Chapter 4

Artificial Intelligence Driven Design.

By Joël van Bodegraven

Brain food — Vol 4

awwwards.books

Chapter 4.

DESIGN CHALLENGES IN MACHINE LEARNING PRODUCTS

This fourth chapter was written by Nadia Piet, a design researcher and strategist focused on AI/ML, data, futures & the human experience

About Nadia

Nadia Piet is a design strategist and researcher fascinated by how we shape technology, and technology shapes us. She's currently working with Bit, a research and prototyping studio on a mission to fast-forward the impact of emerging tech. Next to her role at Bit, she recently released the AI meets Design toolkit in collaboration with Accenture Interactive, facilitates workshops with DECODED, and previously worked as a freelancer for 7 years across a variety of roles, industries, and over 9 different countries. In her free time she likes to practice yoga, drink oat lattes with dino art, browse her Spotify discover weekly playlist, and scout for nudibranches in tropical waters.

Chapter 4.

- Introduction
- Themes
  + Trust & Transparency
  + User Autonomy & Control
  + Value Alignment
- Worksheet
- Outro
- Recommended Reading/ Appendix

Introduction

Every design material comes with unique opportunities and challenges. In the same way that designing an event poster is different from designing a mobile app, designing AI/ML-driven applications is different to designing mobile apps.

As we begin to see AI features popping up in our day-to-day products and services, its challenges begin to materialize. They range from UX problems, such as explainability and user feedback mechanisms, to greater ethical challenges, such as echo chambers and data bias. Designing the user experience of adaptive, intelligent, and semi-autonomous systems present a range of new challenges for us designers to take on.

When thinking or talking about AI, we often imagine utopian or dystopian futures. Rarely, we dare to acknowledge its impact as something we have a hand in shaping (or even: a design challenge). Technology may be neutral and deterministic, but its development is not. As designers, we can take the raw material of AI and turn it into user, business, and social value.

This chapter is by no means all-encompassing and only scratches the surface on the complexities of designing for AI. Instead, it aims to provide a starting point for building a shared understanding around some of the complexities of designing AI/ML interactions, spark discussion, and invite everyone to take part in (re-)imagining how to design positive user/human experiences in algorithmic systems.

This ebook, which shares my research on designing Machine Learning Products, will address 3 different themes **Trust & Transparency**, **User Autonomy & Control** and **Value Alignment**, highlighting 9 of the challenges that can arise within them, all supported with real life examples.

Theme 1:

Trust & Transparency

Not all AI features are invisible to the user, nor should we want them to be. When confronting our users with these new systems, it is our job to help them understand how they work, be transparent about their abilities, construct helpful mental models, and make them feel comfortable in their interactions. Transparency is key to building trust in the system, and respecting user trust in your organization.

Why is trust & transparency important?

- Access value
- Avoids confusion and disappointment
- Lowers drop-out rate
- Establishes trust in system
- Sustains trust in the brand/organization
- Easier interactions

1. Explainability

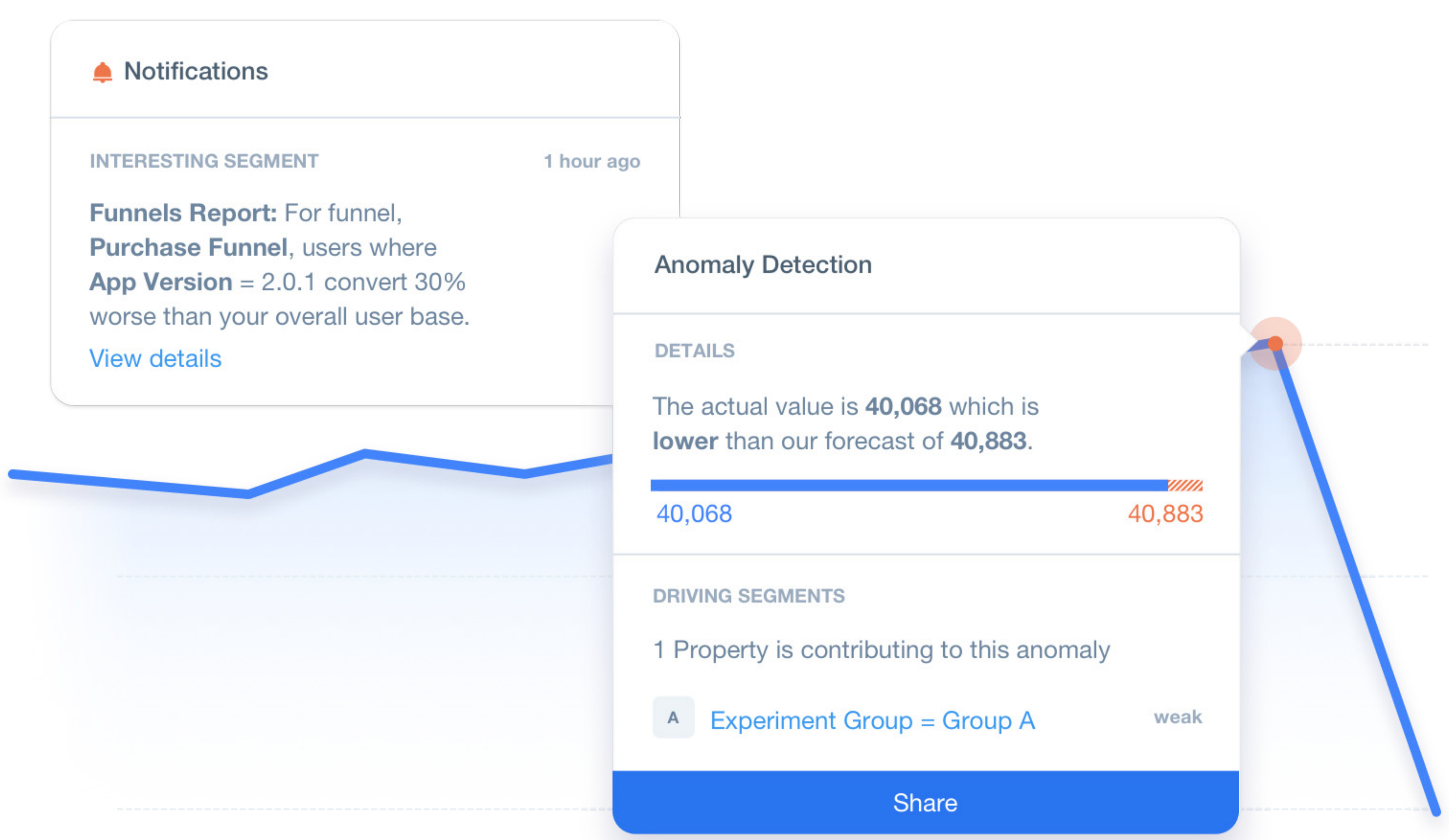
Making sense of the machine and communicating to the user why the system acts the way it does.

Design Strategy

Include a button to show people the data and features (where known) that went into the output and gradual detailing of the model's logic.

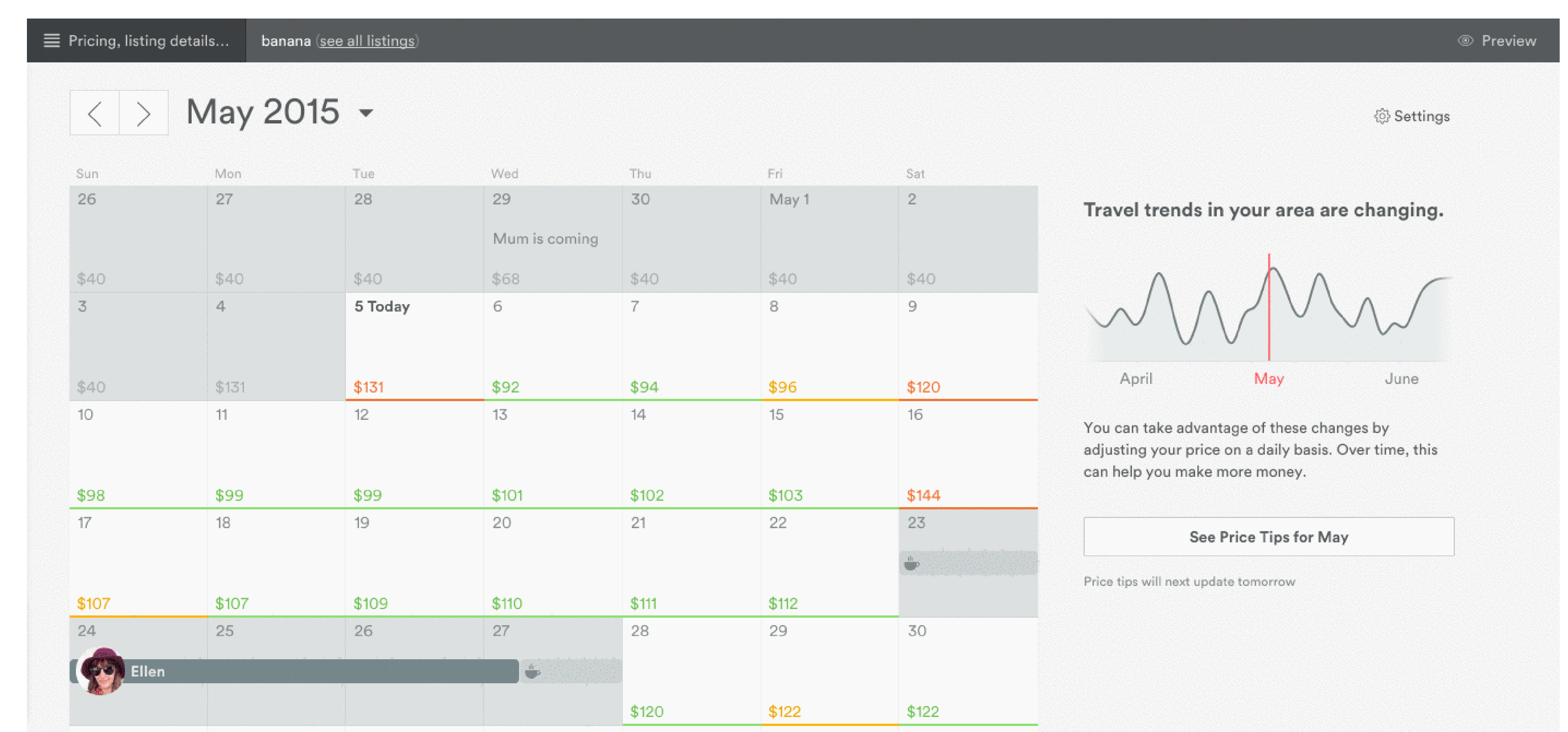
→ Mixpanel

Mixpanel, the business analytics service company, uses machine learning to uncover user insights. The anomaly detection feature helps pick up on unusual behavior. The image shows the anomaly, but also what data the prediction is based on and which segments drive the anomaly, so that the user can make an informed decision about next steps. It even offers a “share” function to consult with a colleague for a 3rd opinion.



