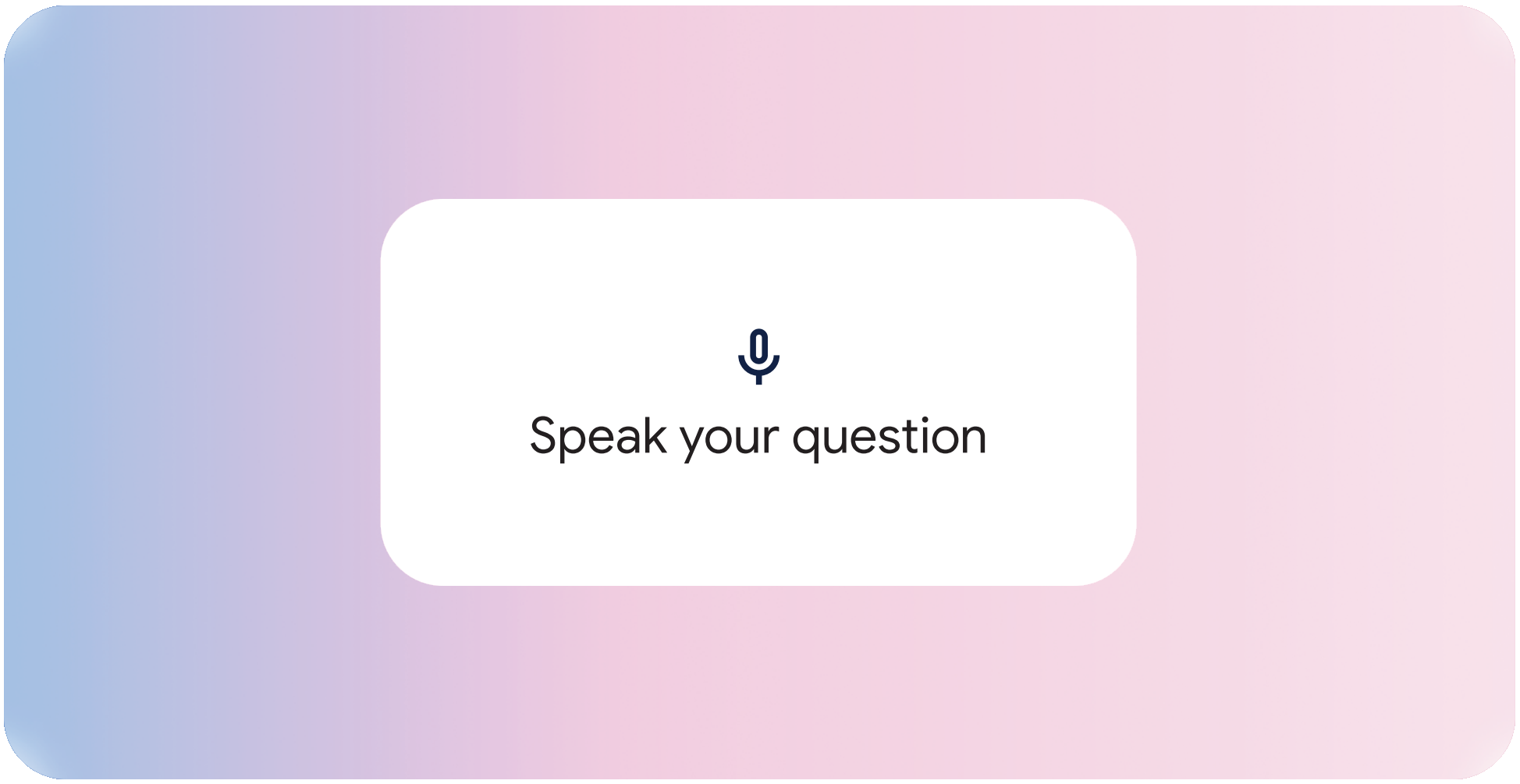
AI Principles Case Study

Responsible Development of

Lookout

Making Visual Media More Accessible With AI



November 2023

10 minutes

A new feature to give users the complete picture.

