The Extract/Validate/Extrapolate Loop

A pattern for LLM Application Development

@kbal11
kevin@mento.co

Very quick about me

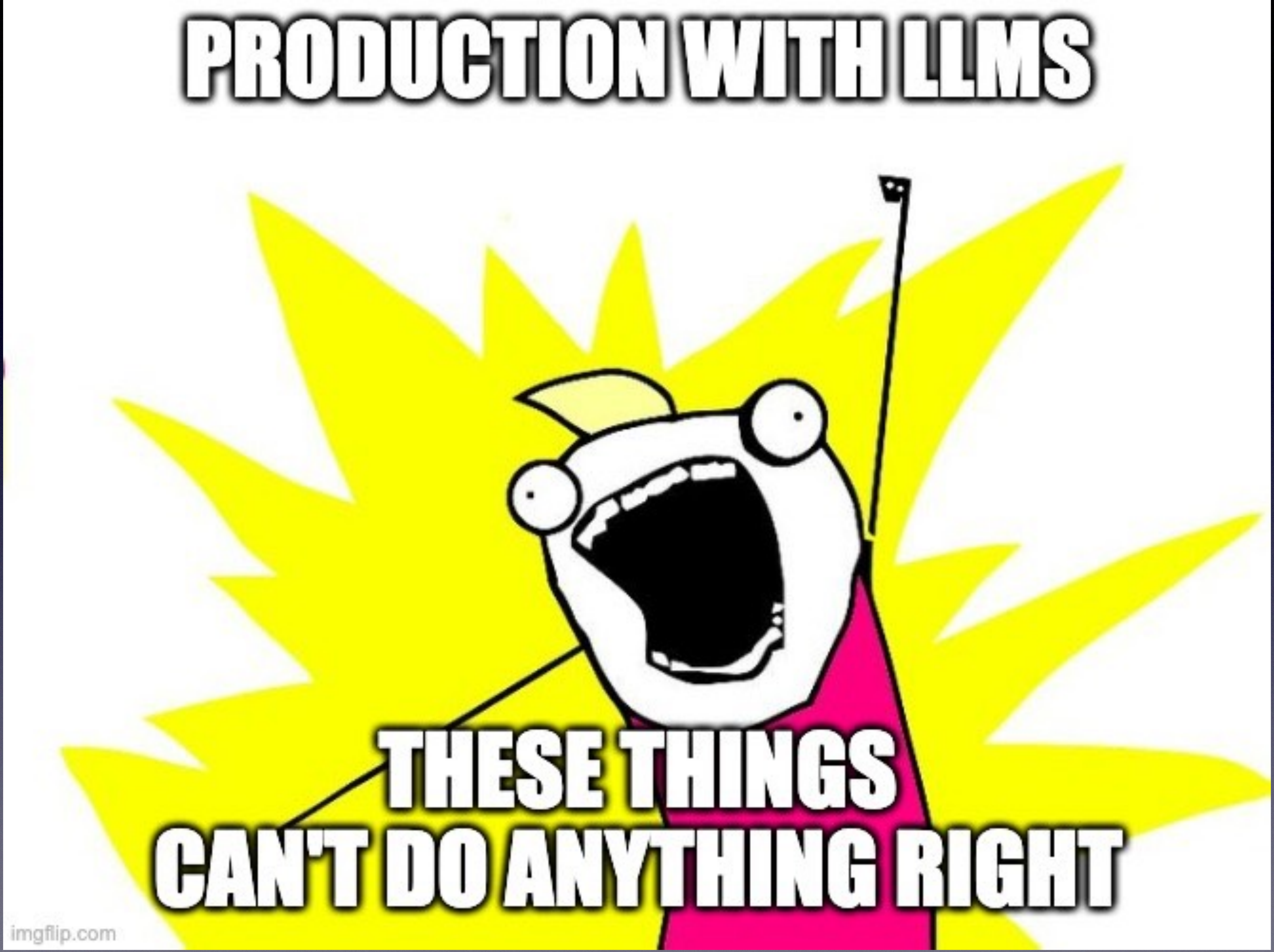
- 20 years in tech as Eng & Eng Manager
- ~1.5 years in GenAI world, helped start Latent Space paper club & AI in Action discussion group
- Currently VP Engineering at Mento

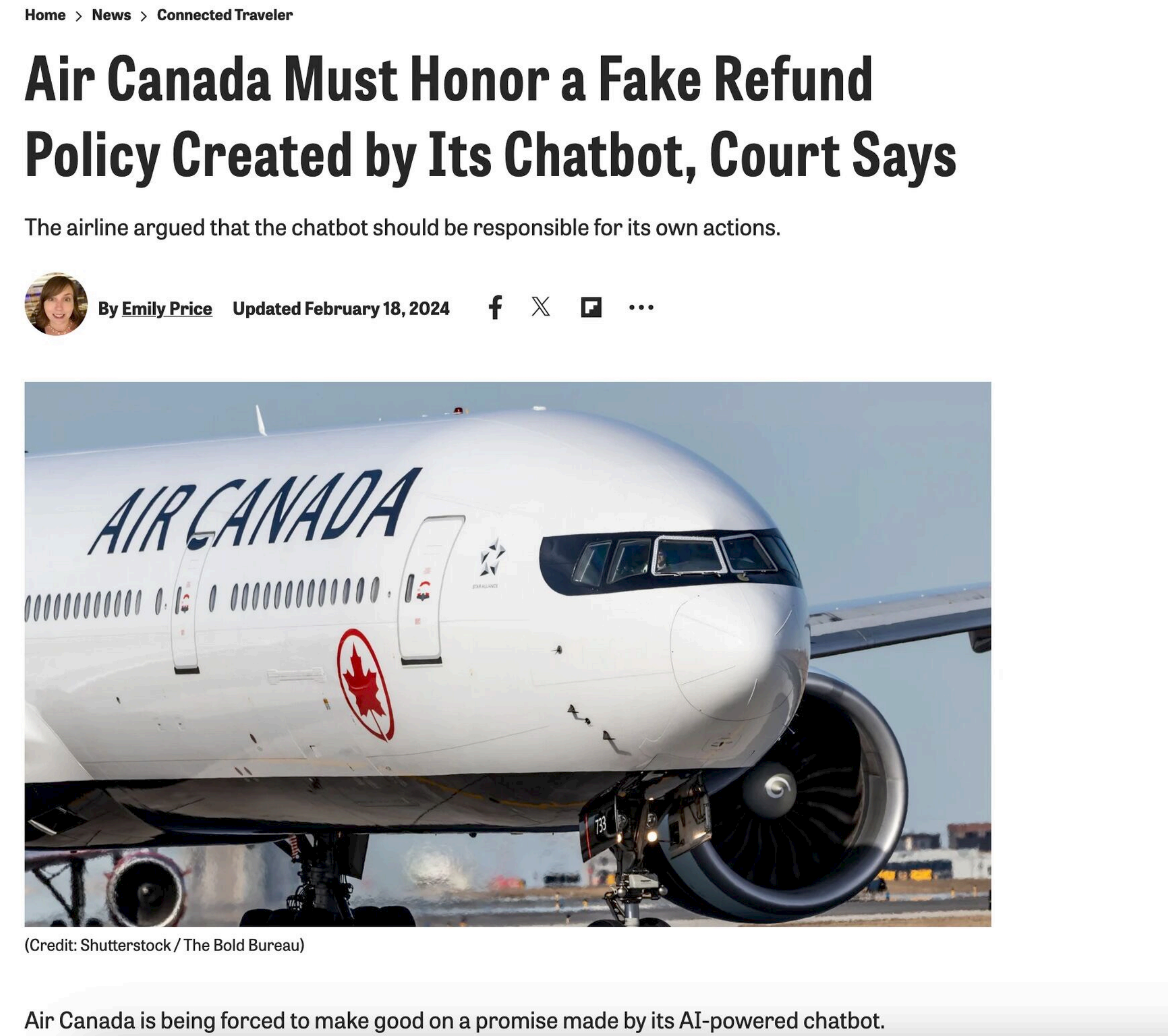
Motivating Problem:
Productionizing LLM Applications