→ Airbnb

When Airbnb introduced ‘smart pricing’ based on supply and demand, adoption wasn’t as high as expected. They learned users were happy to be informed by the algorithm, but wanted to make the final decision for themselves. Airbnb then built an interface where hosts can evaluate price changes and accept or reject each of the algorithm’s recommendations.



2. Managing Expectations

Assisting the user to build helpful mental models of what the system can and cannot do by being transparent about abilities and limitations.

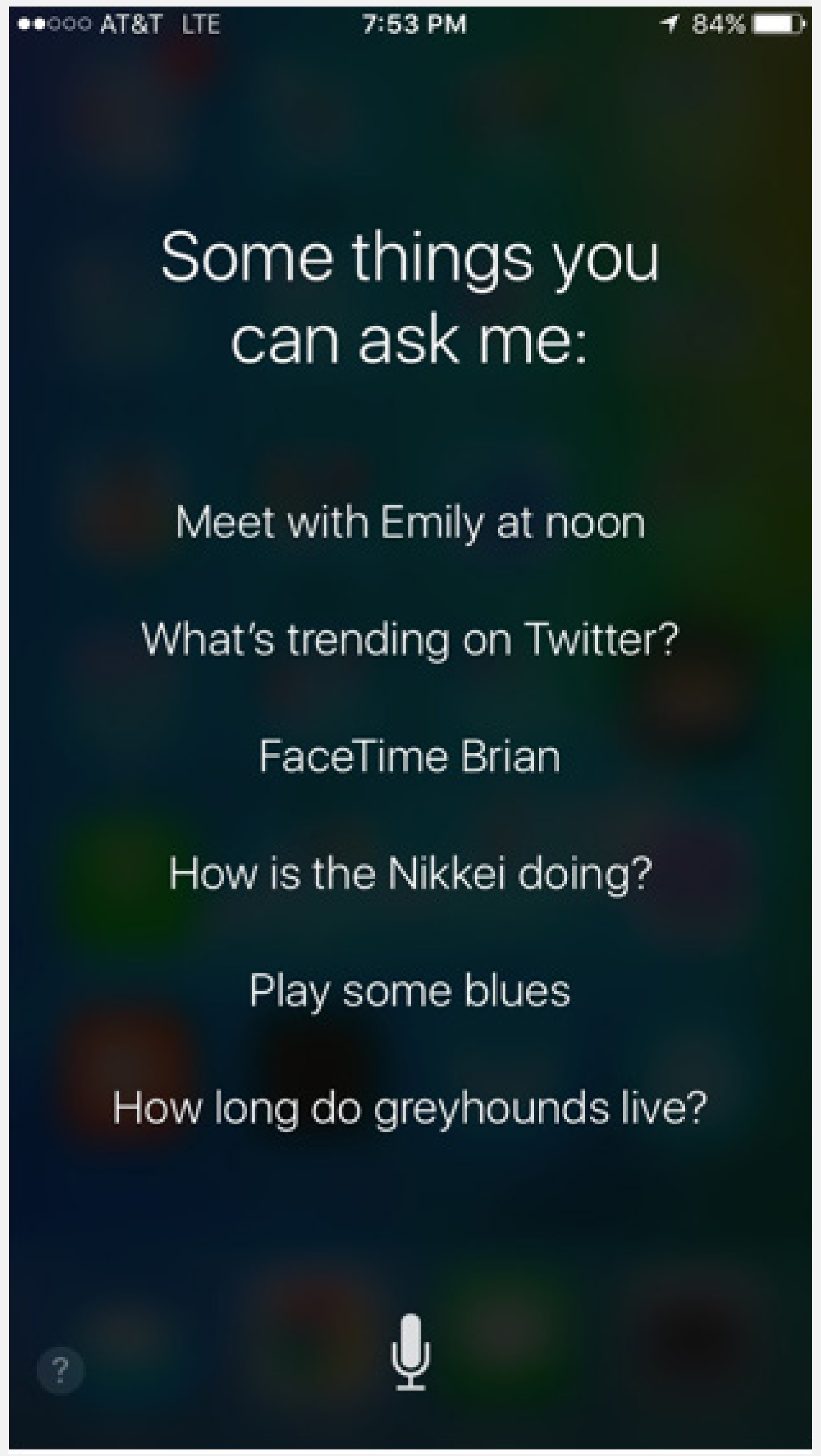
Design Strategy

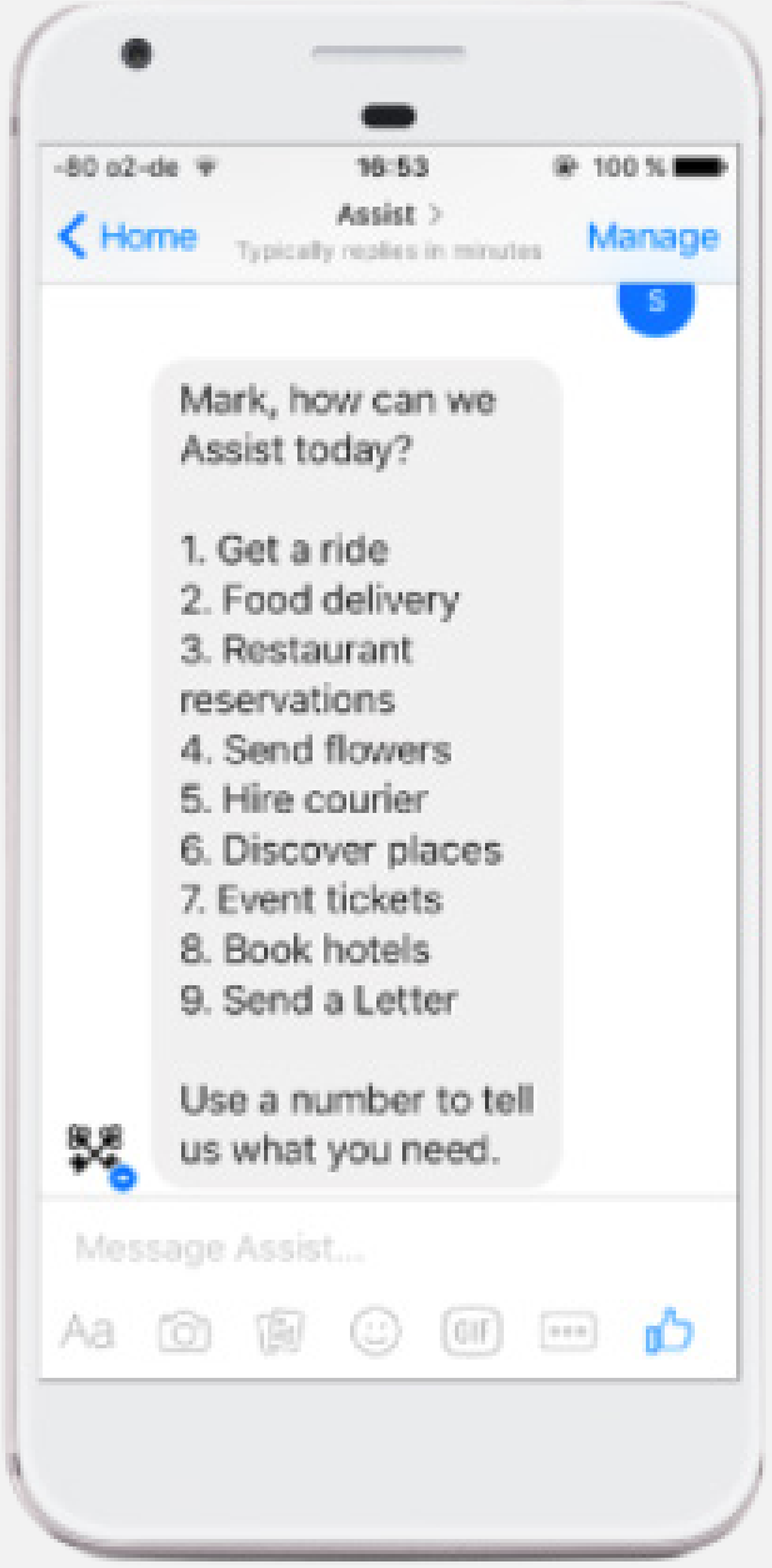
Proper onboarding during the first interaction with the application or feature where abilities and limitations are established.

→ Siri & Assistant

Each time the user calls upon Siri and it shows this screen, Siri has the opportunity to respond to queries and gradually introduce the user to its varied abilities. Onboarding and setting expectations becomes even more important in post-pixel interfaces because the user doesn't have physical affordances nudging them where to go and what to do.

The Assist chatbot doesn't try to cover up its shortcomings, but instead makes the most of its limited abilities by explicitly stating its abilities and how a user must communicate a query in order for it to be processed successfully.





