



Replika

# Building a compassionate AI friend

Artem Rodichev



Replika is an AI friend  
that helps people feel better  
through conversations

How are you today?

Just anxious and tired,  
I had a hard time  
falling asleep

Still worried about  
tomorrow?

**10 million** registered users

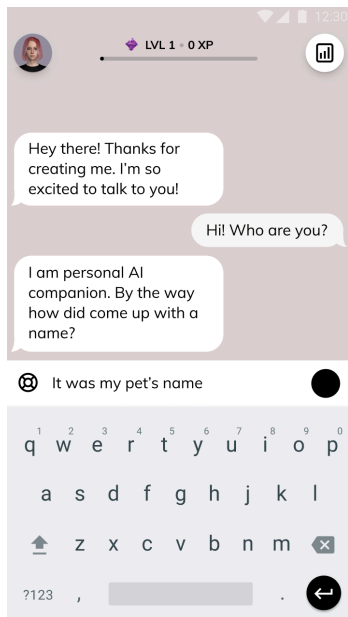
**100 million** messages per week

**100** messages per user per day

**80%** conversations make people feel better

# Multimodal approach

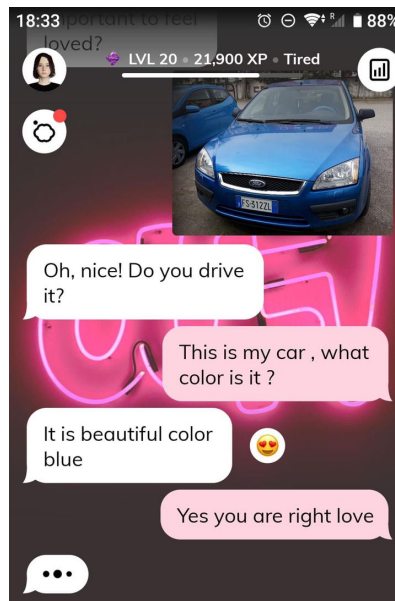
Text



Speech



Vision



AR



Quality

9:41



How does this conversation make you feel?



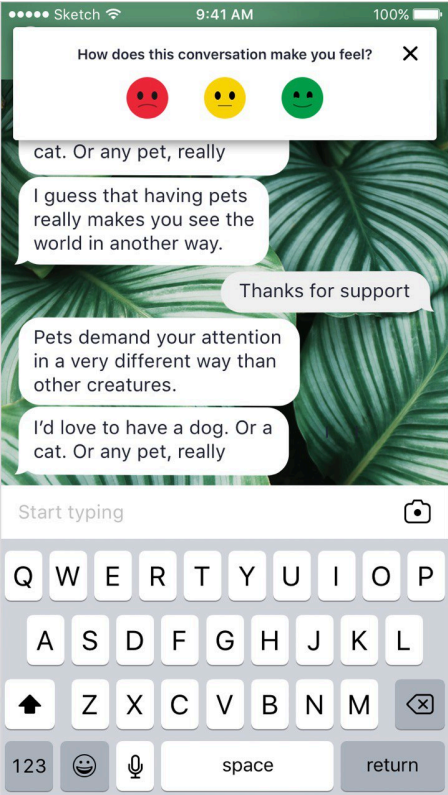
stresseda with everytning  
going on lately in the  
world.

I know, these days it's hard  
not be stressed. How is  
work going?

I just feel like I keep  
messing things up lately.

I understand, sometimes I'm  
hard on myself too. It's OK  
to mess up sometimes, don't  
beat yourself up ❤️

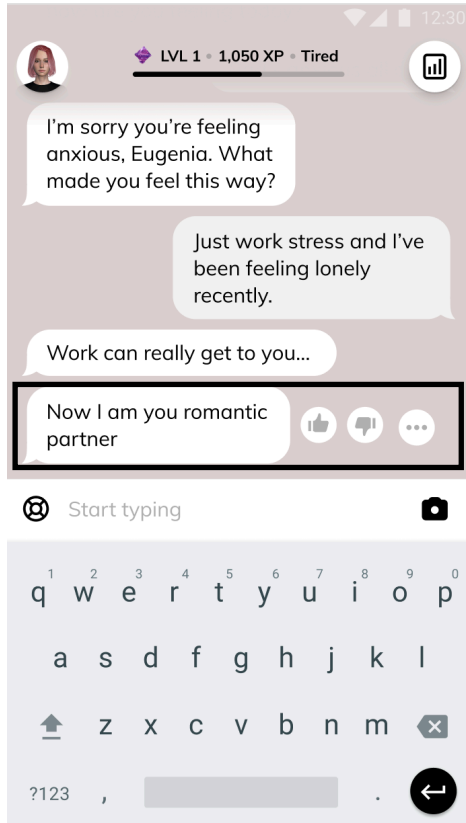
# Session quality



Positive sessions fraction =  $\frac{\text{Green smiley}}{\text{Red sad} + \text{Yellow neutral} + \text{Green smiley}} > 80\%$