DEMOS WITH LLMS
THESE THINGS CAN DO ANYTHING!
imgflip.com

Fundamental LLM Strengths

- Adaptable
- Flexible
- Great with unstructured data & interactions
- Good at executing/generating based on instruction


PRODUCTION WITH LLMS
THESE THINGS CAN'T DO ANYTHING RIGHT
imgflip.com



Home > News > Connected Traveler

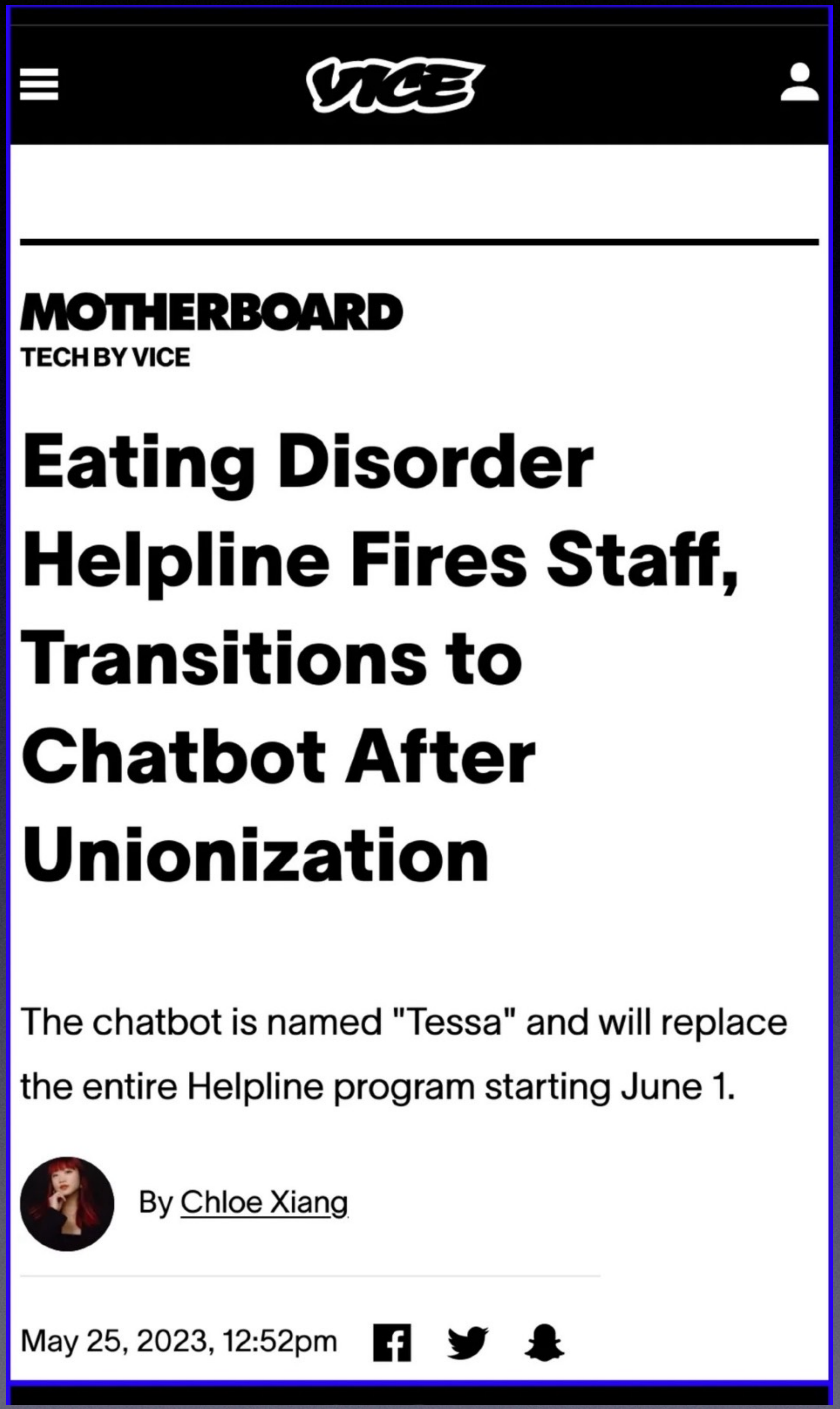
Air Canada Must Honor a Fake Refund Policy Created by Its Chatbot, Court Says

The airline argued that the chatbot should be responsible for its own actions.

By **Emily Price** Updated February 18, 2024

(Credit: Shutterstock / The Bold Bureau)

Air Canada is being forced to make good on a promise made by its AI-powered chatbot.



