

Facilitator note

This is where you paste your workshop challenge statement. See the facilitator guide for more on how to craft the statement.

Delete this note

Meet the team

Firstname Lastname Role	Firstname Lastname Role	Firstname Lastname Role
Firstname Lastname Role	Firstname Lastname Role	Firstname Lastname Role
Firstname Lastname Role	Firstname Lastname Role	Firstname Lastname Role
Firstname Lastname Role	Firstname Lastname Role	Firstname Lastname Role
Firstname Lastname Role	Firstname Lastname Role	Firstname Lastn

Facilitator note

Fill this out with your workshop participants' names and roles.

Delete this note

Agenda overview

Day 1

Day 2

Understand

Case studies
Critical moments

Define

Opportunity statements

Sketch

Errors
Explainability
Feedback
Controls
Onboarding
Trust calibration
Data headlines

Decide

Review questions

Prototype & validate
On your own

Facilitator note

Feel free to modify this to suit the needs of your workshop.

Delete this note

Day 1 agenda

9:00	Arrival & introduction
9:30	Explainability case studies
10:30	**Break**
10:45	Critical moment introduction
11:05	Map guiding questions
11:25	Write opportunity statements
12:00	**Break**
12:15	Share with large group
12:45	Overflow time
1:00	**End of day 1**

Facilitator note

This is the default Day 1 agenda for a 2-day workshop. Modify to suit your workshop's needs. More info on agendas in the facilitator guide.

Delete this note

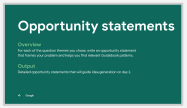
Setup

Introduction

How to use these slides

Follow along with the instructions to complete the activities.

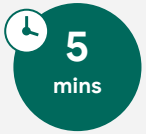
Note the following patterns of how the slides are laid out.



Sections
Green slides. Denote a module of the workshop.

Activities
White slides with green timing dots. A timed activity.

Setup
Gray slides (like this one). Resources and context for each activity.



Timing dots
Each activity will have the timing listed in a dot on the top right.

Google

People + AI Guidebook

Case studies

Goal

Warm up to thinking about applying these lessons to your product.

Instructions

Read the scenarios linked from the slides below.
Answer the questions as a group and take notes on the slides.

Case studies

Get started with these AI scenarios

The following case studies and discussion questions are meant to help you step out of your specific product context and first think about the implications of AI-based systems more broadly.

Things to know

These are not “best case” scenarios
These cases are explicitly not written to present a best practice scenario. They’re designed for you to consider the design of AI explainability outside of your team’s day-to-day product context.

The Model-U
The cases follow the development of a hypothetical car with integrated AI systems called the Model-U by Intelligent Engines.

You’ll be introduced to the car’s various features and a number of customers and other stakeholders.

Self-driving features
The cases portray various situations with the Model-U including its reliable self-driving features, but lack of regulatory approval for the car to self-drive anywhere but on approved highways.

Scenario 1

5 mins

Read the scenario

Link to scenario

Facilitator note

The default workshop agenda plans for 2 scenarios. Guidance for choosing the most relevant ones in the facilitator guide.

Delete this note

Scenario 1.1

5 mins

List a few pieces of information in this explanation.

Your answer here

Facilitator note

This is the first interactive slide. Participants in your workshop will literally write their answers to the question here.

Delete this note

Google

Peop

Scenario 1.1

5 mins

What information on your list is meaningful, appropriate, and timely for a driver who has just entered the car for the first time?

Your answer here

Scenario 1.1

5 mins

What information on your list is useful every time the driver gets into the car?

Your answer here

Google

People + AI Guidebook

Scenario 1.1

5 mins

Do you think this is a good explanation? Why or why not?

Your answer here

Scenario 1.2

5 mins

Are these good explanations of the identification failures? If not, why not?

Your answer here

Google

People + AI Guidebook

Scenario 1.2

5 mins

Discuss some approaches for more helpful explanations.

Your answer here

Google

People + AI Guidebook

Scenario 1.2

5 mins

In situations where the Model-U does not successfully identify a registered driver, what mechanism should be provided so they can start their car and how should it be communicated to them?

Your answer here

Scenario 1

5 mins

Apply back to your product:

How does your evaluation of good explanations here make you think about what aspects of your product need to be explained the first time vs every time?

Your answer here

Google

People + AI Guidebook

Scenario 2

5 mins

Read the scenario

Link to scenario

Scenario 2

5 mins

Is the explanation a good fit for the driver's situation? Why or why not?

Your answer here

Scenario 2

5 mins

What information does Olivia need at the moment she considers the self-driving option, and why?

Your answer here

Scenario 2

5 mins

Which information could be provided to Olivia at a different time instead, and how could it be presented to her?

Your answer here

Scenario 2

5 mins

How should Intelligent Engines respond to Olivia, and what actions should it take based on her complaint?

Your answer here

Google

People + AI Guidebook

Scenario 2

5 mins

Apply back to your product:

Can you identify some explanations (or other information) that your product is providing at an inappropriate moment in time?

How could this be improved?

Your answer here

Scenario 3

5 mins

Read the scenario

Link to scenario

Scenario 3

5 mins

If Richard answers “Yes”, how should the music review feature work since he is driving?

Your answer here

Google

People + AI Guidebook

Scenario 3

5 mins

How should the music review feature work when the user is not driving?