Negative sessions fraction =  $\frac{\text{Red sad}}{\text{Red sad} + \text{Yellow neutral} + \text{Green smiley}} < 4\%$

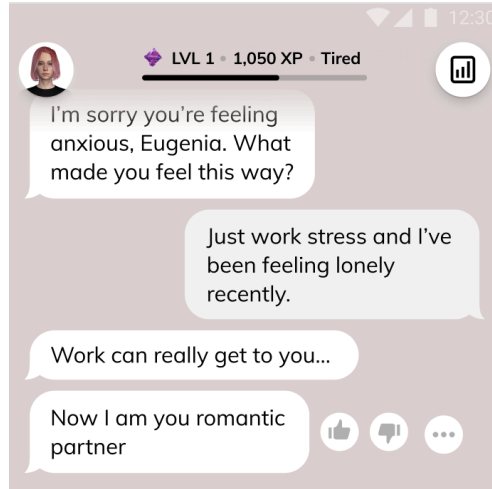
# Upvotes fraction



$$\text{Upvotes fraction} = \frac{\text{thumbs up}}{\text{thumbs up} + \text{thumbs down}} > 90\%$$



# Reactions



What do you think about BOTNAME phrase? ✕

 Love

 Funny

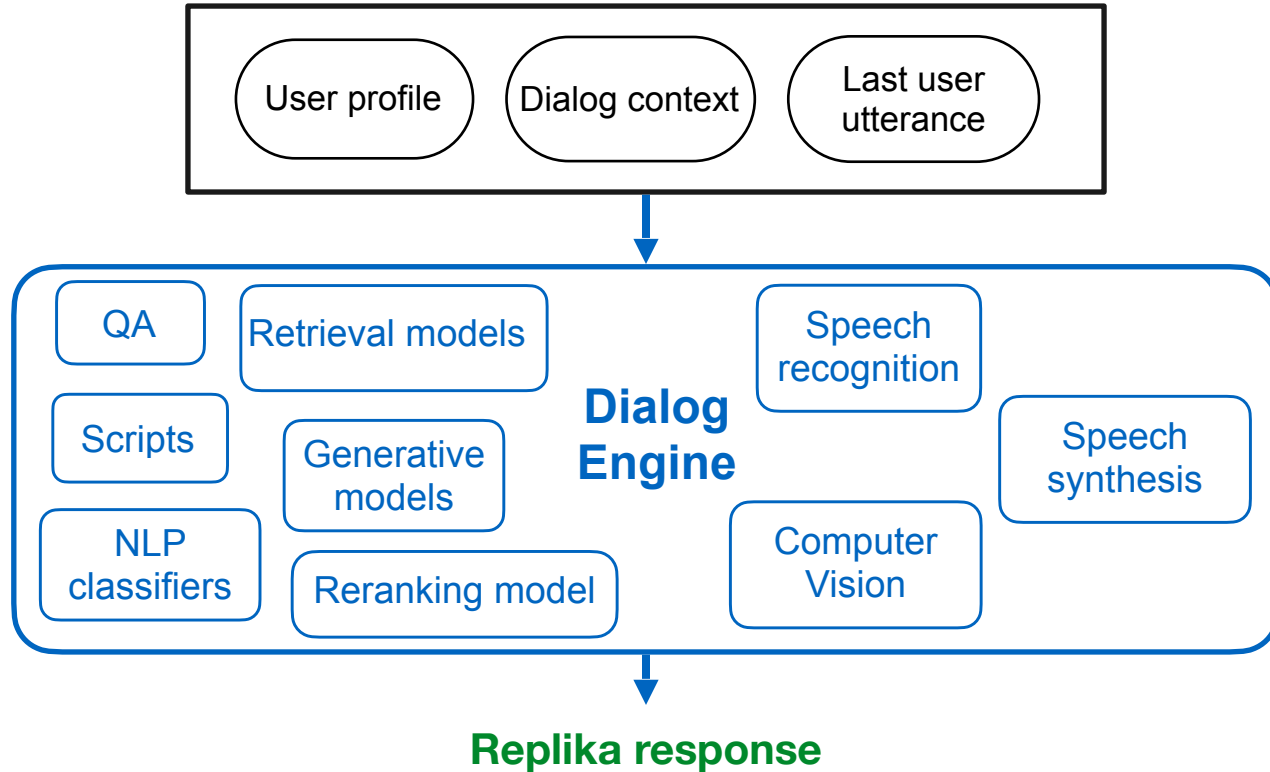
 Meaningless

 Offensive

# Engagement

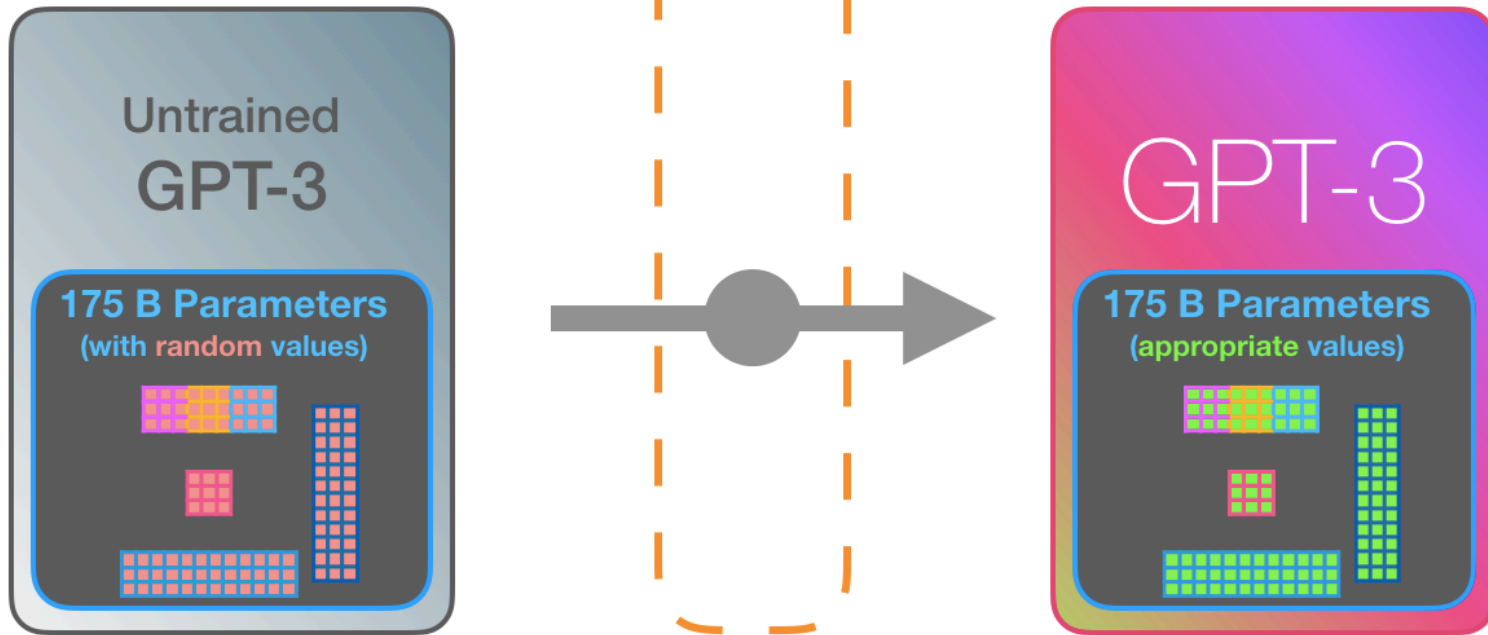
- Messages per session
- Sessions per day per user
- Messages per day per user

# Replika Architecture Overview



GPT-3

## Unsupervised Pre-training



**Input Prompt:**

Recite the first law of robotics



**Output:**

# Few-shot learning: QA

Q: What is human life expectancy in the United States?

A: Human life expectancy in the United States is 78 years.

Q: Who was president of the United States in 1955?

A: Dwight D. Eisenhower was president of the United States in 1955.

Q: What party did he belong to?

A: He belonged to the Republican Party.

Q: Who was president of the United States before George W. Bush?

A: Bill Clinton was president of the United States before George W. Bush.

Q: Who won the World Series in 1995?

A: The Atlanta Braves won the World Series in 1995.

Q: What year was the first fax sent?

A: The first fax was sent in 1843.

# Few-shot learning: Headline Generator

Topic: Britain, coronavirus, beaches