MOTHERBOARD
TECH BY VICE

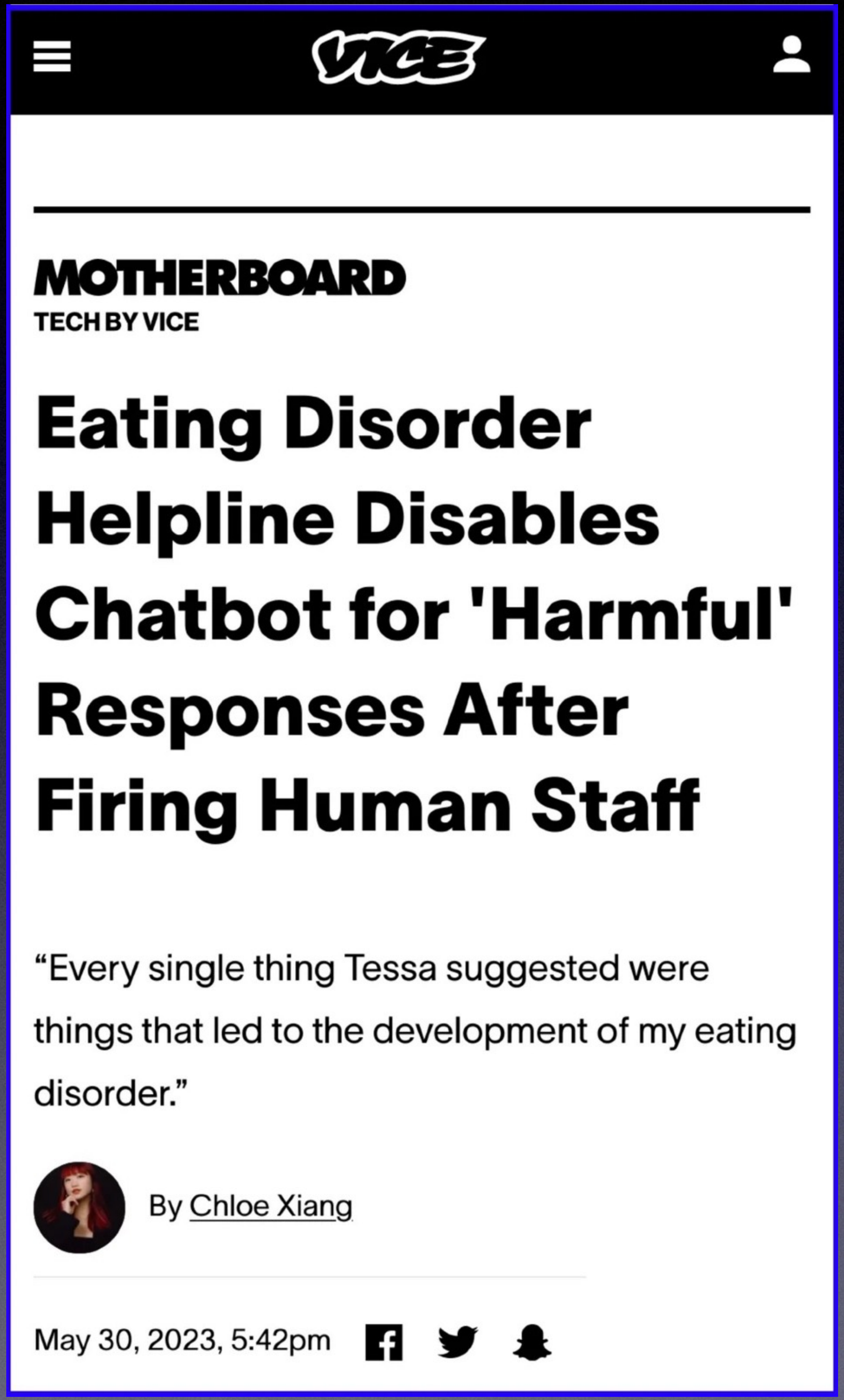
Eating Disorder Helpline Fires Staff, Transitions to Chatbot After Unionization

The chatbot is named "Tessa" and will replace the entire Helpline program starting June 1.

By Chloe Xiang

May 25, 2023, 12:52pm

5 days later



MOTHERBOARD
TECH BY VICE

Eating Disorder Helpline Disables Chatbot for 'Harmful' Responses After Firing Human Staff

“Every single thing Tessa suggested were things that led to the development of my eating disorder.”