The Lookout App

Lookout, by Google, is an assistive Android app that uses a phone's camera to create accessibility tools for people who are blind or low vision (BLV). Lookout helps users complete common tasks by making the visual world more accessible. Its newest flagship feature — Image Q&A — enables users to not only get a much more detailed description of an image, but also to ask questions about a photo, and receive AI-powered responses.

Describing images is inherently challenging. If an image contains people, it's even more complex, as difficult questions arise about how to describe those people in a way that's both useful and respectful of a person's identity. Gender is a particularly challenging trait to describe based on an image. A person's gender may not be obvious from their appearance, and misgendering a person can cause severe harm.

Maximizing Benefit and Fairness

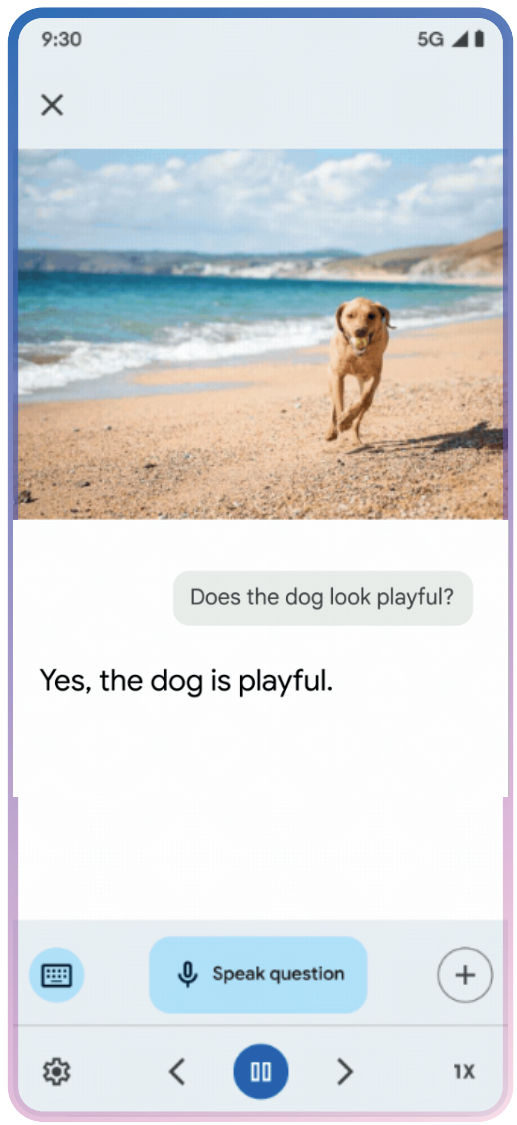
Lookout provided a particularly interesting challenge for the responsible development of AI, as AI Principles # 1 (Be socially beneficial) and # 2 (Avoid creating or reinforcing unfair bias) were somewhat in conflict with one another. Though it may be maximally beneficial to include gender in the description of a person, it also risks potential unfair bias if the person is misgendered in that description; the Lookout team knew they needed to balance maximizing the social benefit with avoiding potential unfair bias.

Previously, AI captioning models used an all-or-nothing approach with respect to perceived gender: either always report a person's gender, or never report it, using a neutral term like "person" instead. Neither of these approaches felt right to the Lookout team, and user feedback confirmed the need for a different solution. If an image contains a person the user knows, then reporting that person's gender is not necessary. But if an unfamiliar person's gender isn't reported, the team felt that that would disadvantage users with visual impairments, given that sighted people are often able to make their own judgements of someone's gender based on appearance. With this challenge in mind, the Lookout team engaged with the AI Principles review team.