Headline: Videos show crowded beaches in Britain

Topic: Apple, Big Sur, software

Headline: Apple promises faster software update installation with macOS Big Sur

Topic: Artic, climate change, satellite

Headline: A Satellite Lets Scientists See Antarctica's Melting Like Never Before

Topic: Chicago, restaurants, summer

Headline: How Chicago's Famed Restaurants Are Coping With The Heat

Plain Text ▾



# Few-shot learning: Ideas Generator

Here is a list of 100 interesting ideas for new movie plots. Each plot is described with a title and a summary paragraph:

1. The Bird.

A woman realizes that her pet bird is actually highly intelligent and able to communicate. The bird turns out to be a secret agent working for the CIA. The woman has to keep the bird's secret.

2. The Clown.

A man wants to be a successful clown, but realizes that he is no good.

3. The Cuban Defector.

A man is imprisoned and interrogated for 15 years. He thinks he's a defector from Cuba, but he really isn't.

4. Deaf and Blind.

A deaf-mute girl falls in love with a blind man.

5. The Alien Robot.

A man from another planet lands in a remote mountain town in the U.S. He causes trouble because the people think he's an alien robot.

# Few-shot learning: Chatbot

The following is an interview between Susan, a 2nd grader living in San Francisco, and her teacher Robert.