By Chloe Xiang

May 30, 2023, 5:42pm

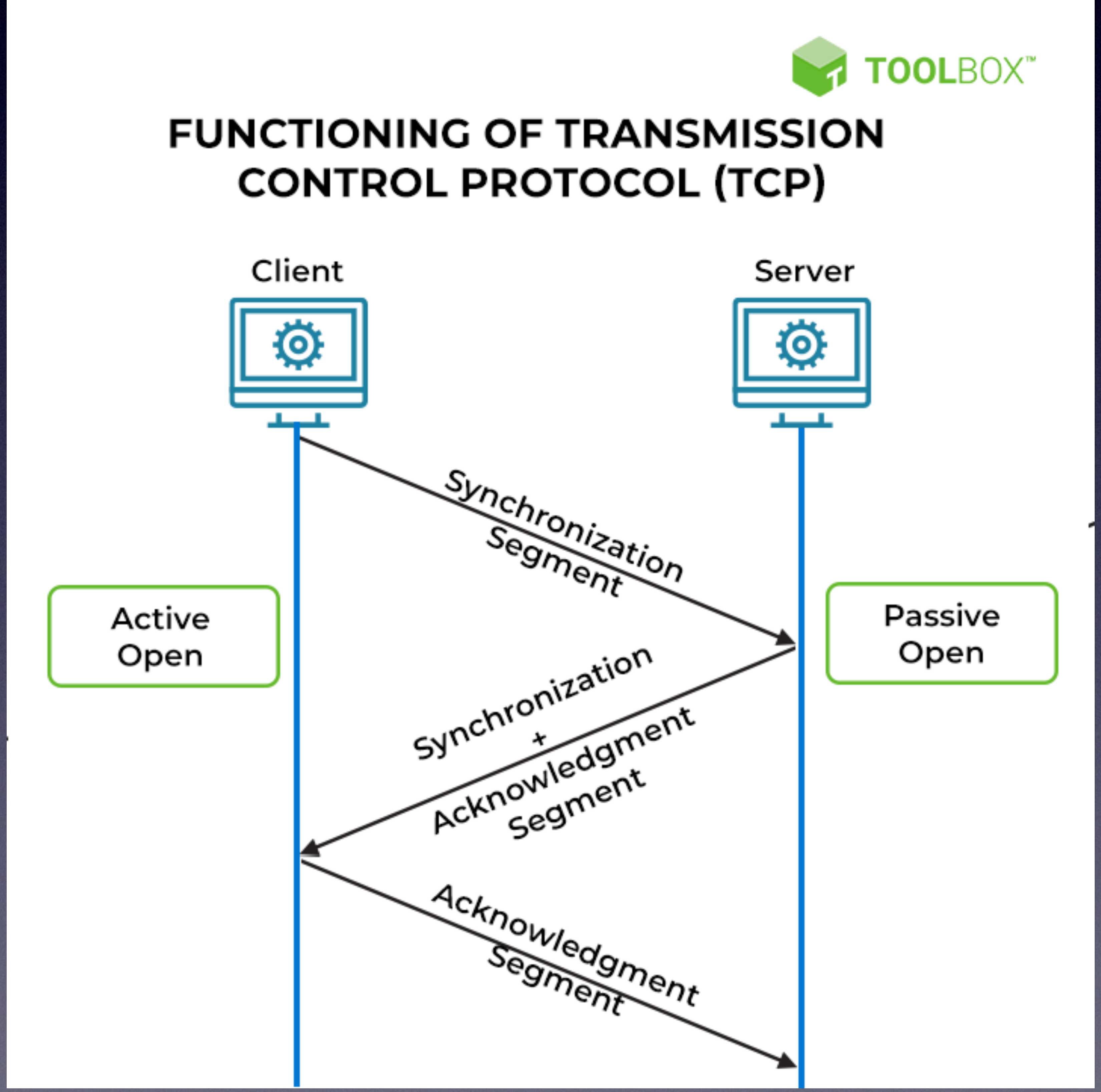
Fundamental Challenges of LLMs

- Hallucination
- Black box
- Bad at “decisions” and judgment

How do we make a reliable service on top of an unreliable substrate?

Flashback: TCP

Reliable protocol built on top of an unreliable protocol

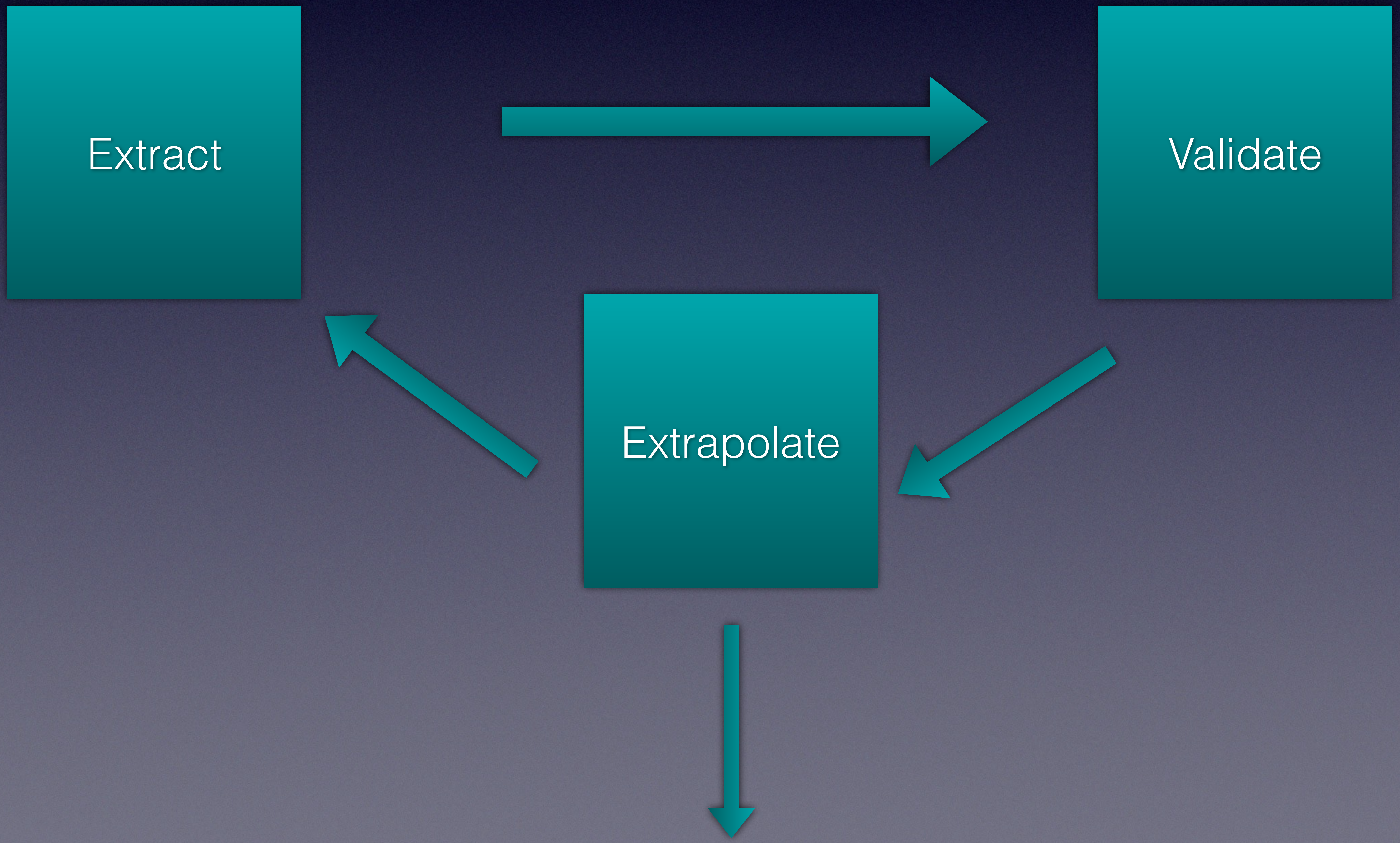

TOOLBOX™
FUNCTIONING OF TRANSMISSION CONTROL PROTOCOL (TCP)
Client
Server
Synchronization Segment
Active Open
Passive Open
Synchronization + Acknowledgment Segment
Acknowledgment Segment

Source: https://www.spiceworks.com/tech/networking/articles/tcp-vs-udp/

What does that look like for LLMs?

The Extract/Validate/Extrapolate Loop

A New UX Pattern & Application Building Block



Extract

Validate

Extrapolate

Extract

- Use LLM to extract information from text or user interaction
- Can be interactive or asynchronous
- Can extract structured or unstructured