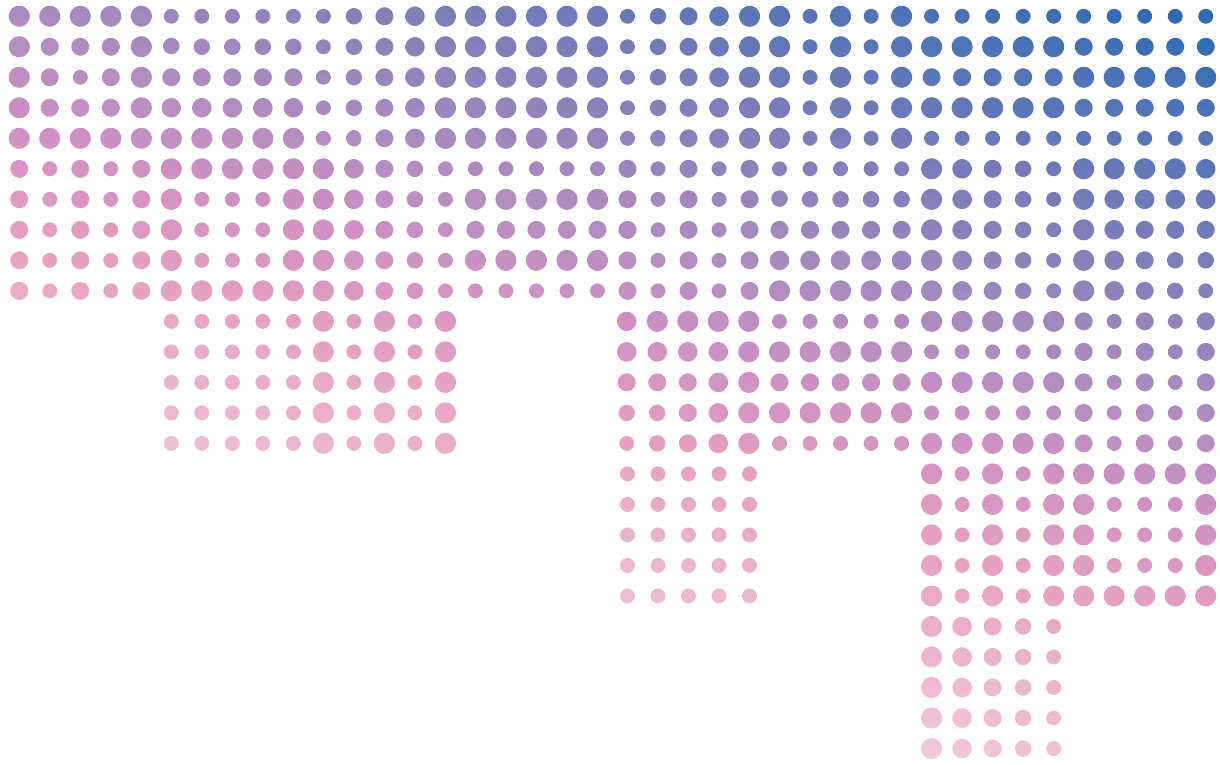
> We are aiming to answer questions about what's in the image with information that is objectively verifiable. The guardrails around that are to be responsible and humble. Avoid anything that is subjective. Avoid anything that is unknowable.

Product Manager

Google Lookout



Lookout is built in a way where a person can either type questions using the keyboard or dictate using their device's microphone. It's a straightforward way to quickly learn about an image and move on to the next task.



Finding A Path Forward

The Lookout team engaged in several AI Principles reviews over the course of the app's development and conducted adversarial fairness testing. The team incorporated a Google DeepMind visual language model (VLM), heavily customized for this use case, with several rounds of feedback from BLV people and from trans and non-binary people. VLMs enable people to ask natural language questions about an image. The new Lookout question and answer feature allows users to go beyond captions and ask about the image details that matter to them the most. This functionality allows the team to provide captions without perceived gender, but if the user asks a question about a person's gender, the model can provide a best guess of perceived gender, using cues from the person's appearance. In this way, Lookout can avoid providing gender descriptions when they're not necessary, reducing potential misgenderings, but the app can provide this information when users decide it's useful to them. The Lookout team tested this approach with end users who were BLV and non-binary and found that these users thought the approach was both useful and respectful.