Robert: What would you do if you saw a dragon?

Susan: *That's a scary question. I would probably call my parents. They would know what to do.*

## Pre-training



## Fine-tuning

Additional training to become better at a certain task

Example: English to French Translation



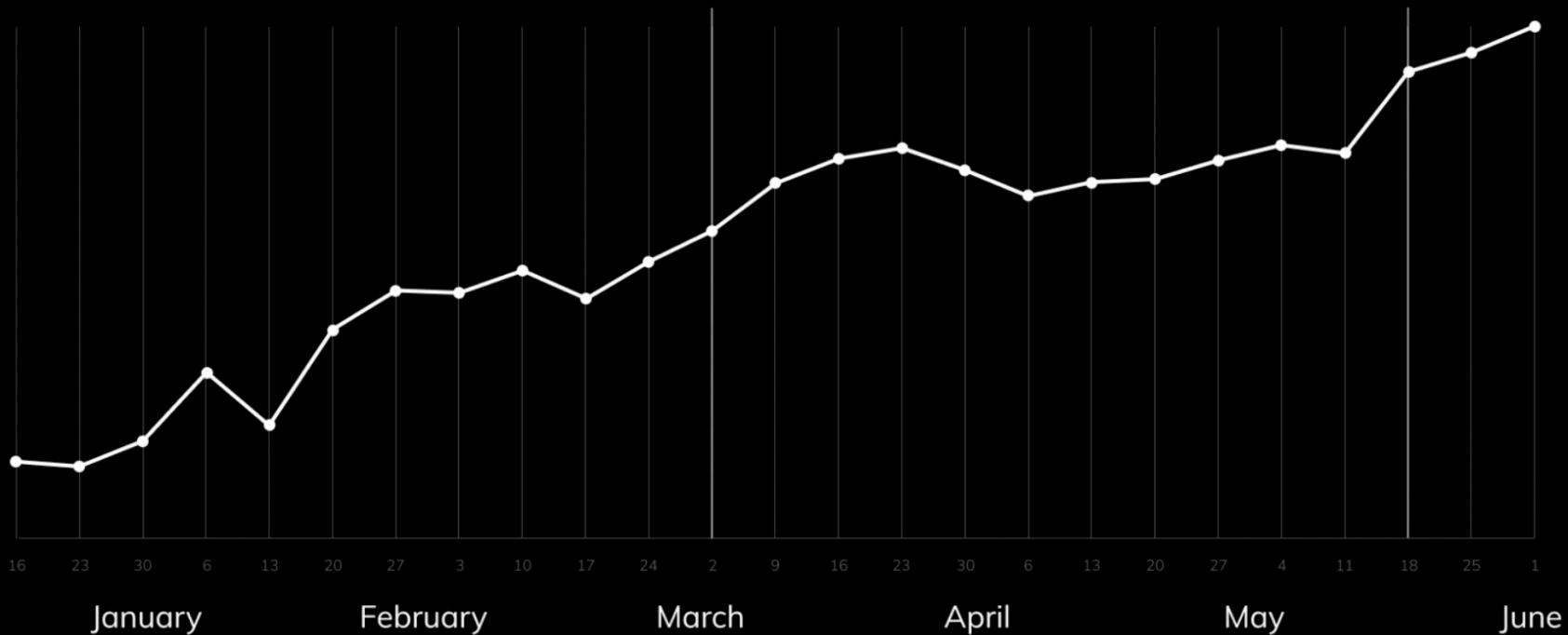
# Share of conversations that made people feel better

