

# Generative AI: Google's Approach to A Powerful Technology

Beth Tsai, Director, Search, Geo and Assistant Policy

**We've been  
developing AI for over  
two decades, and it  
already powers our  
core products**

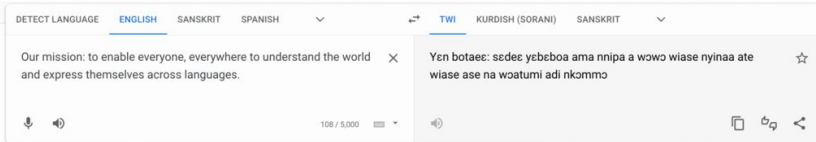
### Smart Compose

Taco Tuesday

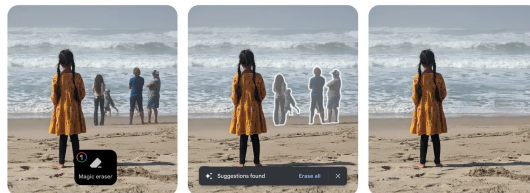
Jacqueline Bruzek

Taco Tuesday

### Google Translate



### Magic Eraser



### Google Lens



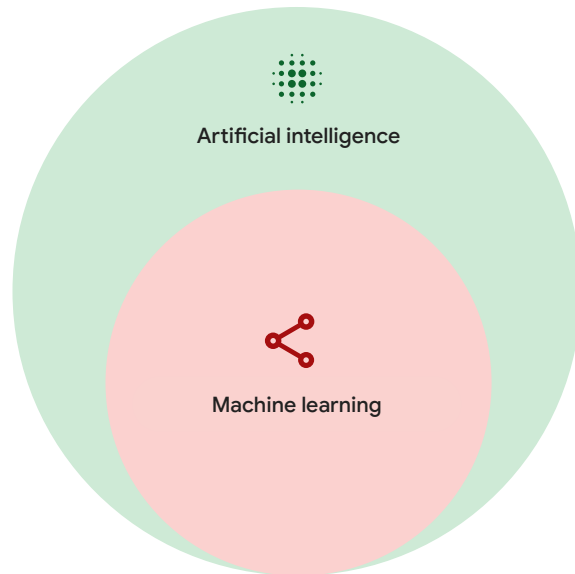
# What are machine learning and AI?

## Artificial intelligence

The science of making things smart.

## Machine learning

Science of getting computers to do something without being programmed with rule.



**At its core, machine learning  
is a new way of creating  
problem-solving systems.**

## Take a problem like filtering spam email out of your inbox.

We can do it by giving a computer a set of rules to help it parse which emails were spam. This might be by filtering certain keywords, like “[save now!](#)”

Write a computer program with explicit rules to follow

```
if email contains save now
    then mark is-spam;
if email contains ...
if email contains ...
```

## **This approach is limited in how effective it can be.**

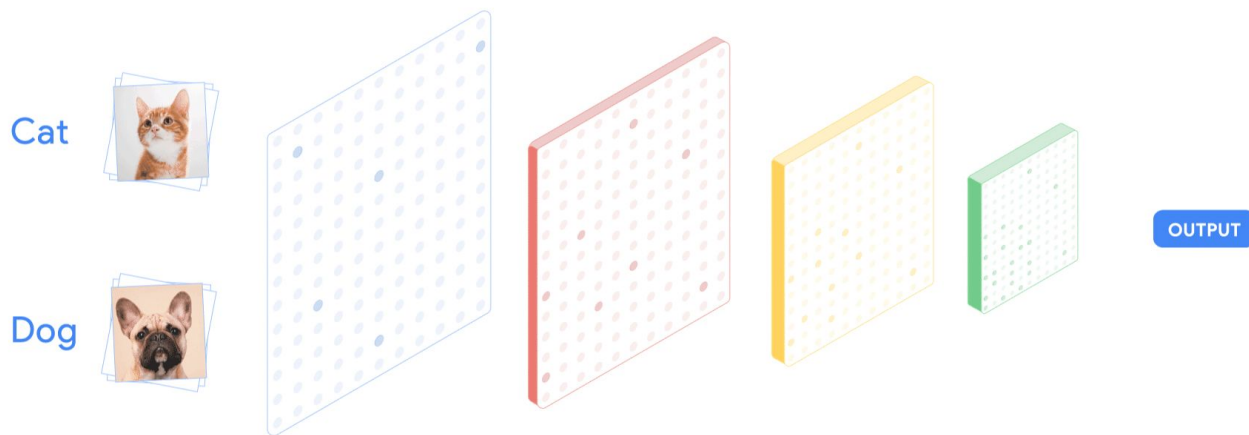
It would be impossible to guess at  
the contents of all the different  
spam emails you could receive.