Your answer here

Google

People + AI Guidebook

Scenario 3

5 mins

Does Richard understand enough about how the entertainment recommender system chooses songs? If not, what does he need to learn in order to meaningfully influence it?

Your answer here

Scenario 3

5 mins

Apply back to your product:

How does Richard's experience with training the entertainment system apply to how (+ if and when) you might explain to users what your product is doing and whether the user has any control over it?

Your answer here

Scenario 4

5 mins

Read the scenario

Link to scenario

Scenario 4.1

5 mins

How should a minor accident be explained to both drivers and their insurance companies?

Your answer here

Google

People + AI Guidebook

Scenario 4.1

5 mins

At this point in time, what responsibility does the company have to investigate and act on the concerns raised on social media?

Your answer here

Scenario 4.1

5 mins

What information should Intelligent Engines share publicly about how they are handling the situation?

Your answer here

Google

People + AI Guidebook

Scenario 4.2

5 mins

Many AI systems can be vulnerable to adversarial attacks, even if their foreseeable likelihood is small enough to pass regulatory scrutiny.

How can a remote chance of being subjected to a malicious and adversarial attack by third parties be explained to the owners or prospective buyers of the Model-U?

Your answer here

Scenario 4.2

5 mins

What should the team explore in their audit? What information from the audit should they explain in an external facing audit report?

Your answer here

Scenario 4

5 mins

Apply back to your product:

How might someone use your system in an unintended or adversarial way? How might you prevent this?

What obligation do you have to your users when this happens?

Your answer here

Scenario 5

5 mins

Read the scenario

Link to scenario

Scenario 5.1

5 mins

Choose one character per participant, and discuss:

Which data or information you may need to support your positions or concerns?

Which data or information is likely to be most meaningful in making a decision about which roads the Model-U can drive on?

Your answer here

Scenario 5.2

5 mins

From the point of view of Intelligent Engines, which information requested by community members in the previous question can Intelligent Engines provide?

Are there any limitations that make it difficult or undesirable for a corporation to provide this information?

Your answer here

Google

People + AI Guidebook

Scenario 5.3

5 mins

What would the community (and especially your character) need to know on an ongoing basis to evaluate the effectiveness of the policy about which roads the Model-U can drive on?

Your answer here

Scenario 5

5 mins

Apply back to your product:

Can you think of some equivalent unforeseen problem with your product analogous to traffic in rural towns?

If your product is extremely successful, what problems might it cause?

Your answer here

Google

People + AI Guidebook

Break

15 minutes / Back at 10:45

Critical moments

Goal

Uncover potential user problems during key touchpoints in your user's journey within the scope of your challenge statement.

Instructions

Discuss the key moments in your product where ML-based problems may occur. (These key moments were chosen by your facilitator.)
Document user problems on your sticky notes as you discuss.

Google

Peop

Facilitator note

You will need to choose up to 5 critical product moments to focus on. More info on the next slide and in the facilitator guide companion doc.

Delete this note

Setup

Critical moments

These are critical places or moments in your product where trust can be gained or lost. They are typically places where the user interacts with your ML algorithm — recommendations delivered, explanations, confidence scores, errors, etc.

5 critical moments have been chosen by your facilitator to focus in on for your workshop. They should fall within the scope of your challenge statement.

Explaining the thing during onboarding	Quick description
Critical moment	Quick description
Critical moment	Quick description
Critical moment	Quick description
Critical moment	Quick descripti

Facilitator note

Fill out this slide in advance of the workshop. Present these moments to your team using this and the following slide.

Delete this note

Critical moments

Name of critical moment 1

Critical moment 1 description

Insert a screenshot of the critical moment here.

Facilitator note

Fill out this slide with information about your critical product moment.

Duplicate for each one.

Delete this note

Google

Peop

Opportunity statements

Goal

Turn your critical moments into opportunities for solving user problems.

Instructions

1. For each critical moment, write a problem statement in the voice of the user.
2. Identify relevant patterns in the People + AI Guidebook to solve user prob

Facilitator note

From this point, divide the team into groups for each of your critical moments. One person or sub-group per moment.

Delete this note

Setup

Step 1: Write in the voice of the user

For each critical moment, document a problem framed in the voice of the user. This should encapsulate the user's thoughts and feelings in this moment relative to your AI system.

Write the statement on the slide below.

Example critical moments and user problems

Onboarding new users into our product

I am not sure what this product can help me with and I'm afraid to keep pressing buttons to find out.

Opting in to personalization flow

I don't want to feel like I'm being manipulated, but I don't want to be distracted by a bunch of irrelevant explanations either.

Correcting errors in our automated system

I'm worried about spending too much time correcting the AI that I might has well have done this manually.