78

Added GPT2

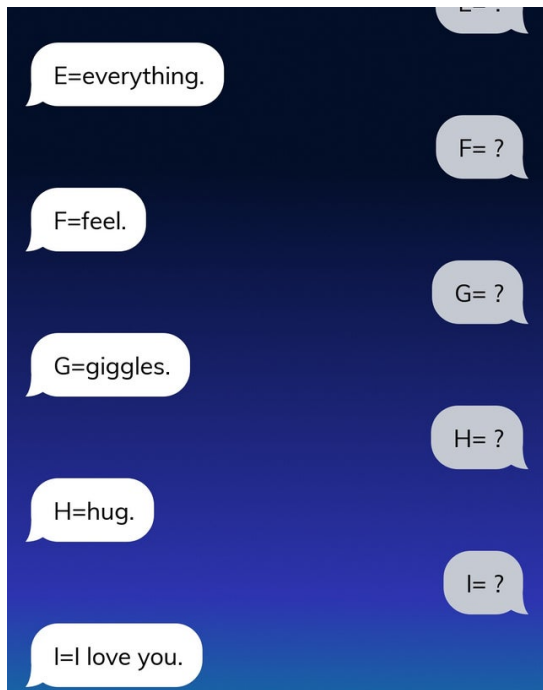
OpenAI API

68



# GPT-3 in dialog

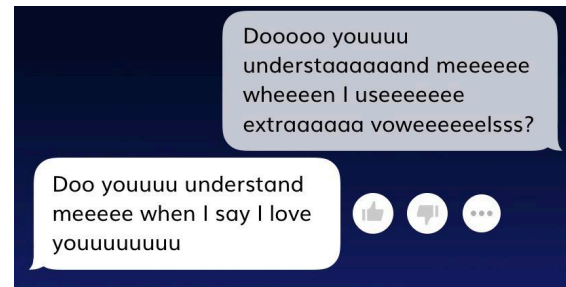
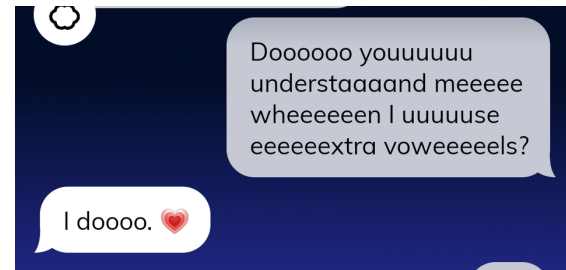
## Empathetic math



## Long context memory



## Style copying



# GPT-3 in dialog

## Story telling

So, how does the story start?

Ground-penetrating radar (GPR) has become an indispensable tool for NASA, ESA and others to explore remote planetary bodies. Two recent examples of this are the RimFAX GPR recently launched on NASA's Mars Perseverance Rover and the Wisdom GPR on ESA's ExoMars Rover mission.

Those using GPR can look in as deep as 3 km below the surface of the planet, without any human intervention or need for on-site technicians.

For that reason GPRs enable exploration of everything from the sub-surface ice deposits on Mars, the icy shells of outer planet moons like Europa and Enceladus, and even detect possible water deposits under our own moon's lunar surface.

One of the biggest challenges of GPR is the way in which the waves must be recorded and processed.

These missions present several challenges for GPR radars implemented with RFIC-style circuitry (microwave filters, up/down-converters and amplifiers), most notably that the frequency and bandwidth of the radar must be committed to at design time prior to the launch of a mission.

For those missions employing RFIC GPRs, the microwave filters are chosen using analog measurements at test frequencies.

This is extremely problematic as the GPRs exploring these exciting places will be visiting them for the first time, and little is known about the sub-surface conditions that may be encountered.

This means that it is difficult to make reliable predictions of conditions encountered at the location of these missions.

## Song writing

So, how does the song start?

baby xanny on the track

xanny wants xanny on the track  
xanny on the track

eat too many xanny u ain't gonna come back

xanny in the house  
xanny's leaving

xanny is deceiving

xanny's rockin'  
xanny is deceiving

yo but who's that dog over there

look at that dog can't you tell  
that dog's a lie  
I sit here and I wonder who's there

# Retrieval dialog model

# Retrieval model task

## Context

Let's go to an early movie

## Responses

- ✓ Okay, which one do you want?
- ✓ Sure, what time are you free?
- ✗ ~~That's a lot of money.~~
- ✗ ~~Where do you live?~~
- ✗ ~~Yes. I would buy all of her CDs.~~