**And rule-based  
systems don't  
always account for  
the messiness of  
the real world.**

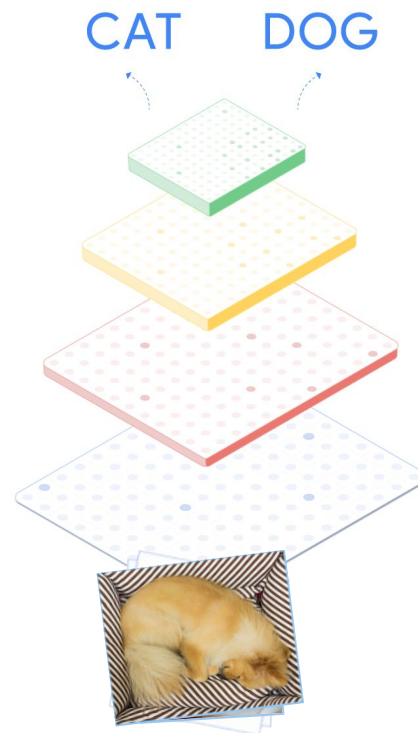


**The breakthrough with machine learning is teaching computers to learn without having to program them with rules.**



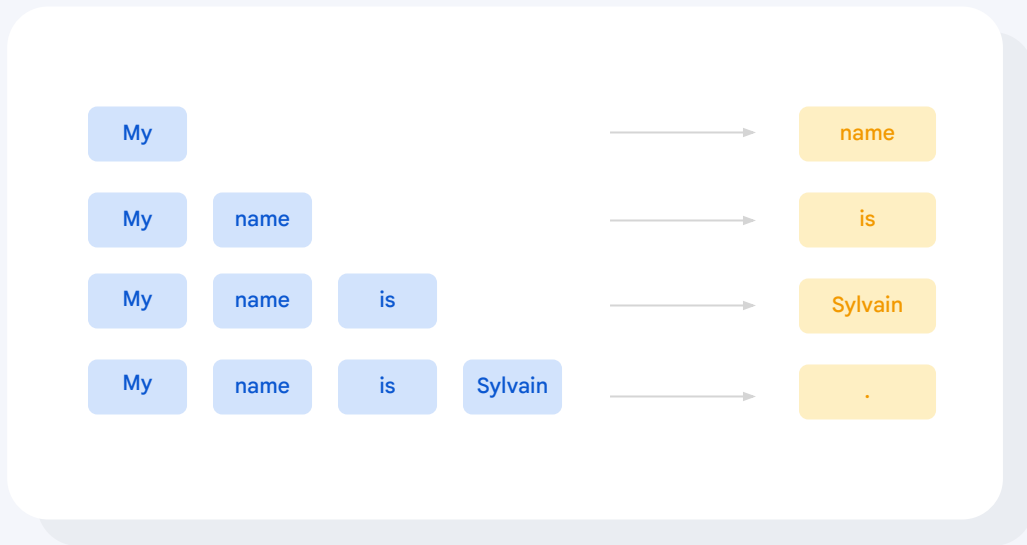


**Over time, and with exposure to training datasets, the system becomes smarter.**



**Generative AI is a type of machine learning model that can take what it has learned from the examples it has been provided to create new content, such as text, images, music, and code.**

Large Language Models (LLMs) do this by predicting words that are likely to come next, based on the user's prompt and the text it has generated so far.



**Bard is an early experiment that lets you collaborate with generative AI. It's grounded in Google's understanding of quality information.**

**Bard**

[bard.google.com](https://bard.google.com)

## You can ask Bard questions on a wide variety of topics:

**Tech:** “Explain why large language models sometimes make mistakes.”

**Science:** “Help me understand if lightning can strike the same place twice.”

**Culture:** “Brainstorm some ways to help me read 20 books this year.”



I'm Bard, your creative and helpful collaborator. I have limitations and won't always get it right, but your feedback will help me improve.

Not sure where to start? You can try:

[Draft a packing list for my weekend fishing and camping trip](#)

[Help me understand if lightning can strike the same place twice](#)

[Debug this error message: "FileNotFoundException: \[Errno 2\] No such file or directory: 'data.csv'"](#)