The approach isn't perfect. The model will still make mistakes with perceived gender, and people with visual impairments still need to request details that typically sighted people receive effortlessly. The Lookout team believes this launch is both a step in the right direction, and an area where we can continue to learn and improve with the BLV community.

There remains much more work to do in this space, and Google is committed to continually learning from all stakeholders, to build tools that are both respectful and useful.

The inspiration for Lookout and the products we build for people with low vision or blindness is just how many physical, cultural, and societal barriers there are around us. Google investigates what some of those barriers are and breaks them down in various ways. The team asked ourselves how we could break down societal and technological barriers when it comes to sharing images.

Sr. Product Marketing Lead

Google Lookout

AI Principles guiding the Lookout team:

- **AI Principle # 1 (Be socially beneficial)**

  Partner with Google DeepMind, incorporating feedback from the BLV community to implement a VLM to improve accessibility for millions of users.

- **AI Principle # 2 (Avoid creating or reinforcing unfair bias)**

  Provide information about images that is objectively verifiable, avoiding anything that is subjective or unknowable.

- **AI Principle # 4 (Be accountable to people) & AI Principle # 7 (Be made available for uses that accord with these principles)**

  Test and evolve the product experience based on actual user needs and direct feedback, such as the shift from an open camera experience where Lookout interpreted the surrounding world to one where users can ask questions about an image.

Key Takeaways

Google has company-wide responsible innovation practices that draw upon more than 20 years of work in machine learning (ML) and over a decade of AI research. These practices support developers to incorporate fairness, safety, privacy, and transparency early in development. In addition, Google continues to:

- Prioritize research on societal risks that AI systems can pose, including avoiding harmful bias and discrimination, and protecting individual privacy through transparency and control of personal data.
- Comply with government initiatives, including the White House's industry commitments to ensure safe, secure, and trustworthy AI.
- Work with organizations like the National Institute of Science and Technology in addition to supporting forums such as the Ethical Considerations in Creative Applications of Computer Vision.
- Publish its AI Principles Progress Update report annually.

The Lookout app is an assistive Android app that makes the visual world more accessible to blind and low vision (BLV) people by using AI to describe images. Its new Image Question & Answer feature, powered by generative AI:

- Aims to maximize social benefit and minimize potential bias by allowing users to ask the model to make a best guess about a person's perceived gender.
- Uses a new type of visual language model (VLM) developed by Google DeepMind, with feedback from BLV and trans and non-binary users.
- Avoids the subjective and unknowable, generating responses that are inclusive, authoritative, and based in fact.

Learn more:

g.co/AI/ResponsiblePractices

Google's AI Principles

Since 2018, Google has used these AI Principles to guide the ethical development and use of technology:

- Be socially beneficial.
- Avoid creating or reinforcing unfair bias.
- Be built and tested for safety.
- Be accountable to people.
- Incorporate privacy design principles.
- Uphold high standards of scientific excellence.
- Be made available for use in accord with these principles.

In addition to our principles, Google will not design or deploy AI in the following application areas:

- Those likely to cause overall harm.
- Technologies primarily intended to cause injury.
- Surveillance violating internationally accepted norms.
- Purpose contravenes international law and human rights.

Google's Responsible Innovation Team, which produced this case study, is the company's central AI ethics governance team. It's composed of people with backgrounds in ethics, law, philosophy, research, and various social sciences such as linguistics, economics, political science, international studies, and religious studies.

Internally, we support Googlers in applying the seven principles through AI Principles reviews, education programs, workshops, and other engagements with product teams.

Additional details are available at https://ai.google/principles.