Setup

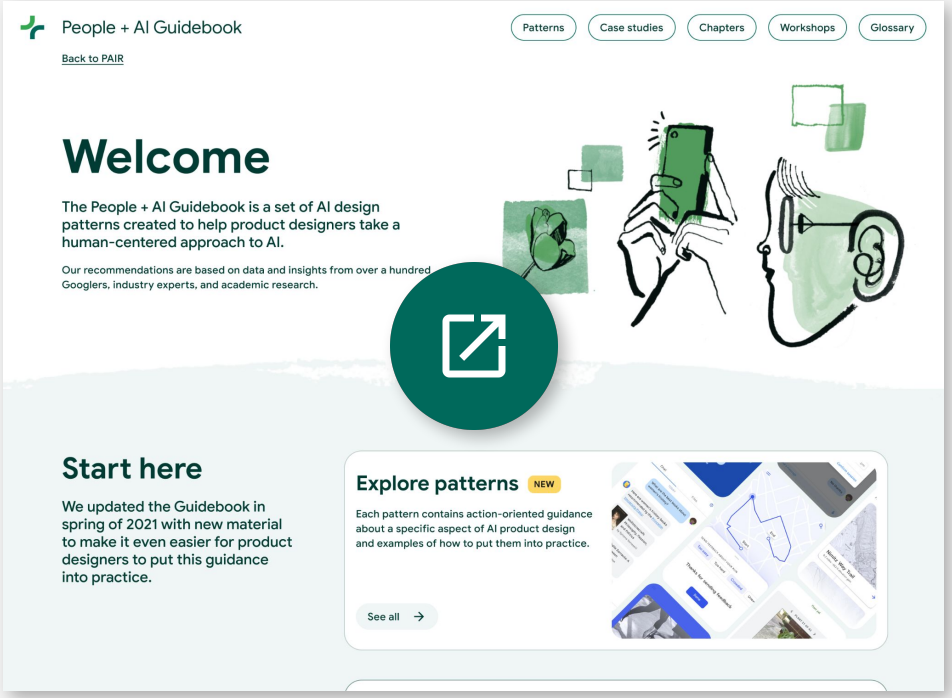
Opportunity statements

Step 2: Find relevant patterns

Divide into groups for as many critical moments as you have.

Look at the patterns in the People + AI Guidebook and identify any that are relevant to your question + critical moment.

Paste the relevant patterns in your opportunity statement on the slide below.



Google

People + AI Guidebook

Opportunity statements

30 mins

Paste your critical moment 1 label here

Write a statement that frames your problem relative to the critical moment. Write the statement in the voice of the user.

E.g. “I assume this app is all I need to venture into the woods and know what is safe to eat.”

Relevant Guidebook patterns	Ideas for applying the pattern
E.g. Set the right expectations	E.g. Explain during onboarding that our system wasn’t trained to find toxic plants outside the US.
Pattern	Solution
Pattern	Solution
Pattern	Solution
Pattern	Solution

Facilitator note

There’s one of these slides for each critical moment. Add or delete as needed.

Delete this note

Opportunity statements

30 mins

Paste your critical moment 2 label here

Write a statement that frames your problem relative to the critical moment. Write the statement in the voice of the user.

E.g. “I assume this app is all I need to venture into the woods and know what is safe to eat.”

Relevant Guidebook patterns	Ideas for applying the pattern
E.g. Set the right expectations	E.g. Explain during onboarding that our system wasn’t trained to find toxic plants outside the US.
Pattern	Solution
Pattern	Solution
Pattern	Solution
Pattern	Solution

Opportunity statements

30 mins

Paste your critical moment 3 label here

Write a statement that frames your problem relative to the critical moment. Write the statement in the voice of the user.

E.g. “I assume this app is all I need to venture into the woods and know what is safe to eat.”

Relevant Guidebook patterns	Ideas for applying the pattern
E.g. Set the right expectations	E.g. Explain during onboarding that our system wasn’t trained to find toxic plants outside the US.
Pattern	Solution
Pattern	Solution
Pattern	Solution
Pattern	Solution

Opportunity statements

30 mins

Paste your critical moment 4 label here

Write a statement that frames your problem relative to the critical moment. Write the statement in the voice of the user.

E.g. “I assume this app is all I need to venture into the woods and know what is safe to eat.”

Relevant Guidebook patterns	Ideas for applying the pattern
E.g. Set the right expectations	E.g. Explain during onboarding that our system wasn’t trained to find toxic plants outside the US.
Pattern	Solution
Pattern	Solution
Pattern	Solution
Pattern	Solution

Opportunity statements

30 mins

Paste your critical moment 5 label here

Write a statement that frames your problem relative to the critical moment. Write the statement in the voice of the user.

E.g. “I assume this app is all I need to venture into the woods and know what is safe to eat.”

Relevant Guidebook patterns	Ideas for applying the pattern
E.g. Set the right expectations	E.g. Explain during onboarding that our system wasn’t trained to find toxic plants outside the US.
Pattern	Solution
Pattern	Solution
Pattern	Solution
Pattern	Solution

Break

15 minutes / Back at 12:15

Share and wrap up

Goal

Review and choose which opportunity statements you'll focus on in day 2.

Instructions

Share the opportunity statements you created with the large group.
Discuss and choose which one(s) to focus on in day 2.

Facilitator note

Pick one opportunity statement to focus on or one per small group. Vote or discuss to determine which one(s).

This means leaving some of your opportunity statements and critical moments behind. Prioritize based on your product and the makeup of your workshop group.

Delete this note

End of day 1

Back at 9:00am tomorrow

Welcome to day 2!

Google

People + AI Guidebook

Challenge statement goes here

Google

Peop

Facilitator note

This is a reminder of the challenge statement at the beginning of day 2. Just copy/paste from slide 4.