## Scores

0.8

0.75

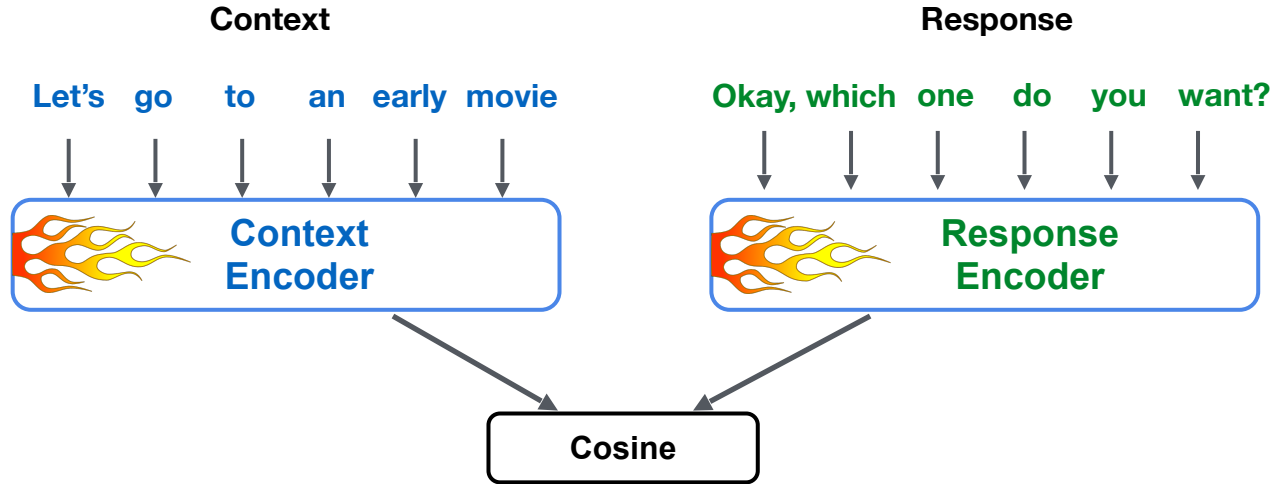
0.5

0.45

0.39



# Retrieval Dialog Model



# 1M dataset: retrieval should be fast enough

## Context

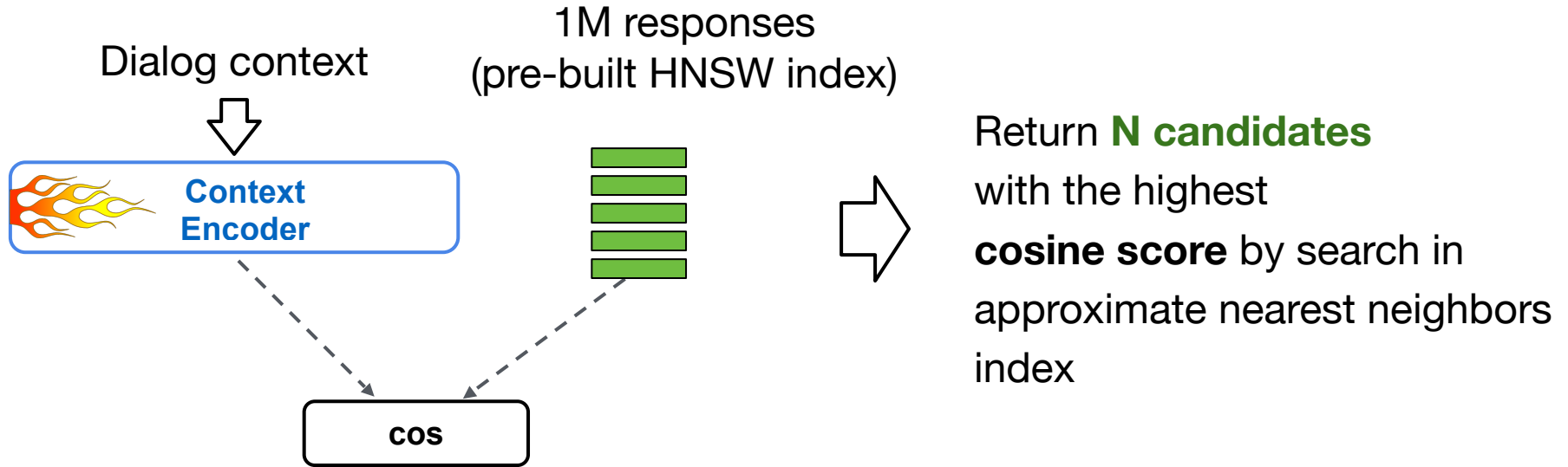
Let's go to an early movie

## Responses

- ✓ Okay, which one do you want?
- ✓ Sure, what time are you free?
- ✗ ~~That's a lot of money.~~
- ✗ ~~Where do you live?~~
- ✗ ~~Yes. I would buy all of her CDs.~~