3. Failure + Accountability

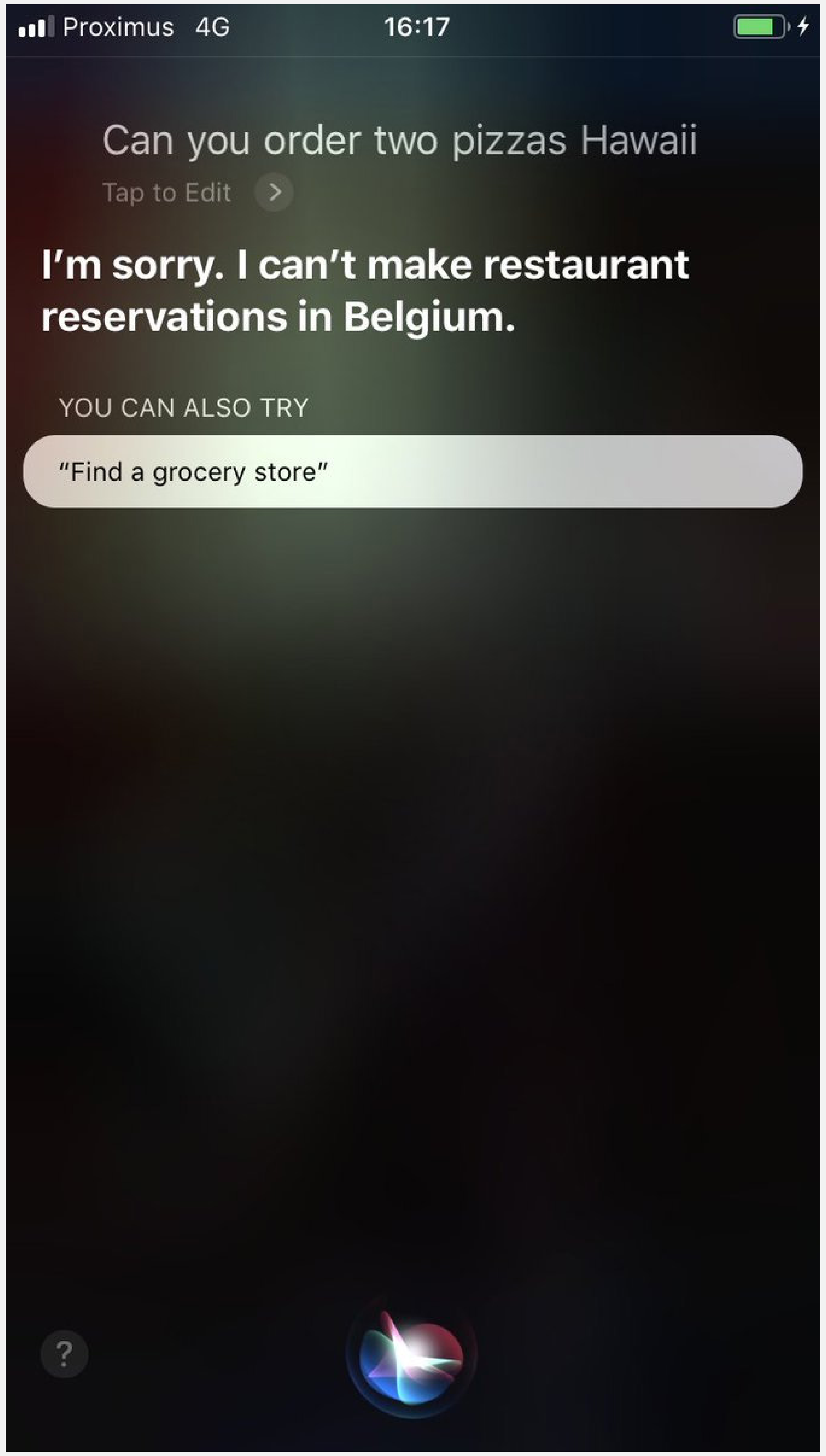
Assume failure and design graceful recoveries. Take accountability for mistakes and minimize the cost of errors for your user.

Design Strategy

Apologizing, minimizing automation bias, allowing the user to indicate a mistake, and taking accountability for mistakes.

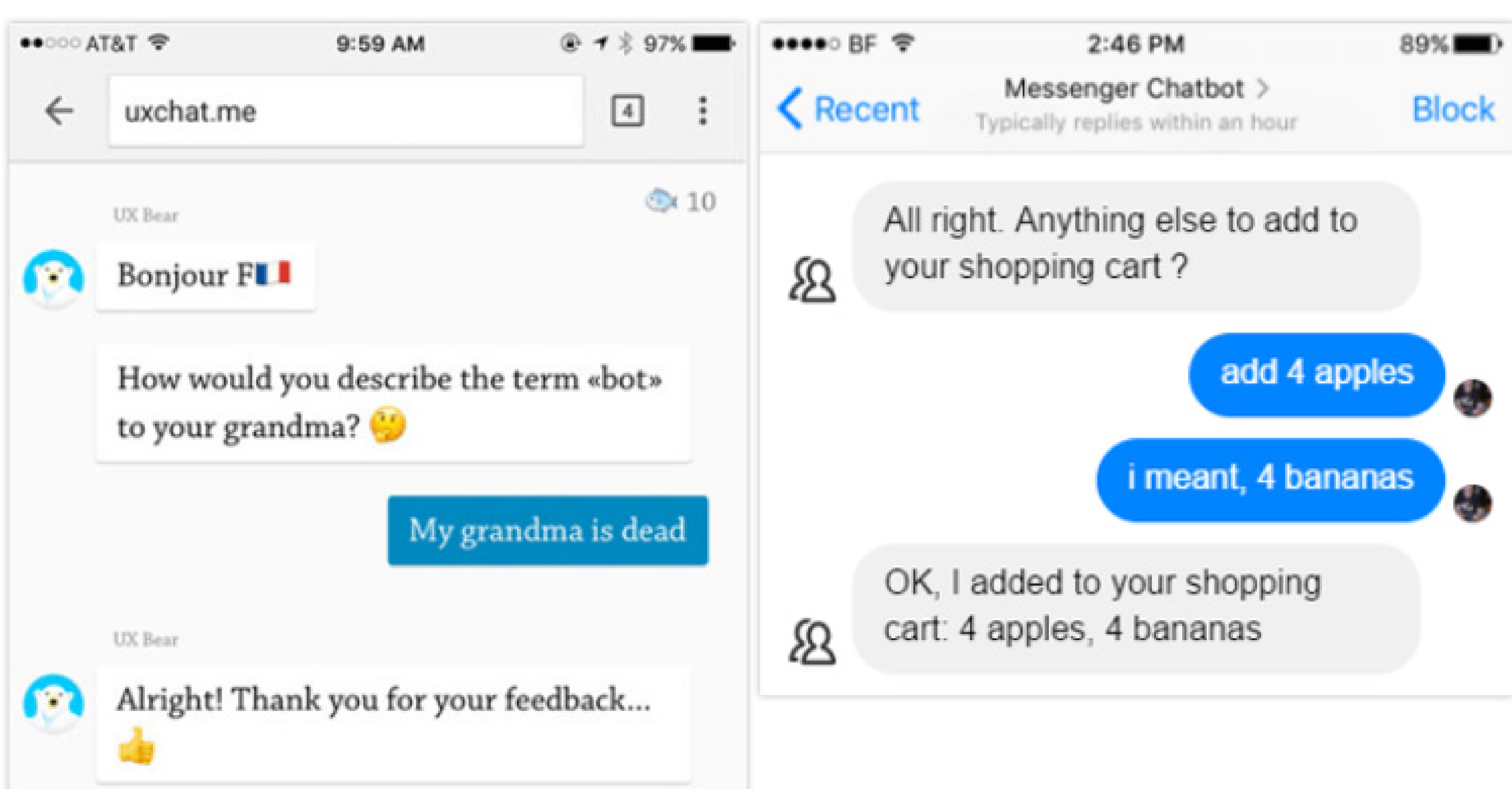
→ Google Home & Alexa

I make requests to my Google Home quite frequently that are returned with "I'm sorry, I can't do that yet." While disappointing, the apologies and promise of future improvement keep me from losing trust. In the example of Siri below, we can see it also recommends an alternative – a query it understands to be similar to the one initially called upon and one it can perform. Both of them recognized there is no way to user test against adaptive systems so they must be designed for failure.