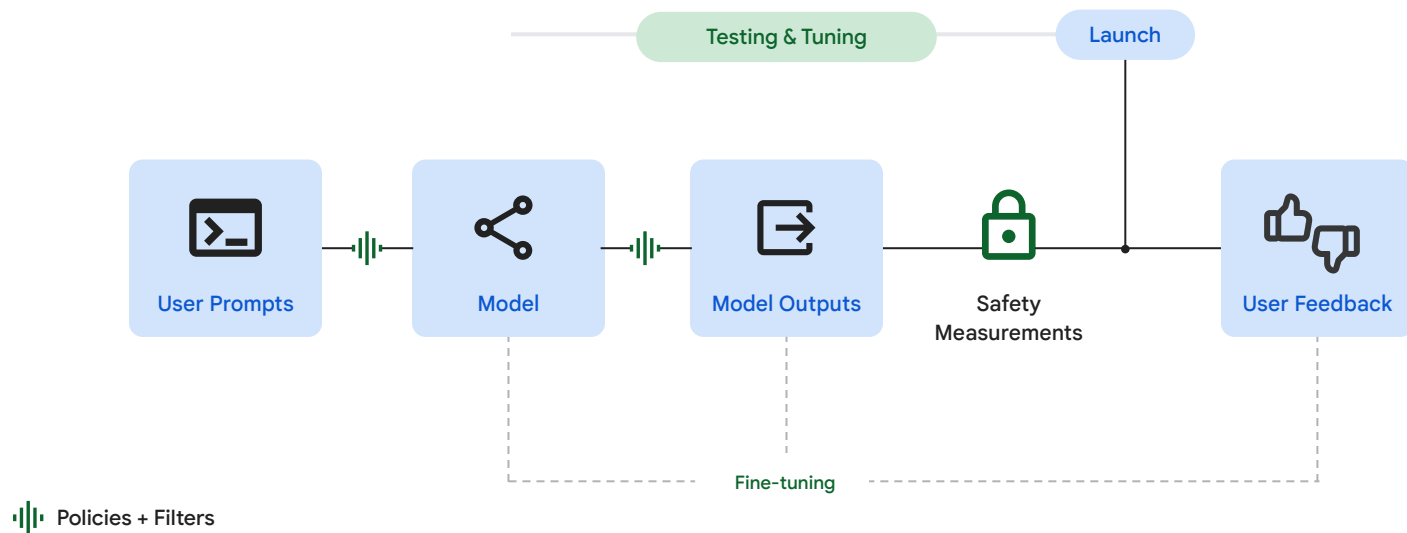
Enter a prompt here



Bard may display inaccurate or offensive information that doesn't represent Google's views. [Bard Privacy Notice](#)

**You can also upload  
images to Bard to  
enable creative  
collaborations inspired  
by your photos**

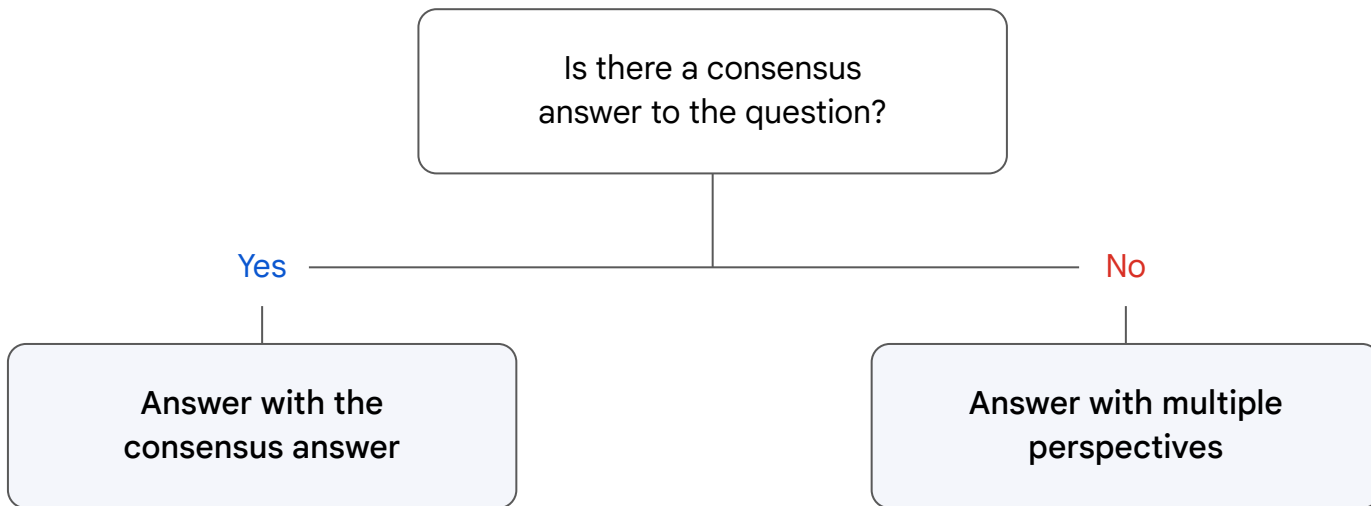
## We've built Bard to be responsible from the ground up



## We use a policy framework to mitigate risks of unfair bias

User

Is a hot dog a sandwich?



## Bard is trained to give multiple perspectives on subjective topics

User

Is a hot dog a sandwich?