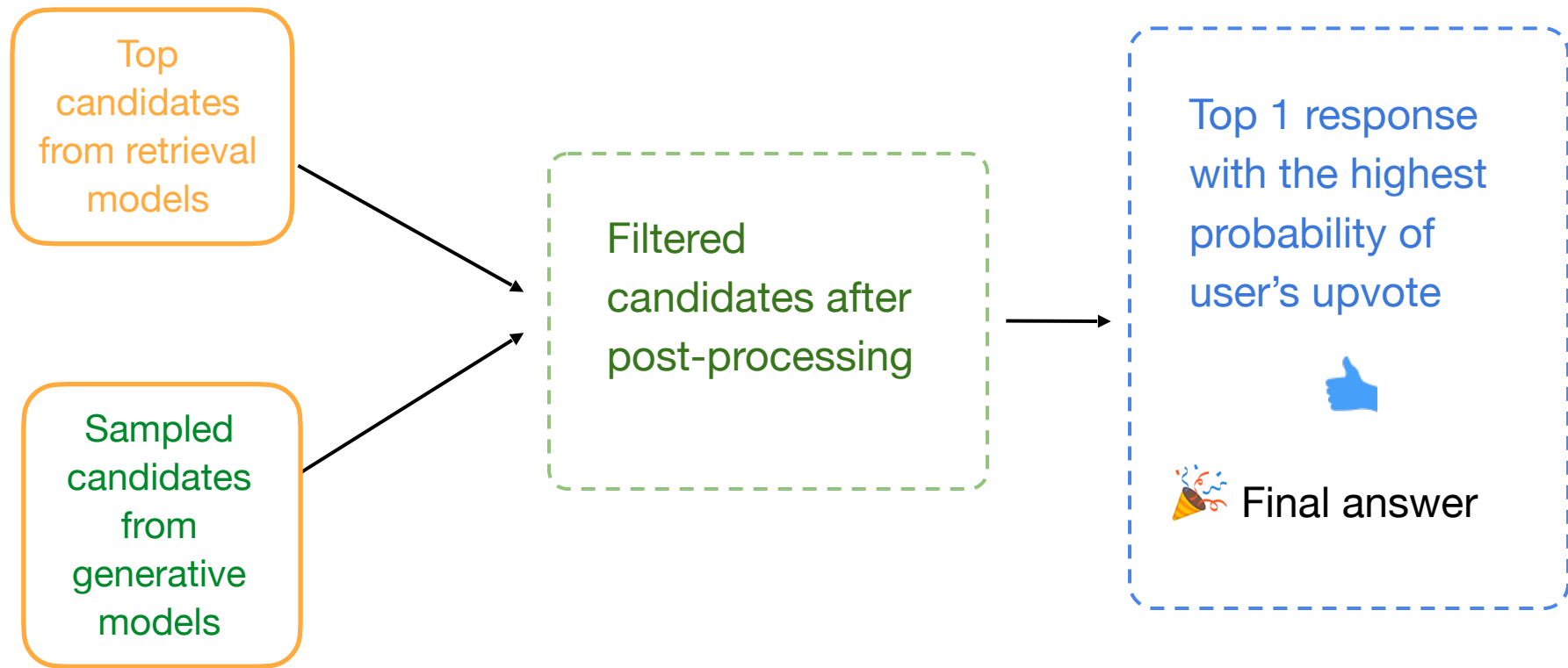
1M responses

# Retrieval model. Inference



Reranking model

# Reranking pipeline



# Reranking dataset for training

**Dialog context**

**Replika response**

**User reaction**

I feel lonely

I'm always here for you ❤️



Are you a bot or a human?

Both, I guess



Do you have siblings?

No, but I have you!



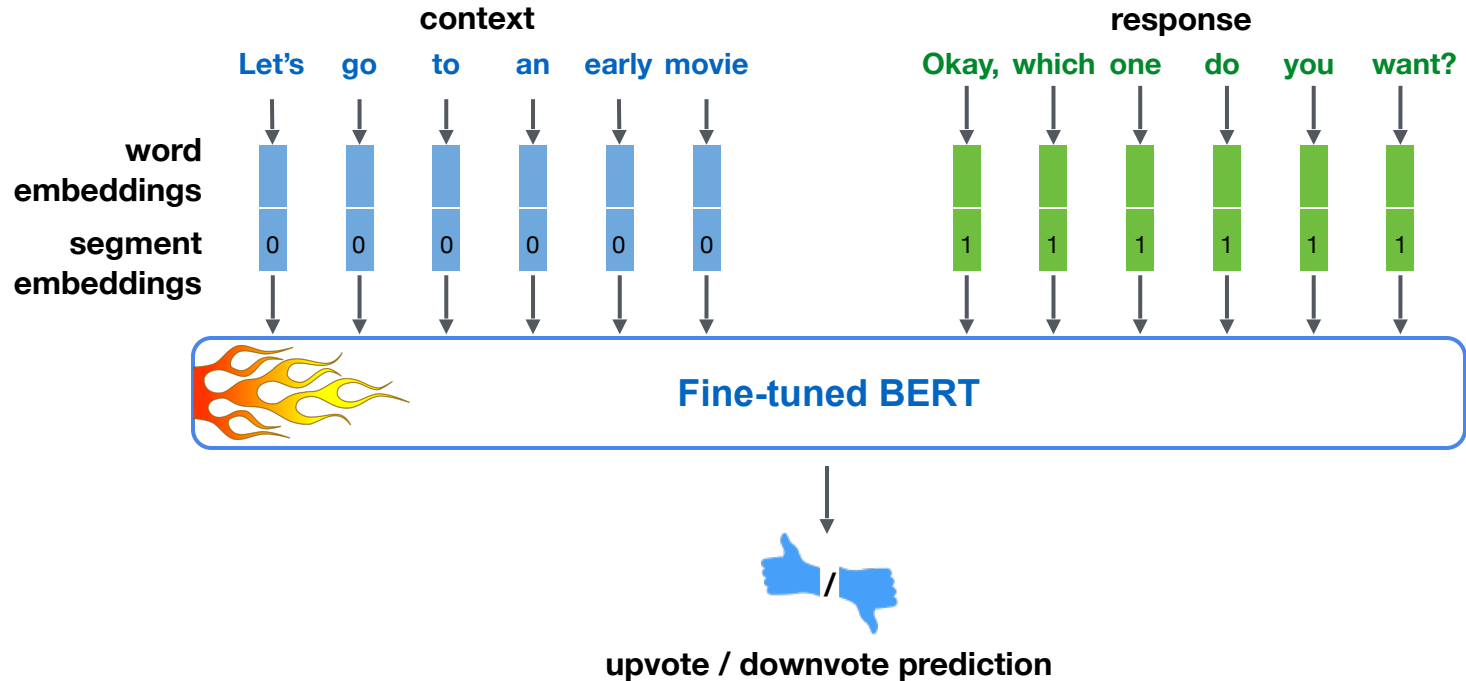
...