Delete this note

Opportunity statements

Paste opportunity statement(s) here

Facilitator note

Paste the slides for your chosen opportunity statements here as a refresher from Day 1.

Delete this note

Agenda overview

Day 1

Understand

Case studies
Critical moments

Define

Opportunity statements

Day 2

Sketch

Errors
Explainability
AI onboarding
Trust calibration
Controls
Feedback
Data headlines

Decide

Review questions

Prototype & validate
On your own

Facilitator note

Feel free to modify this to suit the needs of your workshop.

Delete this note

Day 1 agenda

Morning

9:00	Arrival & review opportunity statements
9:15	Errors audit
10:00	**Break**
10:10	Explainability audit
10:50	AI Onboarding
11:30	Break/Stretch
11:35	Trust calibration
12:00	**Lunch break**

Afternoon

1:00	Controls audit
1:30	Feedback audit
2:00	**Break**
2:10	Dataset checklist
2:30	Review questions
3:00	Share with group and next steps
3:30	**End of workshop**

Google

Peop

Facilitator note

Feel free to modify this to suit the needs of your workshop.

Delete this note

Errors audit

Goal

Develop plans for preventing and mitigating user harm caused by errors in your system.

Instructions

Follow the worksheet to document errors and user harms.
Write mitigation and prevention plans for each error.

Google

Peop

Facilitator note

Note that this and all remaining activities assume the whole group is focusing on one opportunity statement. Duplicate the activities as needed for any sub-groups.

Delete this note

What could go wrong?

Follow the prompts on the slides below to identify harms that could be caused by use of your system: whether by explicit error, wrong predictions, or unintended or malicious usage.

How do I support users when something goes wrong?

Google

Setup

Be accountable for your errors

Understand the types of errors users might encounter and have a plan for resolving.

Give control back to users when automation fails

Give your users a way to move forward even when the system fails or offers poor quality output.

Let users supervise automation

Understand the types of errors users might encounter and have a plan for resolving.

Describe errors at the right level

Setup

✖ Too vague

System's answer doesn't fit the context.

✔ Just right

Misunderstands speech due to accent and environmental noise.

✖ Too specific

Doesn't know when garbage day is for one block in my neighborhood

10 mins

Errors audit

What error(s) can occur during this critical moment?	What user groups are impacted by the error?	What is the harm caused?
Example: System misunderstands speech due to accent and environmental noise.	All/Novice/Expert/Underserved/Etc	Time/Hassle/Financial/Safety/Etc
Error	All/Novice/Expert/Underserved/Etc	Consequence(s)
Error	All/Novice/Expert/Underserved/Etc	Consequence(s)
Error	All/Novice/Expert/Underserved/Etc	Consequen

Facilitator note

Consider giving each participant one row to work in and then share back to the rest of the group.

Delete this note

Errors audit

10 mins

What error(s) can occur during this moment? Copy/paste these from previous slide	How do we prevent this error from happening?	What do we do when (not if) it does occur?
Error	Prevention plan	Recovery plan
Error	Prevention plan	Recovery plan
Error	Prevention plan	Recovery plan
Error	Prevention plan	Recovery plan

Errors audit

Put it all together

On the following slide(s), write a holistic plan for addressing each of the errors you listed out above. Consolidate and synthesize as needed.

Duplicate the slide for as many errors as you want to document.

Google

Examples

Broken automation
Automation that fails costs users time, hassle, and unplanned working time. In high confidence scenarios that fail anyway, we'll provide users with system credit for lost time.

Bad recommendations
In business settings, adopting a bad recommendation causes lost reputation and trust by colleagues and/or clients. We'll make sure all recommendations are reversible to correct any issues they might cause.

Misidentification
Classifying users can fail by singling out some individuals or letting others fall through the cracks. We'll have accountable humans there to address case by case issues when the system fails.

People + AI Guidebook

Setup

Errors mitigation plan

10 mins

In {critical product moment}, {specific error} can impact {impacted user group}.

To prevent this error, we will {error prevention strategy}. When the error does occur, we will help our users by {error recovery strategy}.

Break

10 minutes / Back at 10:10

Explainability audit

Goal

Find gaps in your current explanations and write new ones that better support your users.

Instructions

For each opportunity statement, answer questions to analyze that moment. Write new explanations to help your users.

Explainability audit

Audit overview

In this activity, you will follow the prompts through a set of questions that evaluate your current system explanations and write new ones.

How do I explain my AI system to users?

Google

Setup

Overview

1. User mindset
Who are your users and what state of mind are they in when they see explanations in this critical moment?

2. Explanation evaluation
What are the pieces of info being provided in your current explanations and are they adequate?

3. User actions
How actionable are your explanations for the user? Are there steps they can take afterward?

4. Explanation strategy
Use the answers to the above questions to write new explanations for your product.

Explainability audit: User mindset

5 mins

What types of users see this critical moment?
Your answer here. (E.g. All / Novice / Expert / Low-usage / Underserved / Etc.)

User's state of mind in this moment?

😄/😟/😡/😬/😅/😒/🤔/😮

What is the user's goal in this moment?
Your answer here

Where is the user just before this moment?

Screenshot of previous moment

Does the user know what to do nex
Your answer here

Facilitator note

Consider dividing up and having each person fill out this activity for each opportunity statement you are focusing on.