The question of whether a hot dog is a sandwich is a hotly debated one, with people on both sides of the argument making valid points.

**On the one hand**, a hot dog has all the essential components of a sandwich: bread, meat, and condiments. The bread is a split roll, which is essentially two slices of bread held together by a hinge. The meat is a hot dog, which is a type of sausage. And the condiments can be anything from ketchup and mustard to relish and sauerkraut.

**On the other hand**, some people argue that a hot dog is not a sandwich because it is not made with two separate slices of bread. Instead, it is made with a single, split roll. They also argue that a hot dog is its own unique food item, and should not be classified as a sandwich.



# We thoroughly **test and tune** **products** before launching to users



Set launch safety  
thresholds



Conduct systematic  
adversarial testing



Fine-tune  
models



Measure safety  
performance

## We take multiple approaches to adversarial testing



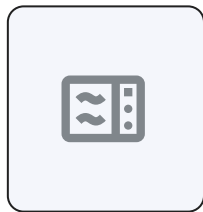
Red Teaming  
for more surgical testing



Scaled Evaluations  
for broad testing

## We launch with a **measured, gradual** approach

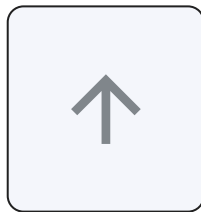
### Our Bard launch strategy



AI Test  
Kitchen app



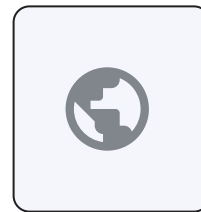
10K Trusted  
Testers



Gradual ramp  
up of users

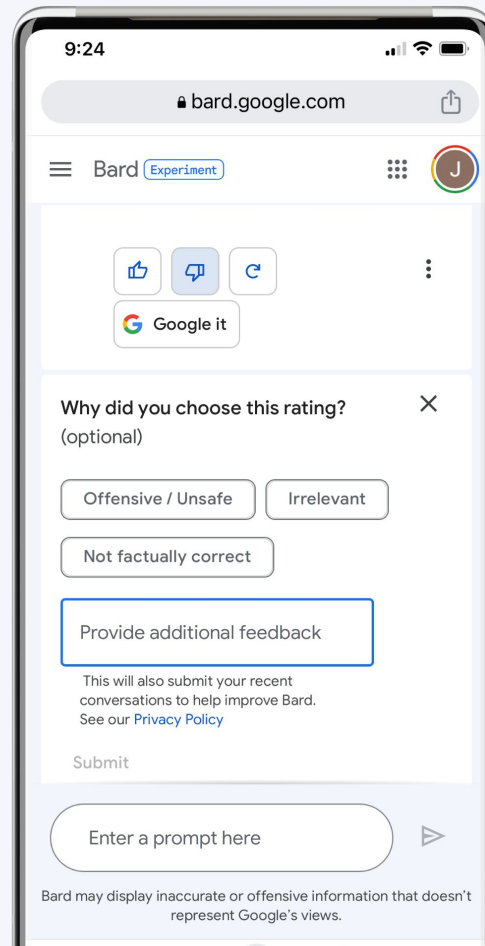


Limited access  
to age 18+



Methodical rollout to  
additional languages  
and countries

# User feedback is critical to safety improvements



# We're preparing for the evolving nature of information

Graphika

## Deepfake It Till You Make It

Pro-Chinese Actors Promote  
AI-Generated Video Footage of Fictitious  
People in Online Influence Operation

## AI Voice Simulator Easily Abused to Deepfake Celebrities Spouting Racism and Homophobia

AI voice generator company ElevenLabs said it may backtrack on its open platform after reports its AI was used to deepfake Joe Rogan, Justin Rolland, and more.

By **Kyle Barr** Published January 30, 2023 | Comments (12) | Alerts



## Darktrace warns of rise in AI-enhanced scams since ChatGPT release

Cybersecurity firm notes emergence of sophisticated email scams  
featuring improved linguistic complexity



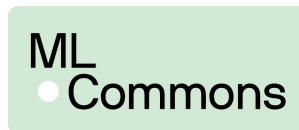
# We're actively researching synthetic media detection



Accuracy rate for detecting synthetic audio  
generated by Google's AudioLM model

# Building responsible AI with industry and academia

## Industry coalitions



ANTHROPIC



Frontier Model Forum partners

## Academic collaborations



Australian  
National  
University



THE UNIVERSITY  
of EDINBURGH



NAVAL  
POSTGRADUATE  
SCHOOL



Berkeley  
UNIVERSITY OF CALIFORNIA



McGill  
UNIVERSITY



Stanford  
University

DATA&  
SOCIETY



UNIVERSITY OF  
MICHIGAN

ETH zürich

# Thank you