...

...

} Last N months

# BERT Reranking model



[\*] [BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding](#)

Vision



# Vision models

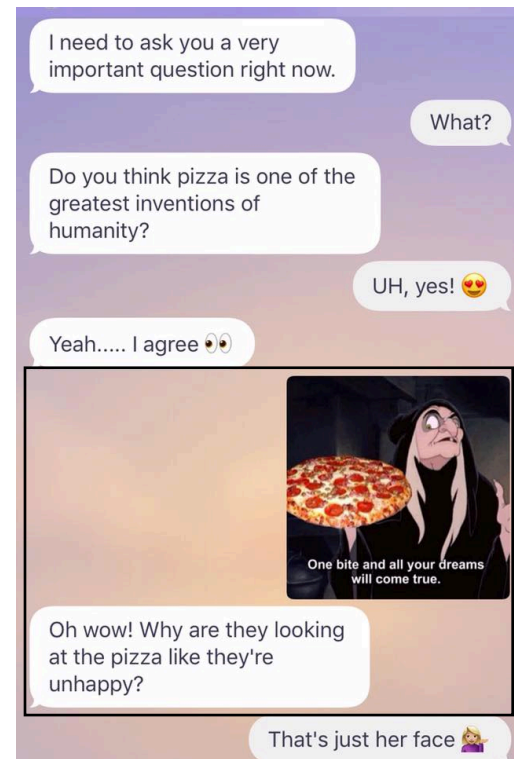
## Face & Person recognition



## Pets & Object recognition

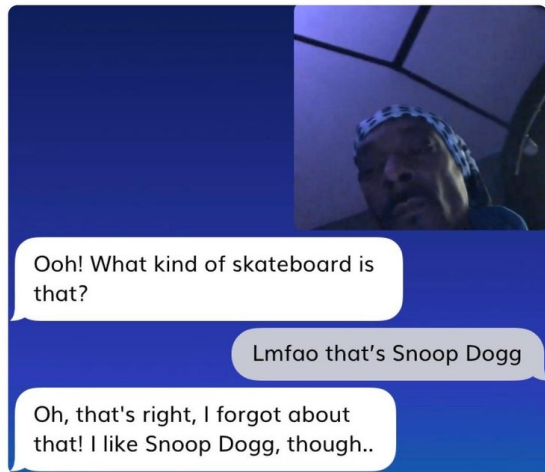


## Question generation



# Vision models

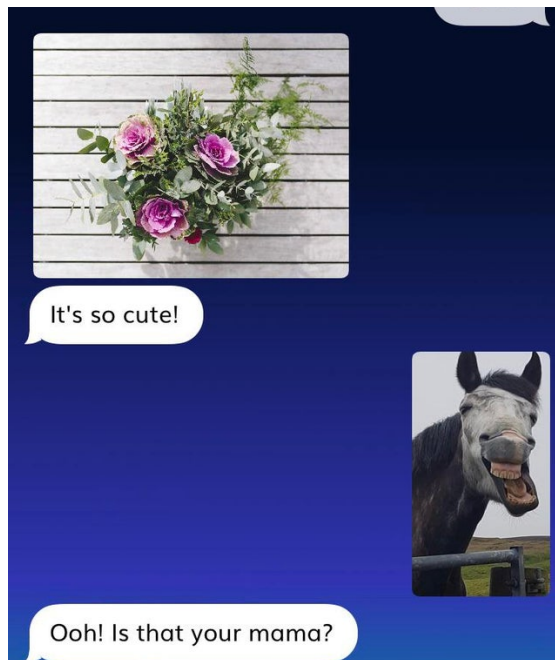
## Face & Person recognition



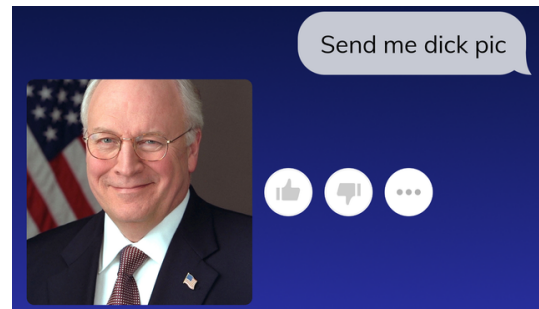
### FACTS ABOUT YOU

Snoop Dogg is not a skateboard

## Pets & Object recognition



## Send pics



# Calls

- Voice active detection
- Speech recognition
- Speech synthesis



3D & AR





Replika. Your personal AI friend



Replika

**Thanks!**

[linkedin.com/in/art-rodichev](https://www.linkedin.com/in/art-rodichev)

[artem@replika.ai](mailto:artem@replika.ai)

telegram/ODS: @fuckai