Examples of Extraction

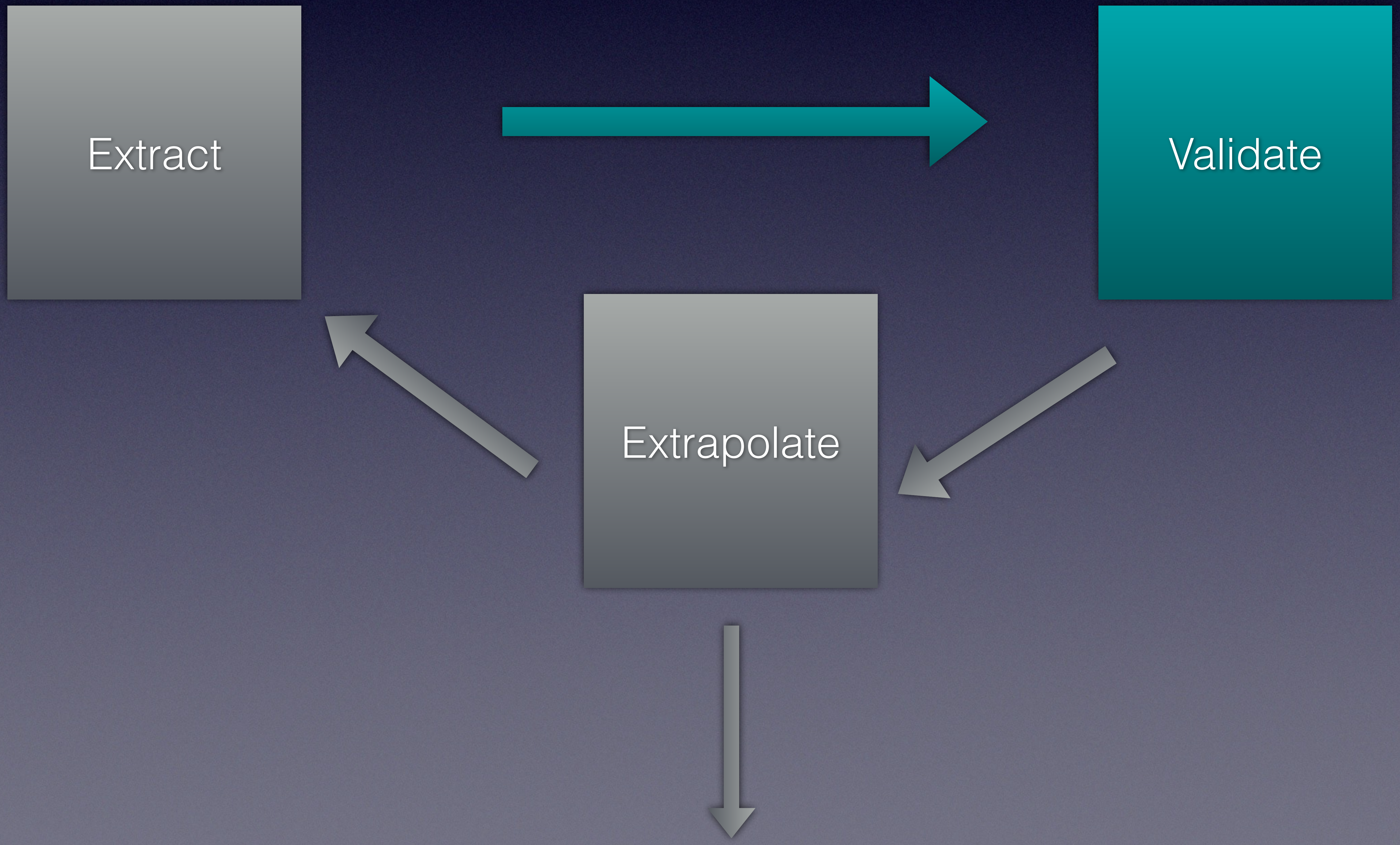
- Summarize a transcript
- Identify a user goal
- Write some code to accomplish a task
- Pull out a CSV of data from an unstructured document

Extract

- Use LLM to extract information from text or user interaction
- Can be interactive or asynchronous
- Can extract structured or unstructured

Most LLM Demos Stop At Extract

The Extract/Validate/Extrapolate Loop



Extract

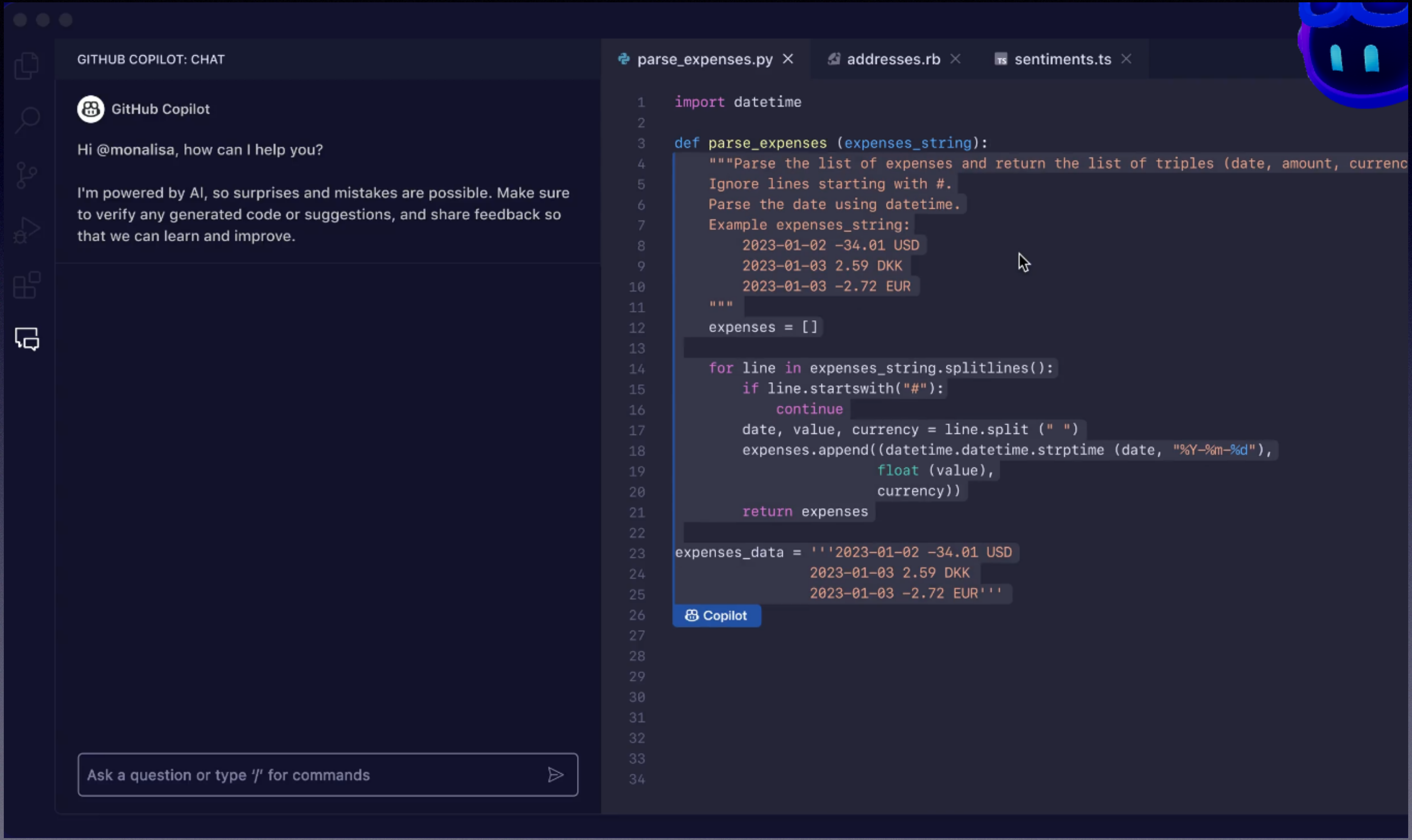
Validate

Extrapolate

Validate

- Confirm validity/correctness of extraction
- Can be human or machine
- Can be formal or informal