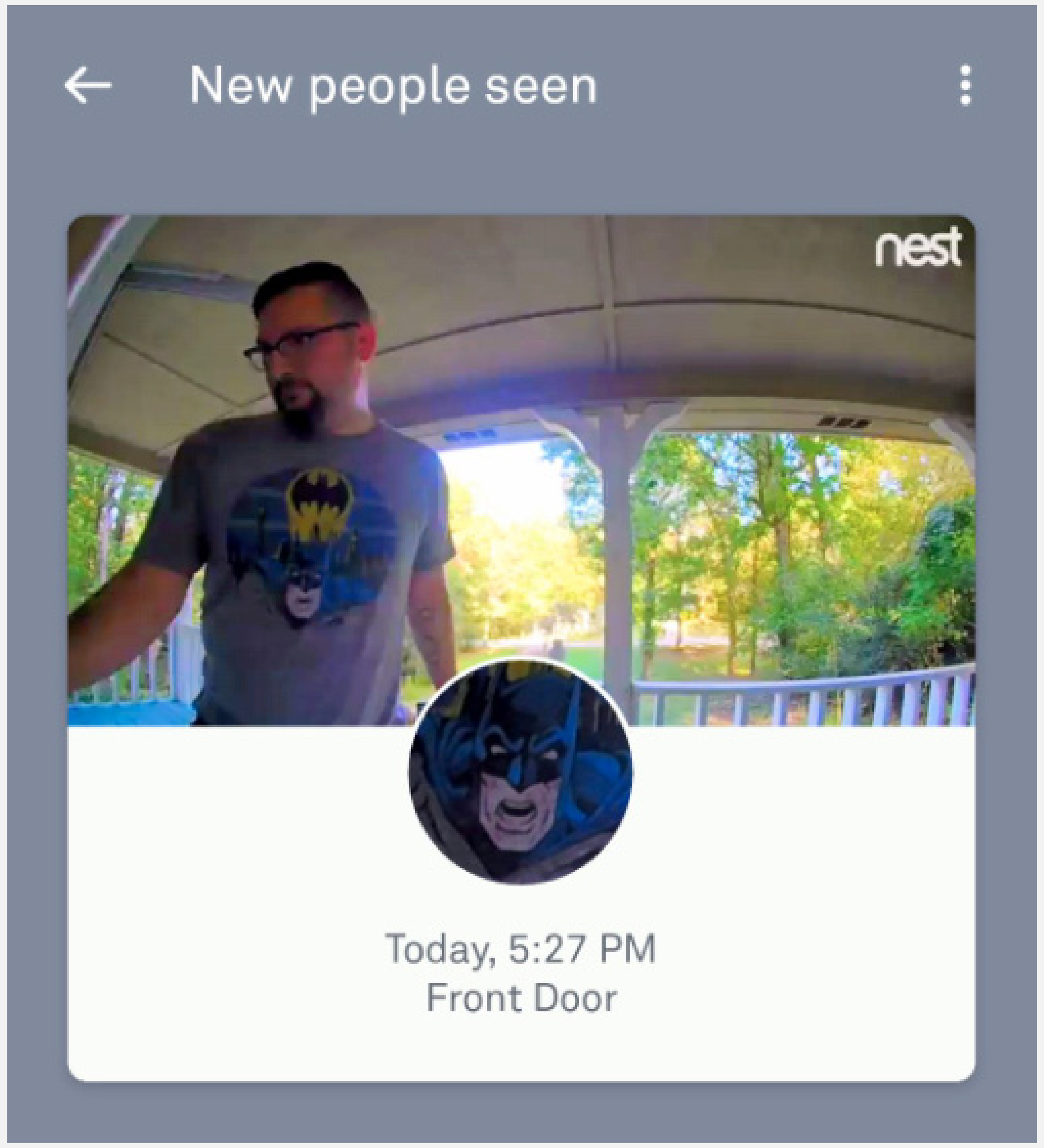
△ Fail! Chatbots

While we can not predict every possible scenario and ML's adaptive nature makes testing a bit more tricky, we can anticipate obvious failures and prevent them from leading to awkward user experiences like below. Test your systems in real-life, out-of-the-lab context to bring to the surface common and obvious mistakes.



Batman at the door

One day as B.J. May approached his Nest doorbell, it wouldn't let him in because the model thought he was Batman. Fortunately, the designers anticipated failure and designed 2 back-up ways for him to intervene and still get inside. Consequently, the failure didn't have many consequences other than a funny Twitter thread.



Quote

“For the foreseeable future, AI models will sometimes fall short. This gap presents an opportunity for UX designers”

Zuliani, 2019

Questions to consider

- **How do we build trust?**
  Consistency is key for building trust, but it can be hard to practice in adaptive systems.

- **What is the right level of trust?**
  Too little trust means the user doesn't get any value from the system. Too much trust might lead to automation bias and poses risks for both the user and the organization.

- **What is the right level of transparency?**
  Too little means the user doesn't trust your system.
  Too much means the user might get confused with an overload of information.

- **How to design an interface that minimizes the cost to the user when the AI makes a mistake?**
  Considering the impact of mistakes in your use case, and how to retaliate from a bad prediction to not harm trust.

- **Who is responsible and liable for the consequences of mistakes?**

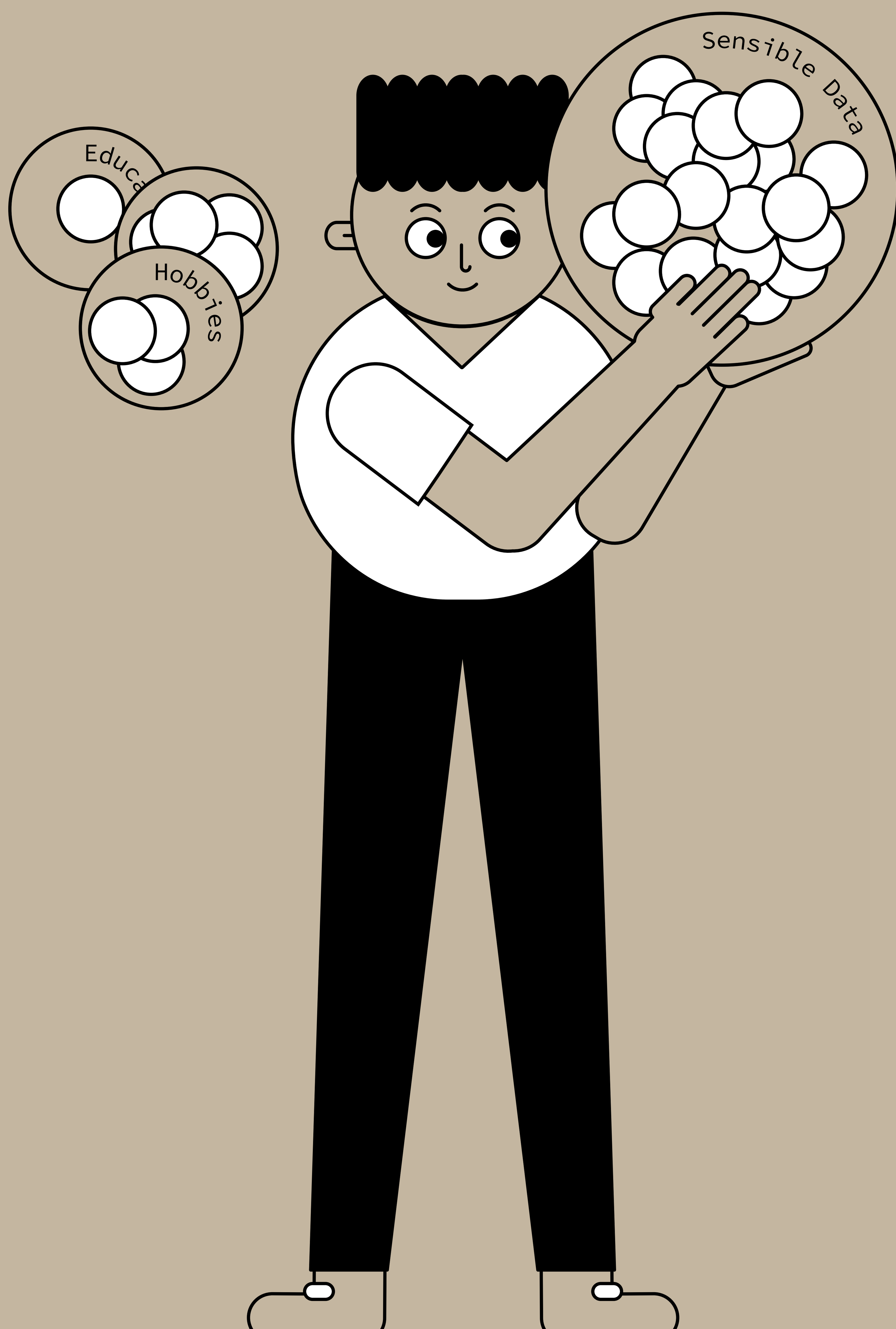
- **What are useful mental models to help users understand the AI?**

- **What are good ways to explain predictions, confidence, and the logic underlying them in the interface?**

- **How to explain those occasions when even the creators don't know how it works?**

Theme 2:

User Autonomy & Control



The user must feel like they're in charge of the system. People have justified concerns about giving up agency to (semi-) autonomous systems, and sharing the personal data required to make them work well. Respecting the human need for autonomy, users need a way to exercise consent and control over the system and their data based on their individual and contextual needs. One size rarely fits all and AI systems are no exception.

Why is user autonomy & control important?

Consent. Avoids feeling out of control. Avoid feeling being surveilled. Creates more user value through customization. Learn about user needs through feedback.

4. Machine Teaching + User Feedback

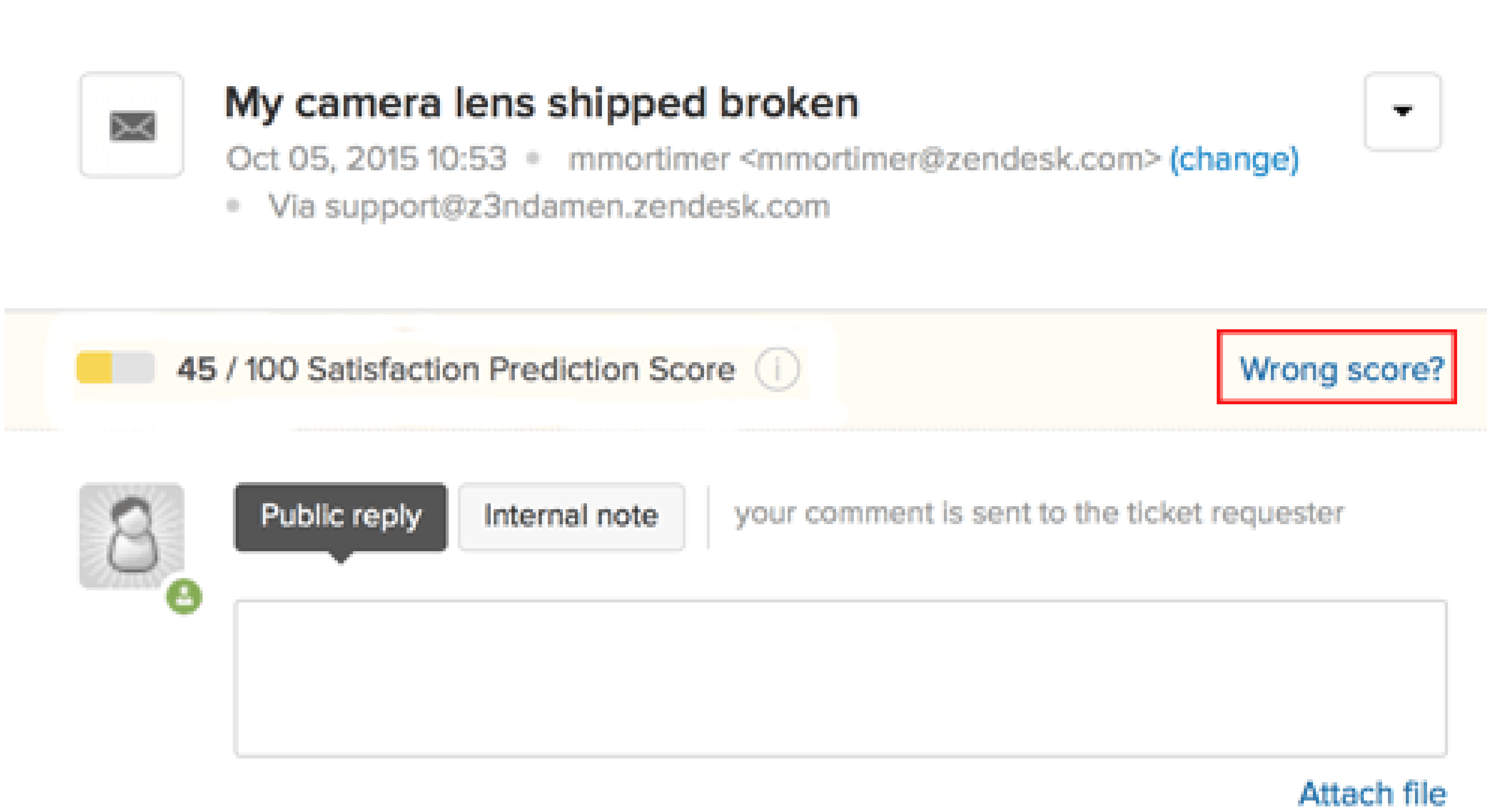
Allowing the user to teach the machine with implicit and explicit feedback loops and collecting direct data input.