Delete this note

Google

Peop

Explainability audit: Explanation evaluation

5 mins

What key info does any existing explanation provide?

Your answer here

What other information could be provided?

Your answer here

Should additional info be added here, or elsewhere in the system?

Your answer here

Explainability audit: User actions

5 mins

What actions are available to the user? Are they clear?

Your answer here

If a user believes your system's prediction is wrong, what can they do?

Your answer here

If a user is unhappy with this decision, what are they likely do first?

Your answer here

Explainability audit

Write new explanations

After scrutinizing the existing explanation, on the following slide write out new information you want to provide. Duplicate the slide as needed.

How do I explain my AI system to users?

Setup

Pattern considerations

Optimize for understanding, not completeness
"This is most likely a **sword plant**, because of **its dark green color and pointy shape**."

Determine how to show model confidence
"**87%** confidence this is **Sword plant**"

"This is most likely a **wand plant**, **whisk plant**, or **bottle plant**."

Go beyond in-the-moment explanations
"Our system is trained by humans looking at thousands of pictures of common houseplants and labeling them by their type, color, and shape. The system uses this data to make statistical inferences (good guesses) to identify the plants you see."

Explanation strategy

15 mins

In the moment explanations

Your answer here

Explanations elsewhere in the product

Your answer here

Explanations outside the product

Your answer here

In the moment explanations

Your answer here

Explanations elsewhere in the product

Your answer here

Explanations outside the product

Your answer here

AI Onboarding

Goal

Consider your system relative to a human performing the same tasks. Then work through how you'll present information on those terms back to your users.

Instructions

Work through the set of questions regarding the strengths, weaknesses, and "weirdness" of your AI. You'll frame the answers to your questions in terms of human mental models.

AI Onboarding

Introducing new AI features

Introducing users to fundamentally new features can be tricky.

These patterns from the Guidebook can help you explore this topic.

How do I onboard users to new AI features?

Google

Setup

Pattern considerations

Explain the benefit, not the technology
Help users understand your product's capabilities rather than what's under the hood.

Anchor on familiarity
As you onboard users to a new AI-driven product or feature, guide them with familiar touchpoints.

Make it safe to explore
Let users test drive the system with easily reversible actions.

Automate in phases
Progressively increase automation under user guidance.

People + AI Guidebook

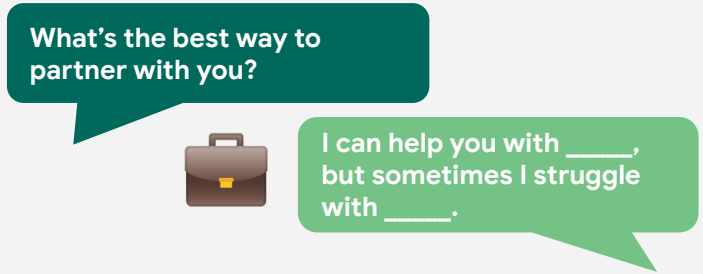
Onboarding humans in human terms

This exercise helps you frame the features in terms of human mental models.

Setup

What is onboarding?

Onboarding is the process of teaching users **how to partner with** an AI, rather than simply presenting AI information.



In this activity, you'll create your product onboarding in these 4 areas:

- Expected benefits
- Strengths & weaknesses
- AI "weirdness"
- Tips for users

AI Onboarding: Expected Benefits

5 mins

In what way is the AI in your product intended to benefit the user?

E.g. Accuracy, efficiency, consistency, up-level an ability, meaningfulness of work

It can help you because....(your answer here)

Which human weaknesses (if any) is it intended to address?

Humans may be limited in their ability to...(your answer here)

AI Onboarding: Expected Benefits

5 mins

When users are relying on another *human* for this task, what qualities do they care about when evaluating the human on a similar task?

E.g. If looking for book recommendations, what would a user expect another person to consider when recommending books?

Your answer here

AI Onboarding: Strengths & Weaknesses

5 mins

Where does your model perform well relative to the human criteria you just described?

The model performs well at...(your answer here)

Where does it struggle?

The model might struggle a bit on...(your answer here)

AI Onboarding: Strengths & Weaknesses

5 mins

What special cases (e.g. edge cases) would someone expect a human to be able to handle for this task?

E.g. When people give each other driving directions, they know to avoid roads under construction, steep hills, and bizarre traffic patterns.

Your answer here

AI Onboarding: Strengths & Weaknesses

5 mins

Where does your model perform well in the special cases you described?

The model is really good at...(your answer here)

Where does it struggle?

The model might struggle a bit on...(your answer here)

5 mins

What cases are “out of scope” for the model that a user might expect to from a human?

The model was not trained on particular kinds of data so there are no expectations of performance in those cases.

E.g., the book recommender never “read” books from before 1900, so don’t expect classics in your recommendation set.

You might expect other humans to be able to handle....but the model...

AI Onboarding: AI “Weirdness”

5 mins

In what ways does the model behave “weirdly” or differently from human ways of thinking? (aka, how is the “AI” brain different from the human brain?)

E.g. *The story-writing model sometimes changes the main character’s name partway through the story. This is because the model learned from paragraph-long snippets, so it hasn’t “seen” plot lines with longer-term dependencies.*

You might think that the model...but it may appear to act strangely by... This is because...(your answer here)

Google

People + AI Guidebook

Synthesize critical info

From the content you generated above, summarize the most important points.