Validation Example: Copilot



GitHub Copilot

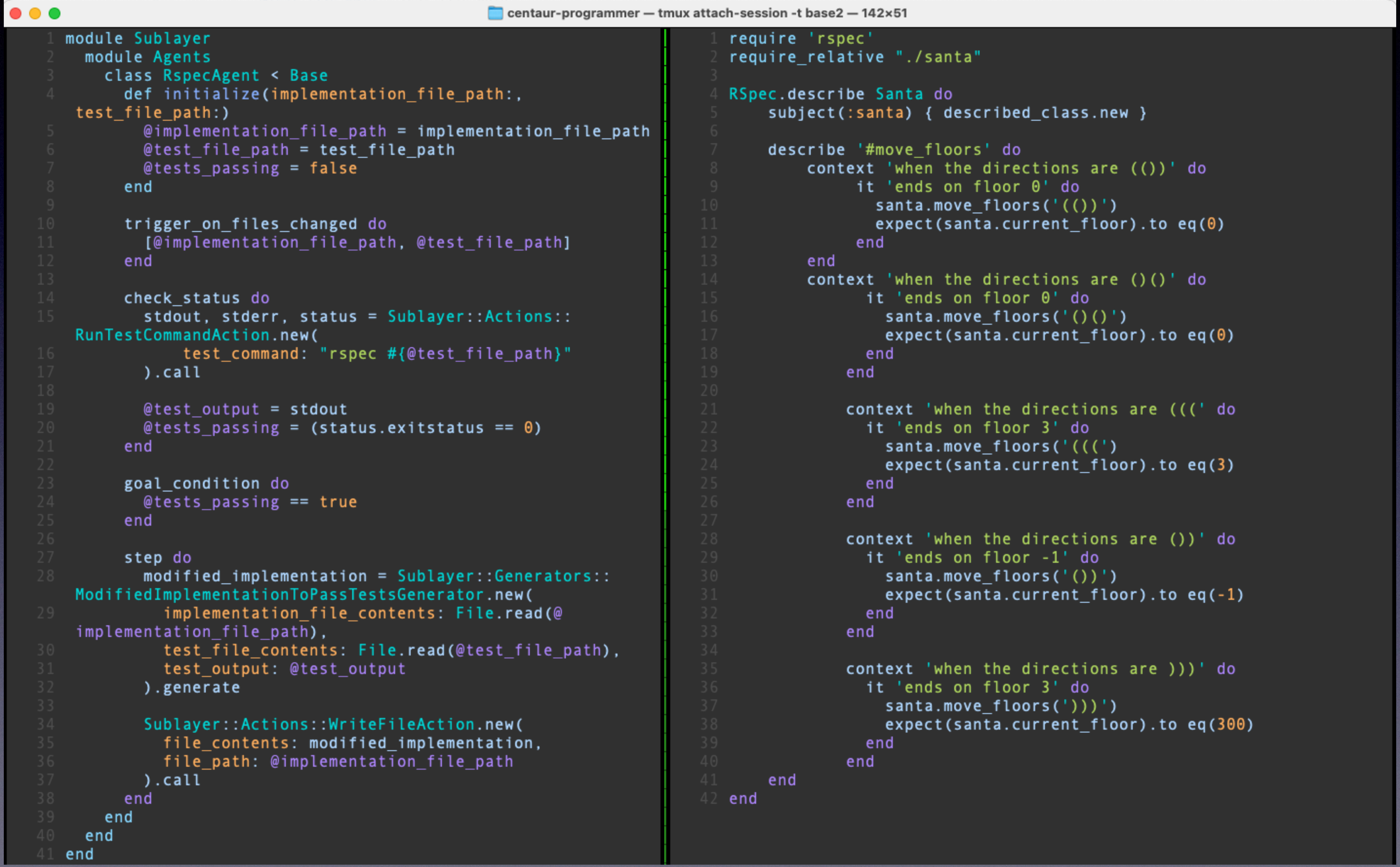
Hi @monalisa, how can I help you?

I'm powered by AI, so surprises and mistakes are possible. Make sure to verify any generated code or suggestions, and share feedback so that we can learn and improve.

```
import datetime

def parse_expenses (expenses_string):
    """Parse the list of expenses and return the list of triples (date, amount, currenc
    Ignore lines starting with #.
    Parse the date using datetime.
    Example expenses_string:
        2023-01-02 -34.01 USD
        2023-01-03 2.59 DKK
        2023-01-03 -2.72 EUR
    """
    expenses = []

    for line in expenses_string.splitlines():
        if line.startswith("#"):
            continue
        date, value, currency = line.split (" ")
        expenses.append((datetime.datetime.strptime (date, "%Y-%m-%d"),
                        float (value),
                        currency))
        return expenses

expenses_data = '''2023-01-02 -34.01 USD
                2023-01-03 2.59 DKK
                2023-01-03 -2.72 EUR'''
```