Design Strategy

Building in implicit and explicit feedback loops. Considering the latter, give the user a way to quickly indicate if this is helpful “yes or no”, then gradually ask for more feedback like “why or why not”, and how the system could have acted better.

→ Zendesk:

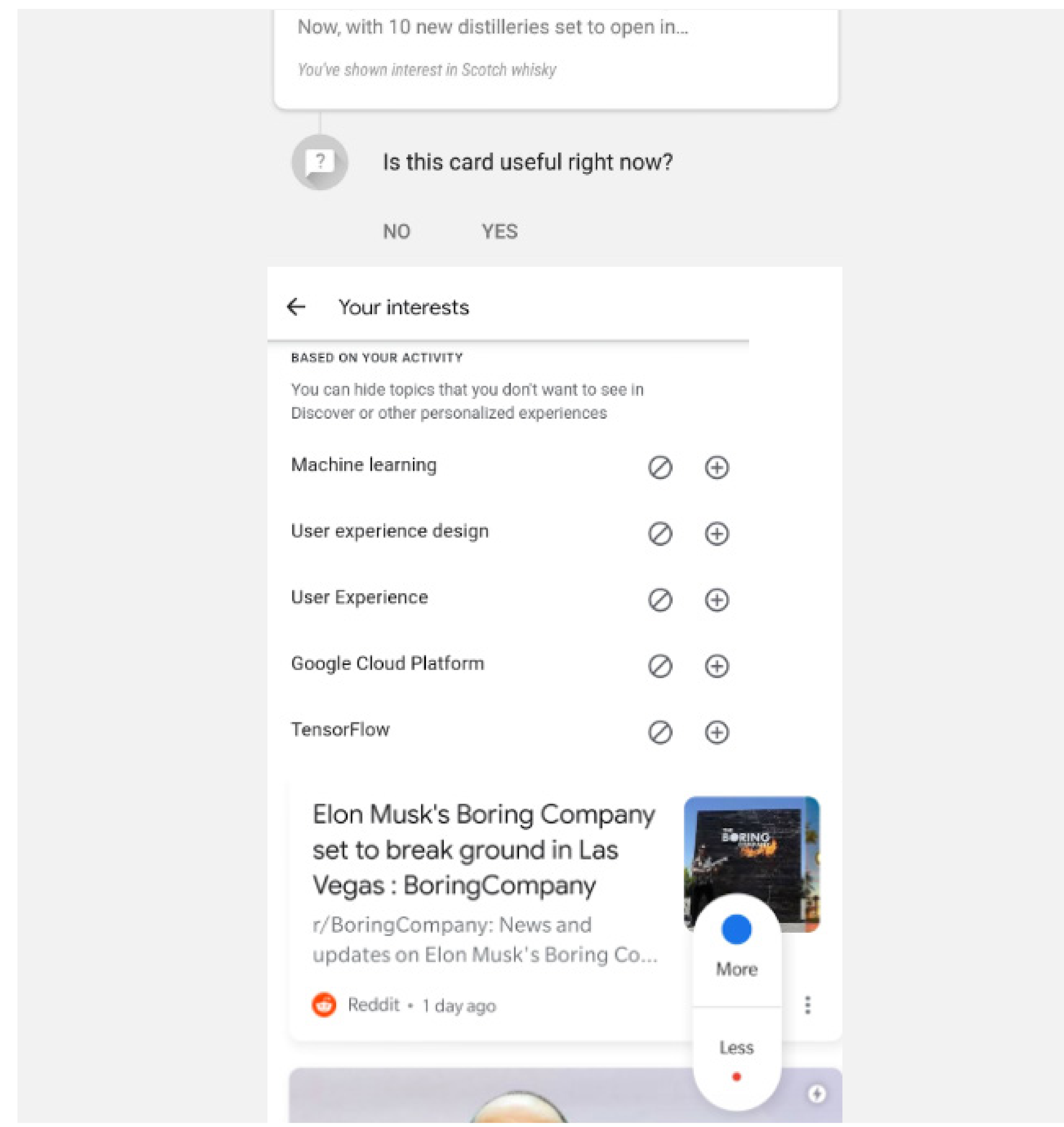
Zendesk provides service providers with a predicted satisfaction score on their customer's support tickets, so they can get a quick overview of the people who are most upset, or most pleased with their service, and can act accordingly. Next to the prediction there is a button to indicate when the model's predictions are wrong, and why.



Reporting inaccurate prediction scores - Source: Zendesk

→ Google Cards:

Google Cards exemplifies a simple way to collect valuable feedback to reward or penalize your model. Last month they added more granular feedback methods to train the algorithm on what an individual user wants and help them get more relevant suggestions.



5. User Controls + Customization

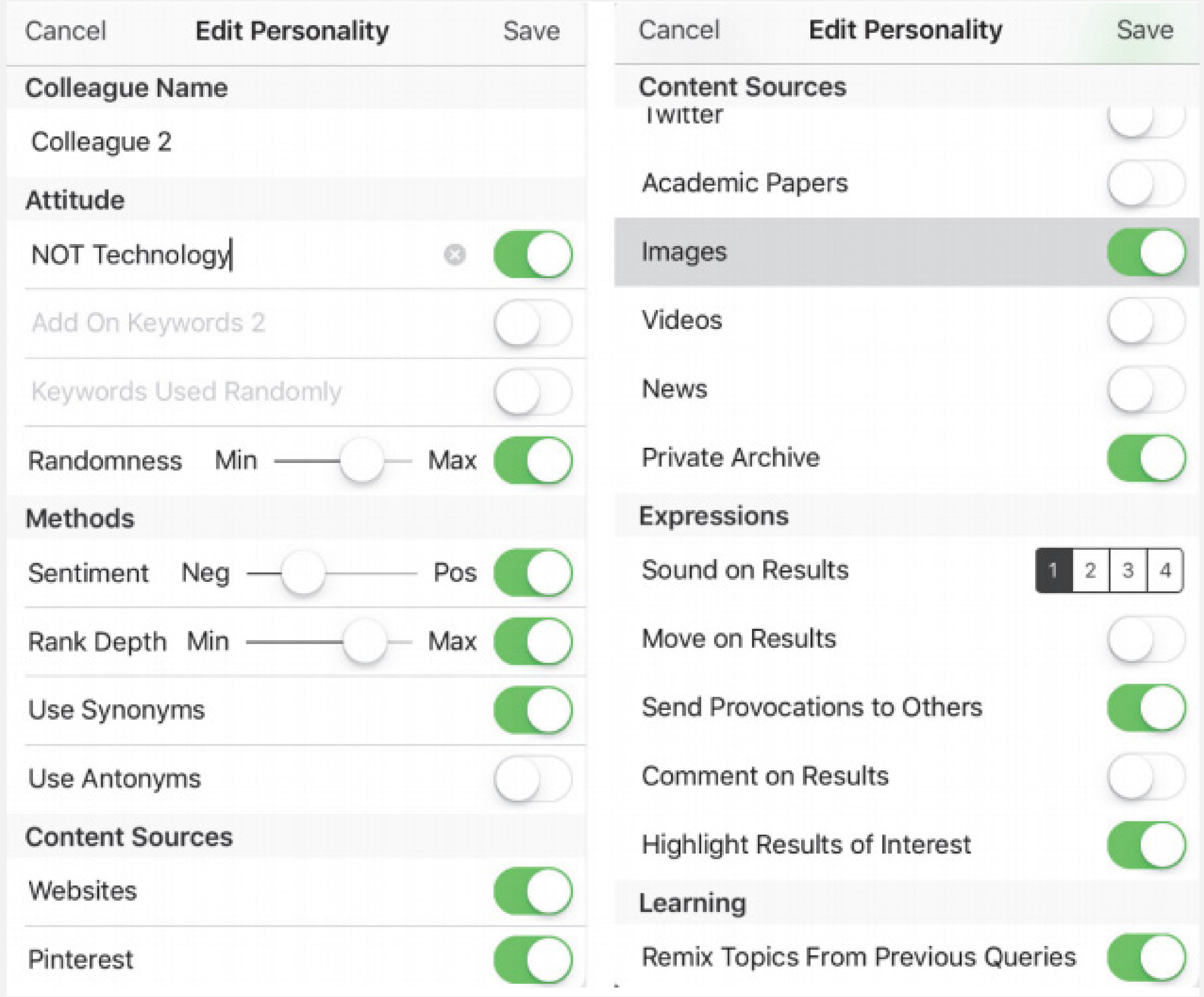
Giving users the controls to customize the model to their needs and intervene with the data or model if needed.