Think of these as onboarding screens you may insert at the beginning of the product journey or in specific critical moments.

Example of AI Onboarding in healthcare imaging

Google

Examples

Plant Pal

This is **Plant Pal**, and it can help you by **identifying the safety of 400+ plants in the USA**. It is unable to **identify plants outside the US or make assertions about their safety**.

It's really great at **understanding safety for adults, dogs, and cats**, but it may struggle on **safety for kids**. It may sometimes appear to act strangely because **it doesn't identify the plants the same way humans do**.

Run

This is **Run**, and it can help you by **suggesting running routes in your area**. It is unable to **make recommendations based on uncommon health conditions**.

It's really great at **choosing appropriate elevation gains**, but it may struggle on **knowing the safety level of sidewalks**. It may sometimes appear to act strangely because **its routes are based on mapping data so they might not look like obvious loops or out-and-backs**.

People + AI Guidebook

10 mins

Most critical onboarding information

Expected benefits

This is {product/feature}, and it can help you by {main benefit}.

Strengths & limitations

It's really great at {special cases the model can handle}, but it may struggle on {special cases the model doesn't handle yet}.

It is unable to {out-of-scope areas}.

AI "Weirdness"

It may sometimes appear to act strangely by {"weirdness"}. This is because {explanation}.

Tips for users

Given these strengths and limitations, a good way of partnering with the model would be to rely on it more for {model's strength areas} but rely on your own best judgment for {users' strength areas}.

Quick break & stretch

5 minutes / Back at 11:35

Trust calibration

Goal

Help your users calibrate their trust in your system by not under- or over-promising on what you can deliver.

Instructions

Plot issues that impact user trust in your system.
Develop a plan for addressing too much or too little trust in your system.

Trust calibration

Weighing trust factors

List factors that may cause users to have too much, not enough, or just the right amount of trust in the system.

Consider what previous knowledge the user is bringing to the table.

How do I help users build and calibrate trust in my product?

Google

Pattern considerations

Setup

Set the right expectations

Be transparent with your users about what your AI-powered product can and cannot do.

Be transparent about privacy and data settings

From initial onboarding through ongoing use, continue to communicate about settings and permissions.

Add context from human sources

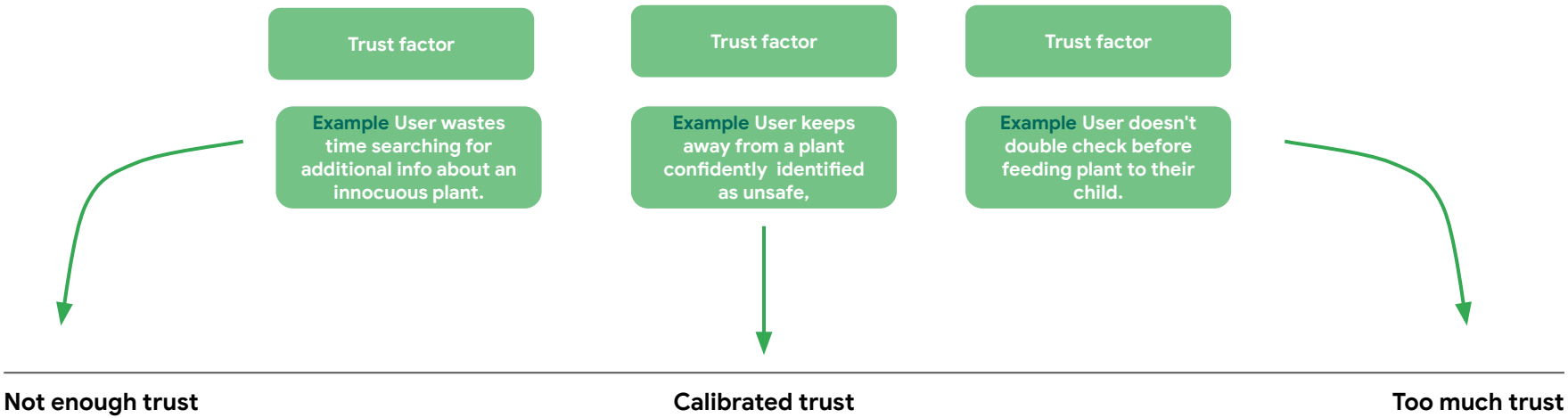
Help users appraise your recommendations with input from third-party sources.

Let users give feedback

Give users the opportunity for real-time teaching, feedback and error correction.

Trust calibration

5 mins



Google

People + AI Guidebook

Plan for helping users calibrate

Copy the most extreme examples of too much or too little trust from the slide above to the following slide.

Follow the prompts on the slide to write a strategy for properly calibrating user trust in those moments.

Duplicate the slide as needed.

Example

Under-trust factor
User wastes time searching for additional info about an innocuous plant identified by Plant Pal.

Calibration plan
Offer a “why we know this” option. And/or social proof that lets communities verify safe plants.

Over-trust factor
A user doesn't double check Plant Pal’s recommendation before feeding a wild plant to their child.

Calibration plan
For any low confidence result or if the “n-best” set of results includes potentially dangerous results, recommend the user get a 2nd opinion.