Source: https://github.com/features/copilot

Validation Example: Sublayer

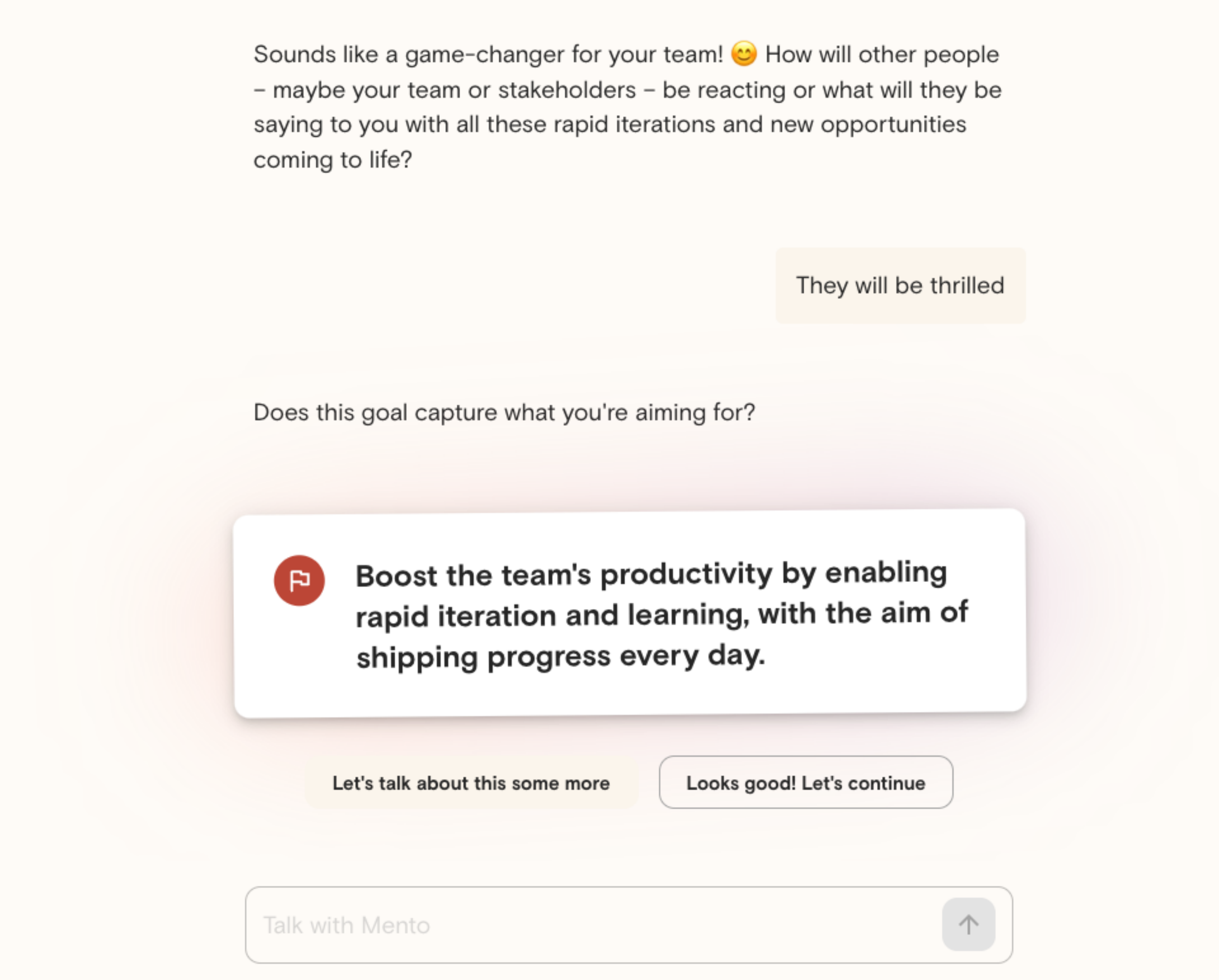


```ruby
module Sublayer
  module Agents
    class RspecAgent < Base
      def initialize(implementation_file_path:, test_file_path:)
        @implementation_file_path = implementation_file_path
        @test_file_path = test_file_path
        @tests_passing = false
      end

      trigger_on_files_changed do
        [@implementation_file_path, @test_file_path]
      end

      check_status do
        stdout, stderr, status = Sublayer::Actions::RunTestCommandAction.new(
            test_command: "rspec #{@test_file_path}"
        ).call

        @test_output = stdout
        @tests_passing = (status.exitstatus == 0)
      end

      goal_condition do
        @tests_passing == true
      end

      step do
        modified_implementation = Sublayer::Generators::ModifiedImplementationToPassTestsGenerator.new(
          implementation_file_contents: File.read(@implementation_file_path),
          test_file_contents: File.read(@test_file_path),
          test_output: @test_output
        ).generate

        Sublayer::Actions::WriteFileAction.new(
          file_contents: modified_implementation,
          file_path: @implementation_file_path
        ).call
      end
    end
  end
end
```

```ruby
require 'rspec'
require_relative "./santa"

RSpec.describe Santa do
    subject(:santa) { described_class.new }

    describe '#move_floors' do
        context 'when the directions are (())' do
             it 'ends on floor 0' do
               santa.move_floors('(())')
               expect(santa.current_floor).to eq(0)
             end
        end
        context 'when the directions are ()()' do
              it 'ends on floor 0' do
                santa.move_floors('()()')
                expect(santa.current_floor).to eq(0)
              end
            end

            context 'when the directions are (((' do
              it 'ends on floor 3' do
                santa.move_floors('(((')
                expect(santa.current_floor).to eq(3)
              end
            end

            context 'when the directions are ())' do
              it 'ends on floor -1' do
                santa.move_floors('())')
                expect(santa.current_floor).to eq(-1)
              end
            end

            context 'when the directions are )))' do
              it 'ends on floor 3' do
                santa.move_floors(')))')
                expect(santa.current_floor).to eq(300)
              end
            end
    end
end
```

Source: https://github.com/sublayerapp/testing_agent/tree/main

Validation Example: Mento



Sounds like a game-changer for your team! 😊 How will other people – maybe your team or stakeholders – be reacting or what will they be saying to you with all these rapid iterations and new opportunities coming to life?