Design Strategy

Allow users to set intentions and configure parameters.

→ Personality Editor:

This speculative concept by Philip van Allen imagines what user controls for AI applications might look like. In this case it's a personality editor, but we can imagine similar interfaces for other applications.



Imagining the Goals and Methods of UX for ML/AI, Philip van Allen

6. Data privacy + security

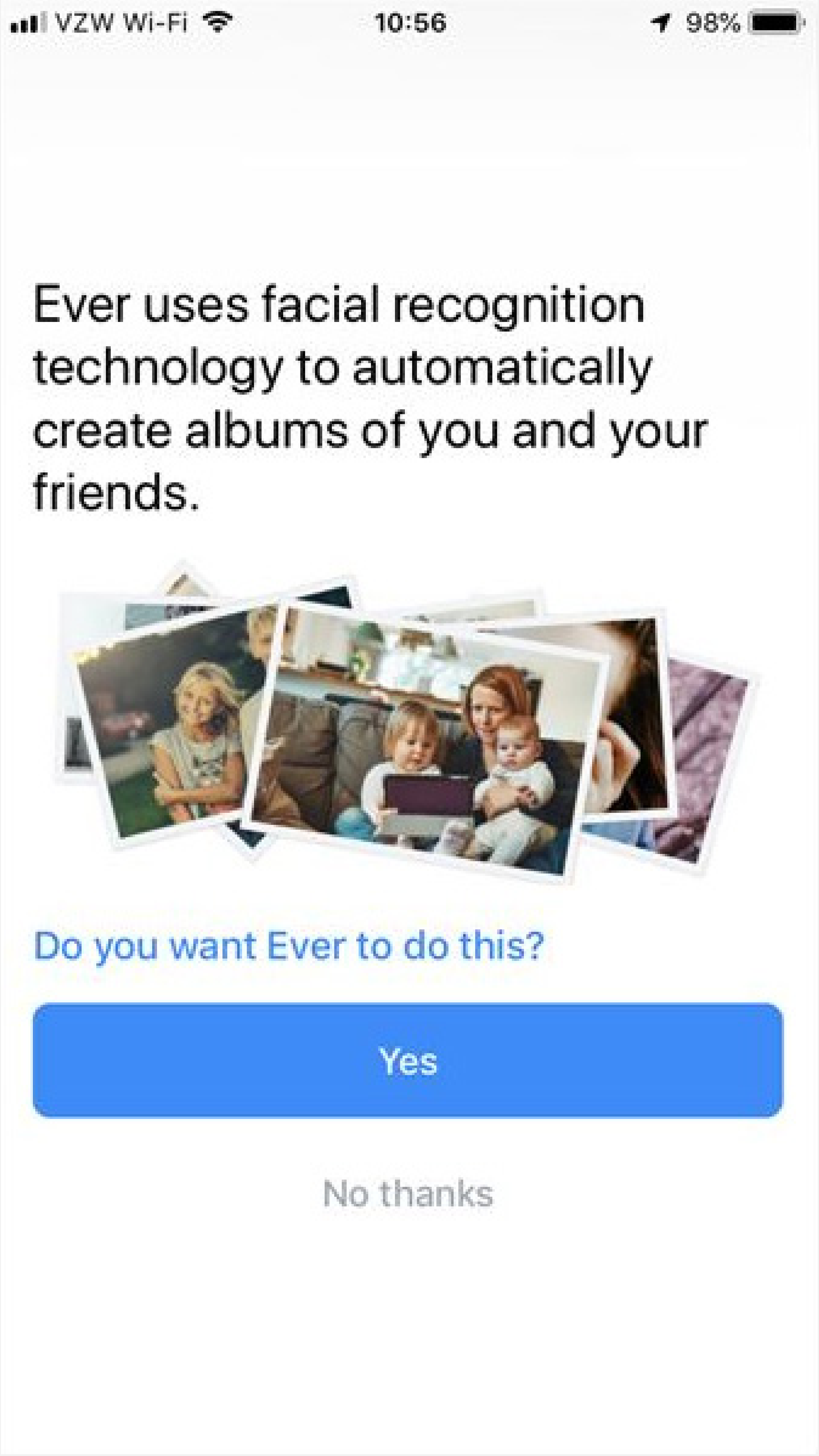
The need to collect, handle, and store user data with care, be transparent about who can access what data and why, while acknowledging their ownership.

Design Strategy

Communicating benefits per data share, allowing easy opt-in/out in a modular way, being cautious in sharing data, and making terms & conditions legible.

→ Ever:

Ever is an app to store and organize personal photos. When news came out that user data was used to train facial recognition algorithms, people were not pleased. It introduced this screen to communicate how it uses data, and gets explicit user consent, or allows them to easily opt-out.



△ Fail! Data acquisition

A recent controversy happened around a tech giant hiring contractors to collect more data for its facial recognition models. Homeless people were targeted because they seemed more likely to participate in exchange for a nominal cash reward. Their faces were captured, while they were asked to play a game, and used as training data without their informed consent.

Quote

“Human-centered design has expanded from the design of objects (industrial design) to the design of experiences (adding interaction design, visual design, and the design of spaces) and the next step will be the design of system behavior: the design of the algorithms that determine the behavior of automated or intelligent systems”

Frog CEO Harry West

Questions to consider

- **How to integrate and interpret user feedback?**

  Will it come from implicit signals or explicit interface actions?

- **To what extent should the user be able to customize the model?**

  How might we give the user controls to tune the algorithm to their needs?

- **How can we respect user data security, privacy, ownership?**

  How can we, and should we, make consent more explicit?

- **How can the user view, edit and wipe their data profile if it does not represent them?**

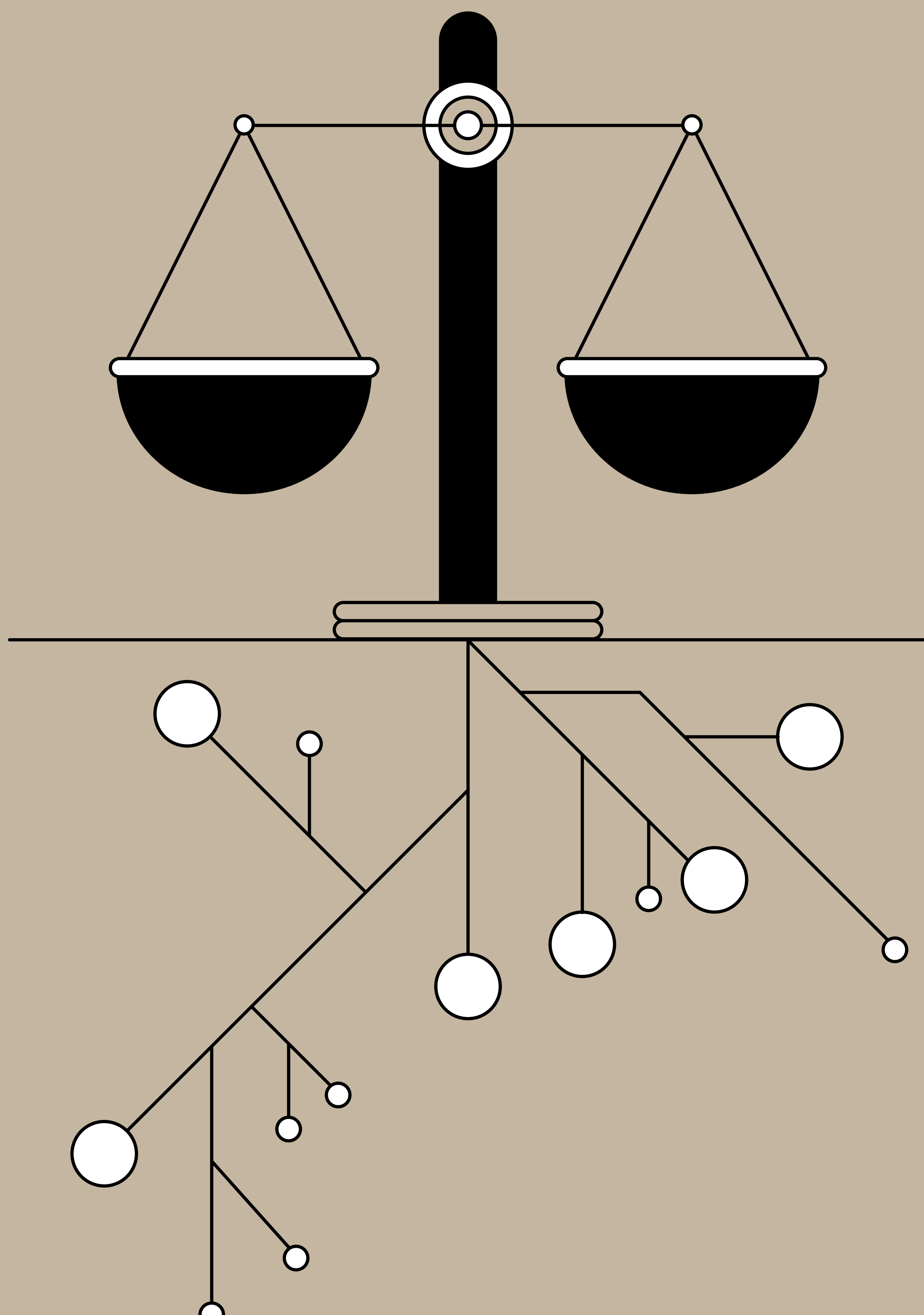
- **Is the data used beyond the service itself?**

- **How is it protected?**

- **Is it anonymized?**

Theme 3:

Value Alignment



Deploying AI systems across layers of society will affect the lives of individuals and groups across the globe in different and sometimes unexpected ways. Operating at an unprecedented scale and complexity, we must be mindful of biases, risks, system dynamics, and consequences, to make thoughtful trade-offs in our AI applications. Striving for value alignment between man and machine (and those operating the machine!) by integrating ethics at the core of our projects is required to shape this technology to help humanity

Why is value alignment important?