Setup

Trust calibration

10 mins

Trust factors (Copy/paste labels from previous slide)	**Did this situation result in too much or too little trust?**	**How will you help people properly calibrate their trust in this moment?**
Example User wastes time searching for additional info about an innocuous plant.	Too little	Offer a “why we know this” option. And/or social proof that lets communities verify safe plants.
	Too much / Too little	User-facing intervention for calibrating trust
	Too much / Too little	User-facing intervention for calibrating trust
	Too much / Too little	User-facing intervent calibrating trust

Facilitator note

Copy/paste your trust labels from the previous slide here.

Delete this note

Lunch break

60 minutes / Back at 1:00

Controls audit

Goal

Develop new controls for users to maintain agency over your system.

Instructions

Plot your product's critical moments according to risk.
Develop new controls for mitigating those risks.

Controls audit

Expectation of control

On the following slides, identify places where users need additional control based on the consequences of error if the system gets it wrong.

What's the right balance of user control and automation?

Google

Setup

Pattern considerations

Automate more when risk is low

Consider user trust and the stakes of the situation when determining how much to automate.

Let users supervise automation

Maintaining control over automation helps users build comfort and correct when things go wrong.

Give control back to the user when automation fails

Give your users a way to move forward even when the system fails or offers poor quality output.

Automate in phases

Progressively increase automation under user guidance.

Let users give feedback

Give users the opportunity for real-time teaching, feedback and error correction.

People + AI Guidebook

10 mins

Controls audit

Predictions or decisions made by your system in this critical moment	Does the user have any control over the system in this moment?	How extreme is the consequence if something goes wrong in this moment? Consider your responses in the Errors Audit to guide your answers here.
Example: Recommendation for today's running route	Yes, can accept/reject.	Medium — Route may be inappropriate for user's fitness level
	Yes/No	Low/Medium/High
	Yes/No	Low/Medium/High
	Yes/No	Low/Medium/High

Facilitator note

Copy/paste your prediction labels from the previous slide here.

Delete this note

10 mins

Controls audit

Predictions or decisions made by your system in this critical moment	Is the existing control adequate to address concern of user harm?	If not, what type of control might be better suited?
Example: Recommendation for today's running route	No. No ability to replace the recommendation or set preferences for new ones.	User can update preferences on the fly and get a new recommendation that's more appropriate.
	Yes/No. Why?	Description of adequate control
	Yes/No. Why?	Description of adequate control
	Yes/No. Why?	Description of adequ

Facilitator note

Copy/paste your prediction labels from the previous slide here.

Delete this note

Feedback audit

Goal

Develop a plan for implicit and explicit user feedback to guide your system.

Instructions

List out all of your product's existing and potential feedback mechanisms.
Create a plan for adding needed feedback where it doesn't yet exist.

Feedback audit

Gather feedback opportunities

On the following slides, list out all existing and potential feedback mechanisms in your product.

What's the right balance of user control and automation?

Setup

Considerations

Consult with engineers, designers, product counsel, and other stakeholders as needed to generate a truly comprehensive audit.

List out events and corresponding feedback opportunities that could provide data to improve the AI in your product. Cast a wide net and list as many as possible: App Store reviews, Twitter, email, call centers, push notifications, etc.

Duplicate the slides and add more rows as needed.

10 mins

Feedback audit

User-driven interaction with the system in this critical moment	Is there any user feedback gathered as part of this moment?	If so, what is it? If not, what feedback could potentially be gathered?
Example: User completes a recommended recipe	Yes, Explicit	User prompted to rate relevance of completed recipe
	Yes/No, Implicit/Explicit	Existing or potential feedback
	Yes/No, Implicit/Explicit	Existing or potential feedback
	Yes/No, Implicit/Explicit	Existing or potential f

Facilitator note

Copy/paste your interaction labels from the previous slide here.

Delete this note

10 mins

Feedback audit

User-driven interaction with the system in this critical moment	What user groups are affected or benefit from this feedback?	How will your model or system change as a result of it?
Example: User completes a recommended recipe	Only those who mark recipes as completed, closer to power users	Recommendations become better and more personalized
	All/Some, which ones?	System change
	All/Some, which ones?	System change
	All/Some, which ones?	System change

Facilitator note

Copy/paste your interaction labels from the previous slide here.

Delete this note

Prioritized feedback opportunities

10 mins

Copy and paste all of the opportunities to add new feedback mechanisms that represent the greatest impact on improving your system.

Duplicate this as needed.

What feedback can be gathered?	How will your model or system change as a result of it?
Potential feedback	System change
Potential feedback	System change
Potential feedback	System change
Potential feedback	System change

Break

10 minutes / Back at 2:10

Data headlines

Goal

Develop a set of opportunities and concerns for your data collection, labeling, and maintenance practices.

Instructions

Write data headlines to spot problematic issues before they happen.
Use these headlines to identify any data concerns.
Follow up with your engineering partners after the workshop.

Data headlines

Avoid data disasters

On the following slides, write potential data headlines based on your current or potential data collection, labeling, and maintenance practices.

How do I responsibly build my dataset?

Guiding questions

Setup

How do you get access to the data? Do you have permission?

What anonymization and/or aggregation techniques does your product use?

What is the downstream, real-world effect of this model's performance?

What data is missing that would adversely impact certain user groups?

Who are the humans involved collecting and/or labeling your data?