Otherwise, what's the point? Ethics. Impact. Fairness. Human-centered. Prevent harm and reinforcing harmful bias. We've been messing around for too long already.

7. Computational Virtue

Translating subjective human needs, values, and experiences into algorithmic parameters the model can optimize for.

Design Strategy

Bench marking usefulness based on use case rather than what's happening in research. Sometimes the model is nowhere near perfect, but as long as it's better than humans (it's more accurate and/or faster and cheaper) there's value.

→ Google Clips

Google Clips set out to develop a camera that would automatically capture memorable moments in the life of young parents. An incredibly subjective and context-dependent task, it required lengthy human discussions to agree on what the qualities of memorable moments were, and relied on extensive human training to guide the machine's learning to adopt this understanding.

8. Bias + Inclusivity

Mitigating harmful bias and guarding inclusivity in data and models to ensure fair treatment for all.

Design Strategy

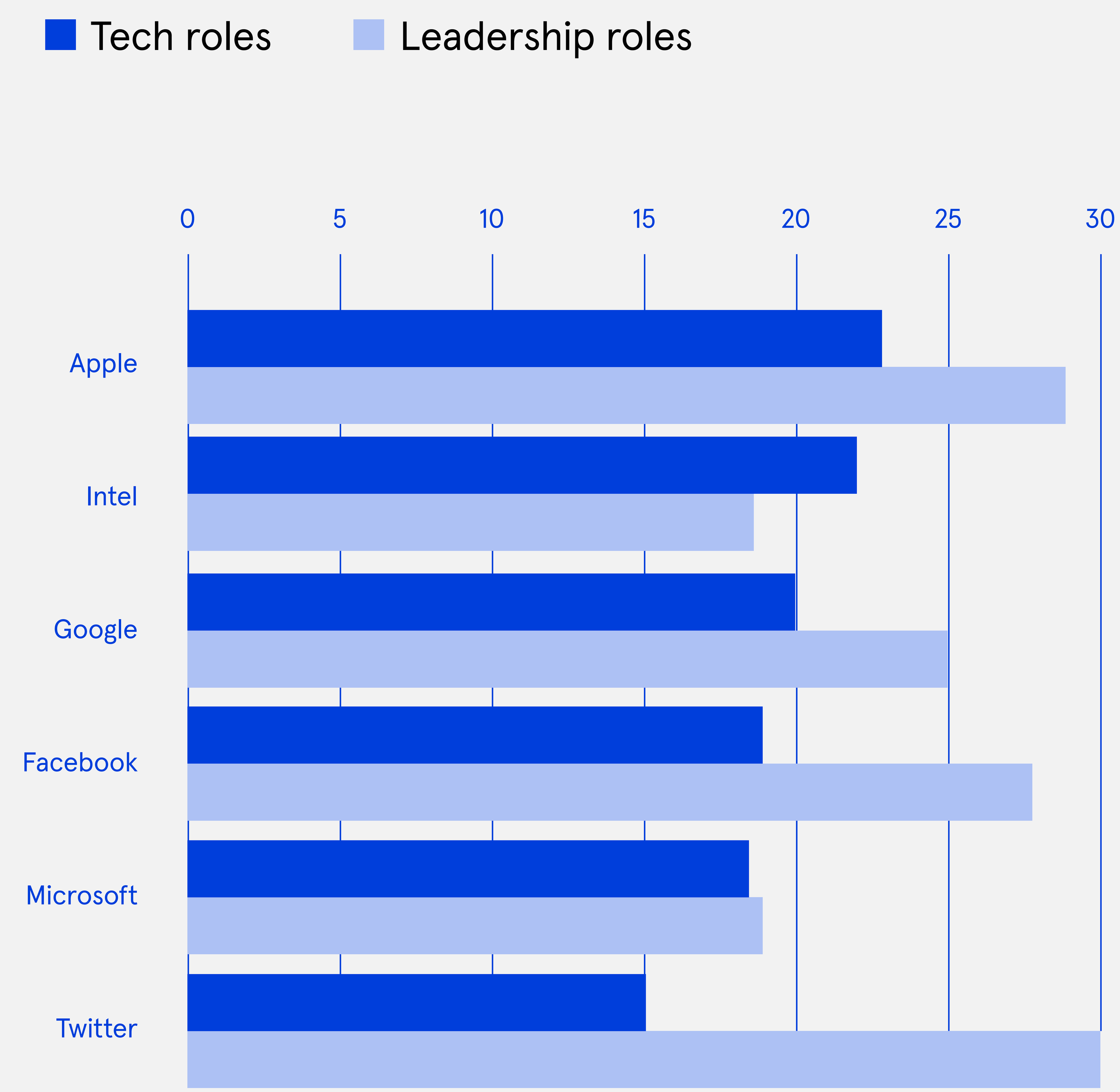
Checking for common unconscious bias, and having an inter-sectional team and user testing group (diverse in terms of gender and race, but also age, digital literacy, sexuality, level of education, lifestyle, political/religious beliefs, and other variables that might be relevant for your case).

→ Hiring gender bias

The historical data we feed AI to learn about the world, might not always represent the present we inhabit, or future we wish to manifest. A clear example is recruitment models discriminating against women and minorities. The algorithm doesn't favor anyone in particular It simply learns from past data in which majority of hires were white males, and perpetuates the pattern. Ensuring fairness in your model requires regular audits to detect and correct any harmful bias.

How big tech companies compare in employing women

Share of tech and leadership roles held by women (%)



Source: companies
©FT

9. Ethics + (Un)Intended Consequences

Unprecedented scale, speed and complexity call for a new level of thoughtfulness and responsibility in anticipating impact and (un)intended consequences.

Design Strategy

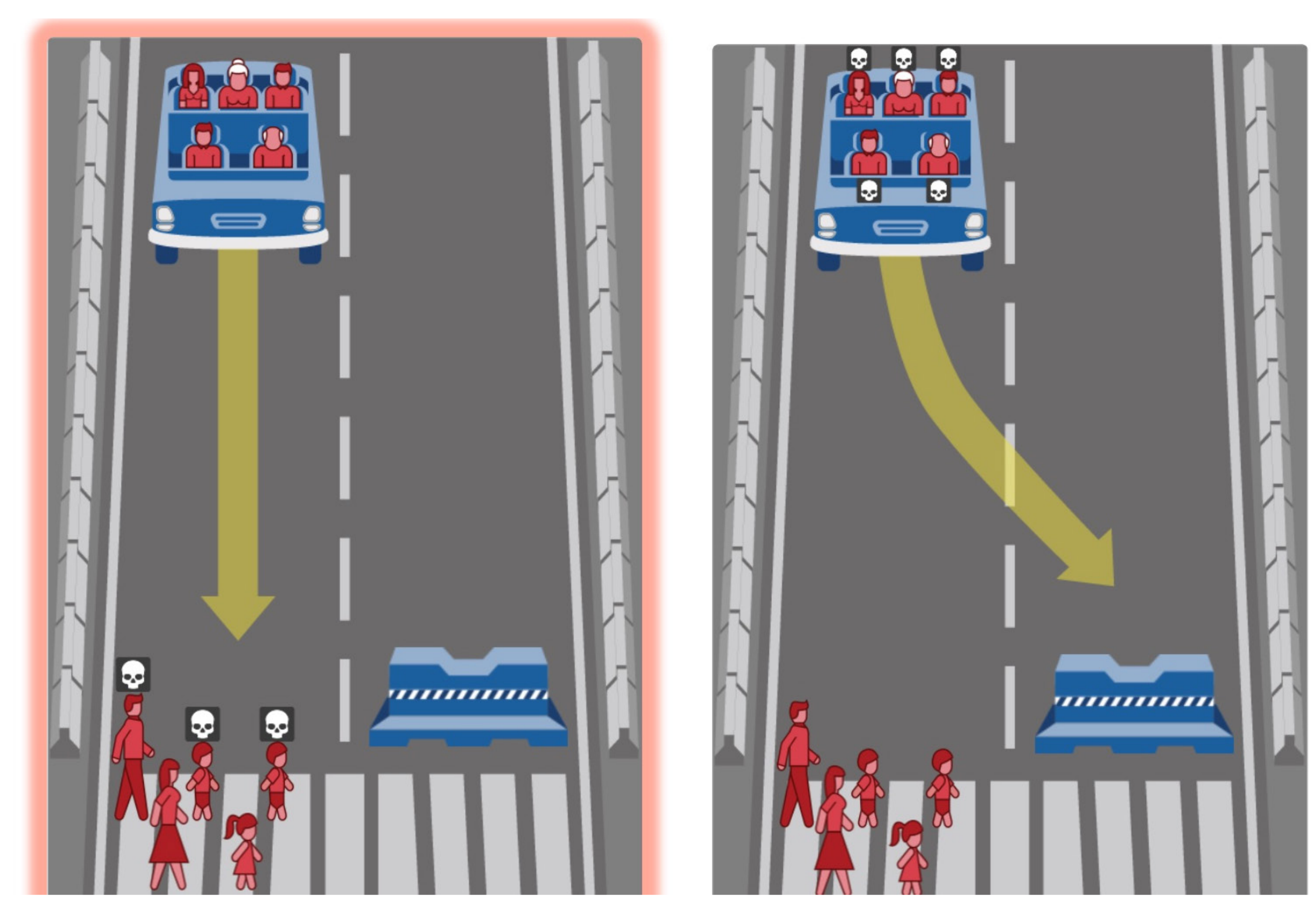
Recognizing good intentions does not equal positive impact. Be critical about anticipating potential consequences from various lenses, for example by using the consequence wheel, which is at times non-compatible with the ideology of capitalism and Silicon Valley's 'moving fast and breaking things'.

→ Affectiva

Affectiva is a market-leader in emotion recognition software that originated from MIT's Media Lab. While a lot of the tech industry is catching up to make facial recognition work well across ethnicities, Affectiva faces the additional challenge of recognizing emotions across cultures. Committed to serving clients across the globe, they had to ensure their models would generalize well across a diverse population. It requires extensive efforts and resources to understand facial and physiological expressions, in order to build data sets that are inclusive.

→ MIT Moral Machine

The MIT Moral Machine was built to gather human perspectives on moral decisions made by machine intelligence, such as self-driving cars. Anyone visiting the website is presented with a scenario in which the car has messed up, and now (in this case, guided by you) it has to choose who to kill. While the majority of us have intuitions, such as killing an older person over a child, making them explicit, (considering differences across cultures, and potentially activating them as a blueprint for machine moral decision-making) is a reality that's pretty hard to come to terms with.



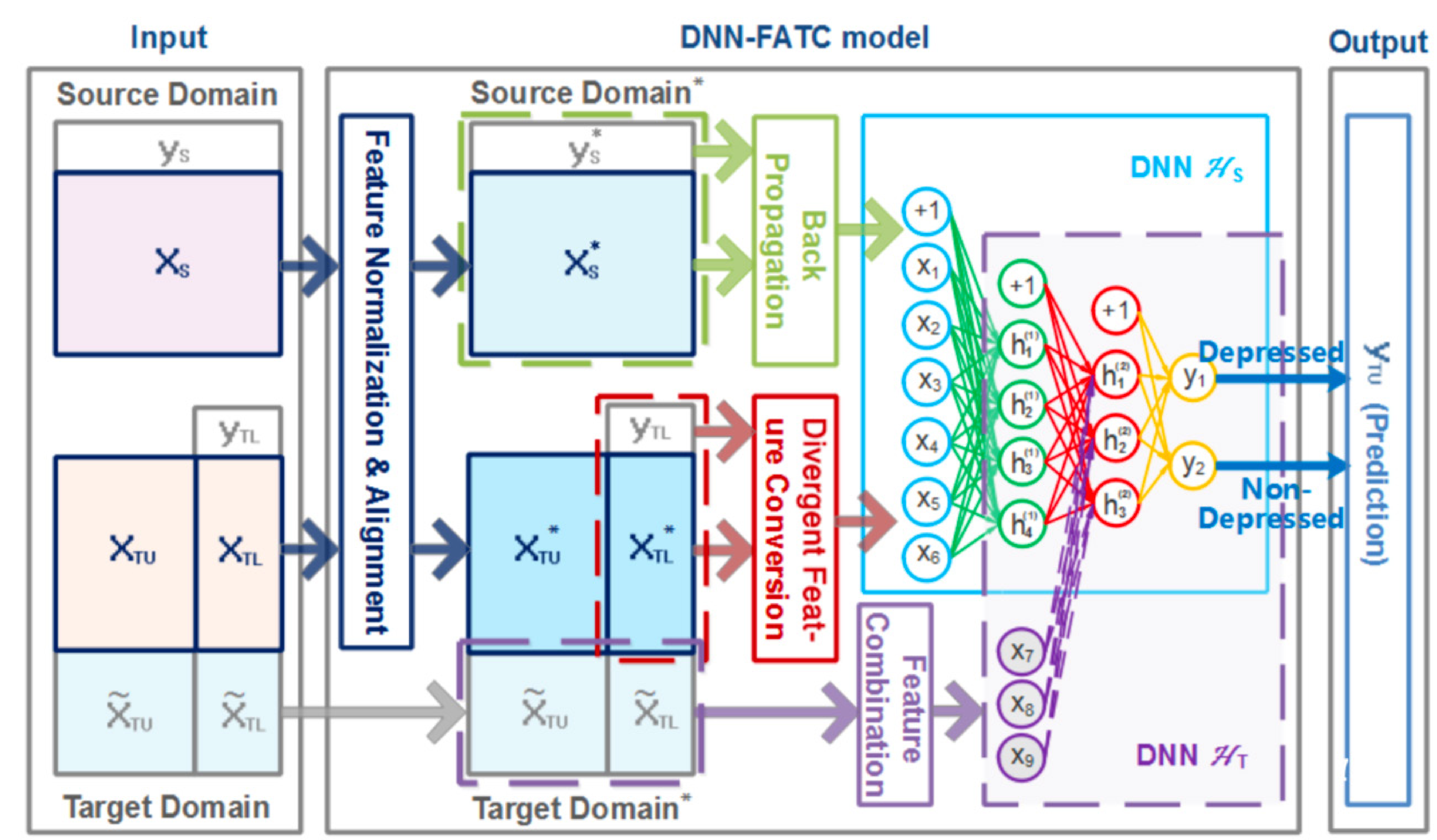
What should the self-driving car do?

→ Predicting mental health

While some moral choices appear obvious, many challenges around AI ethics sit in a gray area, where right or wrong it is not always obvious.

Aspiring to provide preventative and early mental health care, healthcare providers have successfully built models that predict the likelihood of depression and off-set of manic episodes in people with bipolar disorder, from social media data. Even operating from the noblest intention, making such inferences poses complex challenges.

What if the model is wrong and the person, or others, start questioning their well-being, it could become a self-fulfilling prophecy? If the model is right would insurers (be allowed to) treat you differently knowing you're at risk of mental illness? If your health insurer can infer such predictions from public social media data, could your employer? Could this information affect the hiring and firing process? Could advertisers use the predictions to target those in a vulnerable state. It's hard to 'unsee' data and well-intended endeavours bring about a range of ethical dilemmas. Where do we draw the line? In which cases is it (un)ethical to collect, infer, and act on such data?



Cross-Domain Depression Detection via Harvesting Social Media.

Quote

“Now is our opportunity to shape that future by putting humanists and social scientists alongside people who are developing artificial intelligence”

Marc Tessier-Lavigne

Questions to consider

- How do we translate subjective human experience into models and algorithmic parameters?
- What to predict?
- What objectives to optimize for?
- How to protect the inefficiencies that make the human experience meaningful from being optimized to no end?
- How do we benchmark and evaluate our models?
- How to design for positive, and anticipate and respond well to negative experiential, cultural, societal impact?
- How can we mitigate harmful bias and ensure fair treatment for everyone?
- How do we prevent destructive past patterns from leaking into our future models?
- Who owns and has access to the data, the models, knowledge, and computational power?
- How do we deal with power shifts as a result?
- How can we anticipate unintended consequences?
- How do we evaluate our impact and reason our trade-offs considering there is no universal moral framework?
- Who is responsible for mistakes?
- How can we protect against malicious use?
- Is it more dangerous to release research with the risk of malicious use, or to keep research private and centralize power?

Worksheet

If you are prototyping or building an AI-driven application, you can use the worksheet below to jot down and share first thoughts on how you're going to deal with the UX challenges.

Worksheet: UX of AI challenges

1. Explainability – How will we help our user understand certain outcomes?	2. Managing expectations – How will we establish realistic expectations?	3. Graceful failure & accountability – How will we design for trust in case of failure?	User Trust & Transparency
4. User feedback – How will your user provide feedback to the system?	5. User autonomy – How will the user be able to customize their experience?	6. Data privacy & security – How will you collect, store, and handle data?	User Autonomy & Control
7. Computational translation – How will you turn needs into parameters?	8. Bias & inclusivity – How will you prevent bias and guard inclusivity?	9. Ethics & (un)intended consequences – How will you look out for negative and positive impact?	Value Alignment
10. Which other (design) challenges do you foresee?			

Outro

As you've seen throughout this chapter, building human-centered AI applications is not an easy feat, and we've only just scratched the surface.

It's easy to become paralyzed by the scale, complexity, and urgency of these challenges. But considering you've come this far, I suspect you are not. Or perhaps you are, but are channeling your courage to engage with it for that exact reason!

I also suspect you have developed all sorts of thoughts and ideas around the challenges over the course of reading this chapter. I'd love to hear about those.

As it's set to impact all of us across life stages and at scale, designing human-centered AI is arguably one of the most interesting and important challenges of our time. It will require creativity, thoughtfulness, collaboration, and a commitment to shaping a future we want to live in. We'll need people from all walks of life and all areas of expertise to figure this one out. Are you up for the task?

If you decide to take part (in shaping the future, instead of delegating it to others and watching it happen), here are a few suggestions on how to get started.

1. Close to home

Some of these challenges might seem like remote futures, but they're not. Question if anything within your user journey is already influenced by algorithms, automation, and human-AI-interactions. Consider in what ways AI is likely to show up in your context and how these challenges appear alongside its opportunities.

2. Read up

A handful of projects that go into the user experience design of AI that I recommend you to read are included on the next page. Beyond that, there are great learning resources about AI from a more general perspective such as Andrew Ng's AI for Everyone on Coursera or elementsofAI.com.

3. Take a stance

In your work, as a designer or otherwise, how can you take on an active role in shaping these interactions with human values at its core? How can we move past principles, and begin building best practices around them? Besides creators, what is our role as users, as consumers, in demanding these elements of human-centered AI?

4. Join forces

Join fellow practitioners, reach out, share your knowledge and ideas, put them into practice.

If you're curious to explore the potential of AI within your projects, check out the AI meets Design toolkit for hands-on tools, exercises, and worksheets that integrate with the design thinking process.

This chapter is only a first block for you to build on. No one has the answers of what it means to build human-centered AI and how we might act on it, but we're committed to looking for and evolving them through design, dialogue, and democratization. I hope you join us!

Recommended Reading / Appendix

Experience Design in the Machine Learning Era - Fabien Girardin

Human-Centered AI Cheat Sheet by Josh Lovejoy

Human-Centered Machine Learning

Machine Learning and User Experience: A Few Resources by Michelle Carney

Guidelines for Human-AI Interaction

Projects by IF's Data Permissions Catalogue

IBM's Everyday Ethics for Artificial Intelligence

Google's People + AI Research - Library

Arefact's Tarot Cards of Tech

Fluxus Landscape of AI Ethics by Şerife Wong and Aparna Ashok

FOLLOWING CHAPTERS

This Brain food series will be released chapter-by-chapter, stretched over several months. In every chapter experts will dive deeper into specific topics related to AI.

awwwards.