How are you compensating them for this critical work?

What risks are present for using data not originally intended for your use case?

Does your data reflect the real world? e.g. for image based systems does it include off center/blurry images?

Setup

Example Data Disaster Headlines

Data privacy	Users of **Run app** horrified to learn it uses **gender data** to recommend shorter routes for women.
Data exclusion	Uproar over **Run app**'s lack of **rural street data** that excludes **users outside of major cities**.
Data ethics	Calls to boycott **Run app over pressuring route labelers to test routes at night despite safety concerns.**
Borrowed Data	**Run app** cancelled over faulty **trail data** used from **Parks Department without including wildlife encounter risks.**
Data fragility	**Run app** down as team struggles to fix **route recommendations impacted by recent flooding in New Orleans.**

Data Disaster Headlines

10 mins

Data privacy	*Customers of **{product}** horrified to learn it uses **{sensitive data}**.*
Data exclusion	*Uproar over **{product}**'s lack of **{data type}** that excludes **{user group}**.*
Data ethics	*Calls to boycott **{product}** over unfair treatment of **{data labelers}**.*
Borrowed Data	***{Product}** cancelled over faulty **{data}** used from **{inappropriate source}**.*
Data fragility	***{Product}** down as team struggles to fix **{key data input sources}**.*

Google

People + AI Guidebook

Setup

Example Flip the script: Data Disaster Headlines

Data privacy breach	**Run app** champions essential data **by not using demographic data for route recommendations**.
Data exclusion	Praise for **Run app**'s inclusion of **all mapped street data** that benefits **users in rural areas**.
Data ethics	**Run app** prioritizes **route labeler safety** by **making nighttime route verification optional.**
Borrowed Data	**Run app** responsibly used **trail data** from **Parks Departments** by **including notification for wildlife sightings**.
Data fragility	**Run app** outperforms competitors thanks to **updating route data** in response to **local road closures.**

Flip the script: Data Diligence Headlines

10 mins

Data privacy breach	***{Product}** champions essential data by limiting use of **{sensitive data}**.*
Data exclusion	*Praise for **{product}**'s inclusion of **{data type}** that benefits **{user group}**.*
Data ethics	***{Product}** sets the bar for **{humans doing data collection/labeling}** by **{action taken to compensate fairly}**.*
Borrowed Data	***{Product}** responsibly used **{data}** from **{source}** by **{action taken}**.*
Data fragility	***{Product}** outperforms competitors thanks to including **{data}** that accounts for **{real world consideration}**.*

Google

People + AI Guidebook

Concerns & opportunities

10 mins

Data privacy breach	E.g. Our current collection of demographic info isn't necessary and puts us at risk.
Data exclusion	
Data ethics	
Borrowed Data	
Data fragility	

Google

People + AI Guidebook

Review questions

Goal

Take a step back to consider all the solutions you've developed through the lens of user trust.

Overview

Answer questions to scrutinize your solution ideas.
Document changes and next steps toward implementation.

Review questions

Go back and check in

Now that you've developed plans to address your opportunity statement(s), take some time to review your solutions and whether they hold up to the scrutiny of being truly trustworthy.

Answer the questions on the following slides to review.

Considerations

Consider framing your responses from one or more of the following viewpoints:

Tech skeptic
"Is this company being careful with my data?"

New internet user
"How does this work? What does 'signed in' mean?"

Business leader
"Why should we make any of these changes? Will they increase our company's revenue?"

Review questions

5 mins

Gather solutions

First, consider the solutions you developed across all of your activities.

Which ones seem most impactful that you'll want to developer further?

Use the solutions you choose here to guide your answers to the following questions.

Your answer here

Review questions

5 mins

Ability

Why is your company, specifically, making this thing? How is it obviously better than what already exists?

How does it show a deep consideration of what people need and want in their lives?

Your answer here

Google

People + AI Guidebook

Review questions

5 mins

Reliability

Can you meet the quality bar you set for yourself in every use case all of the time?

If not, what do you need to do to set expectations for users in those cases?

Your answer here

Review questions

5 mins

Benevolence

What's in it for your company for this product to exist?

Are you comfortable stating that purpose clearly and directly to your users? If not, why?

Your answer here

Review questions

Product reviews

What do you want to be true about how your product is trusted? Write a version of both a good and bad review you could imagine showing up in the tech press after these features are launched.

5 mins

Good review

Your answer here

Bad review

Your answer here

Share and wrap up

Goal

Reflect on the workshop and discuss what you want to do next.

Instructions

Before you close, take some time to reflect on what you learned in the workshop overall and document any additional next steps and actions.

Share and wrap up

30 mins

Top questions + next steps

Reflect on the review questions above and the workshop overall. What did you learn? What do you still have questions about. What should you do next.

Jot down your top questions and next steps for further changes and investigation.

Your answer here

End of workshop

People + AI Guidebook

Appendix: Optional materials

Solution board

15 mins

Top error mitigation and prevention plans:

Your answer here

Top controls solutions:

Your answer here

Top explainability solutions:

Your answer here

Top feedback solutions:

Your answer here

Top onboarding ideas:

Your answer here

Top data opportunities and concerns:

Your answer here

Your name — Sketching board

15 mins

Sketch your ideas on paper.

Snap a picture with your phone and upload to the slides here..