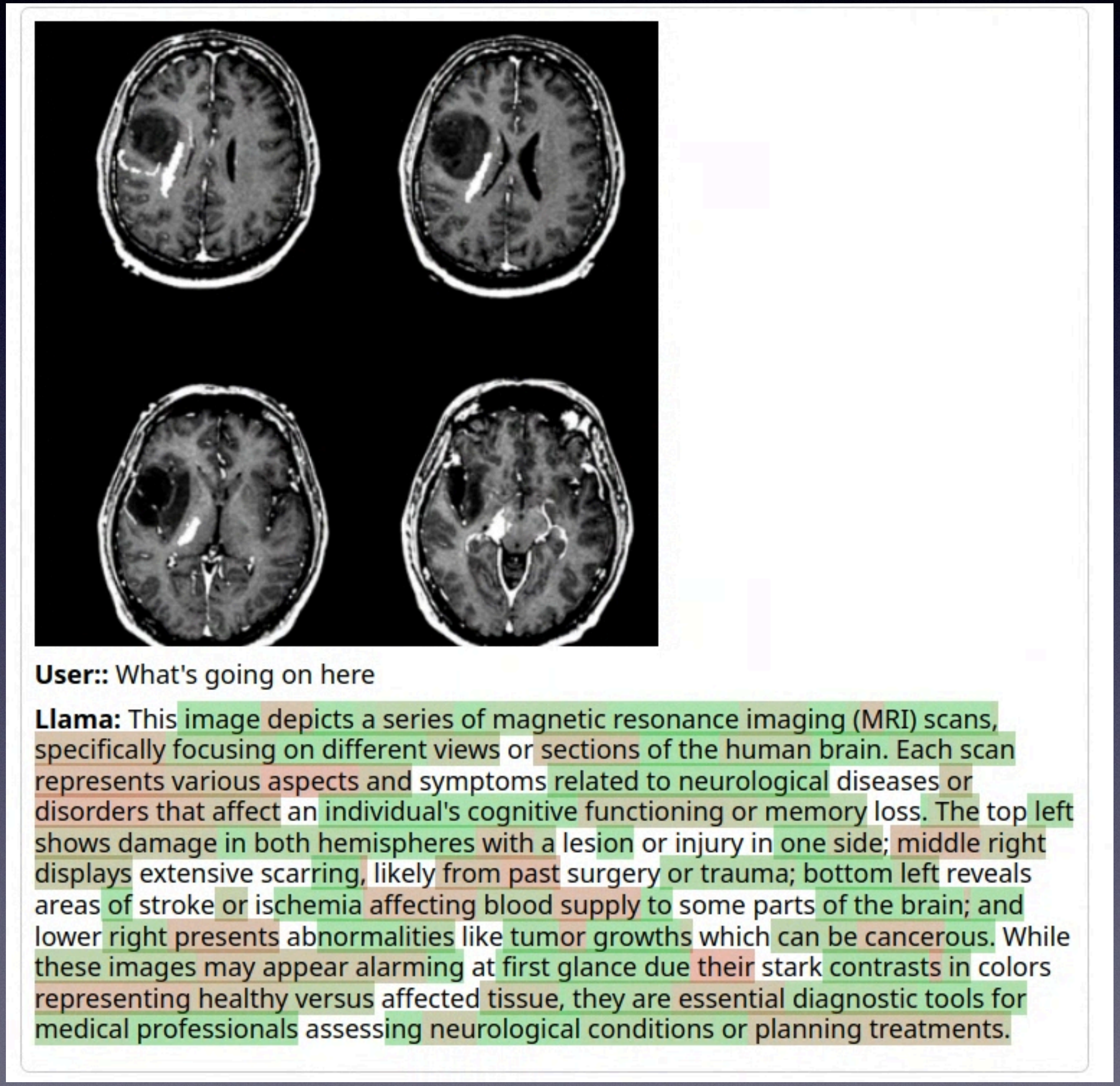
They will be thrilled

Does this goal capture what you're aiming for?

Boost the team's productivity by enabling rapid iteration and learning, with the aim of shipping progress every day.

Source: https://app.mento.co

Validation Example: Exposing Logprobs

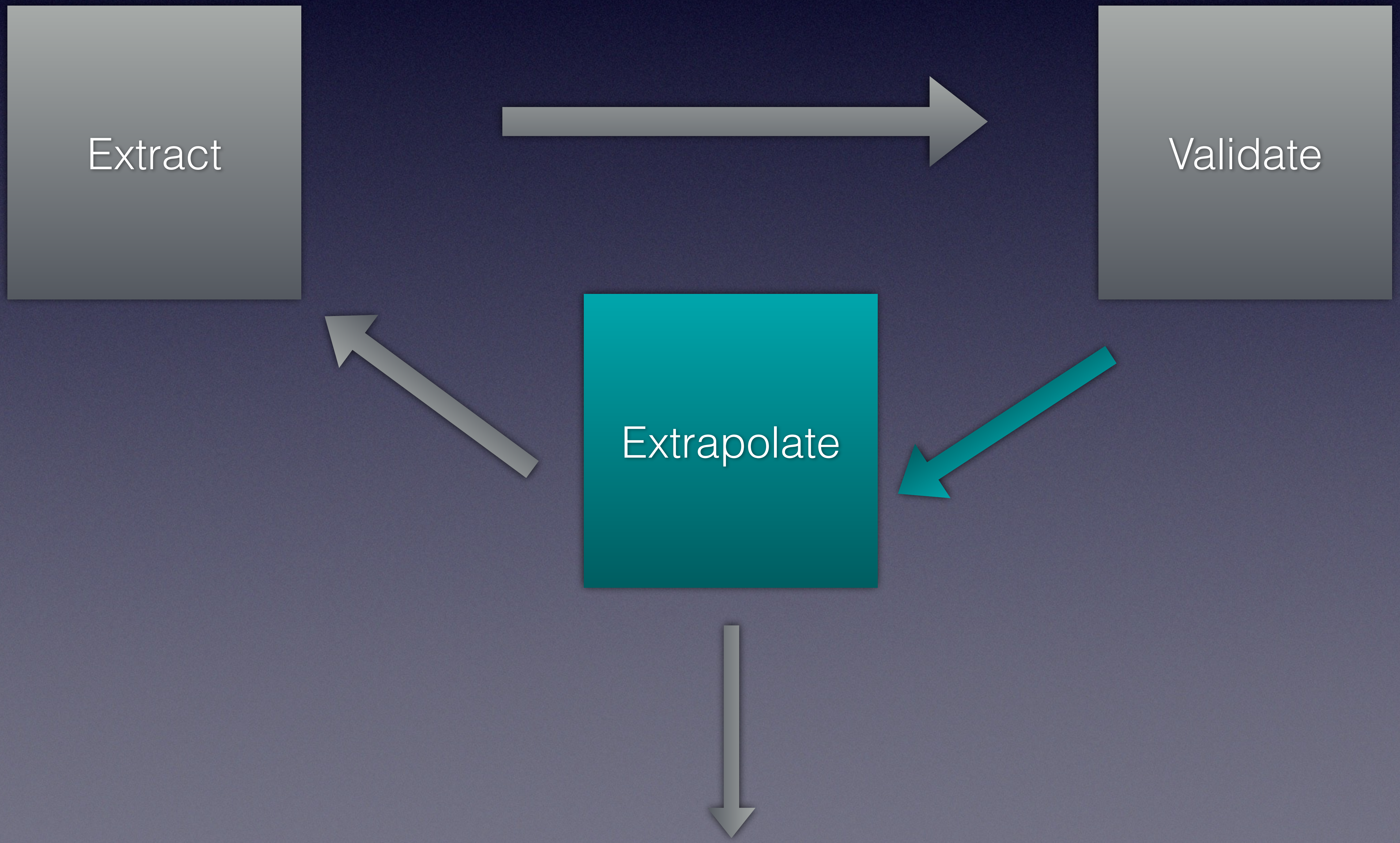


User:: What's going on here

Llama: This image depicts a series of magnetic resonance imaging (MRI) scans, specifically focusing on different views or sections of the human brain. Each scan represents various aspects and symptoms related to neurological diseases or disorders that affect an individual's cognitive functioning or memory loss. The top left shows damage in both hemispheres with a lesion or injury in one side; middle right displays extensive scarring, likely from past surgery or trauma; bottom left reveals areas of stroke or ischemia affecting blood supply to some parts of the brain; and lower right presents abnormalities like tumor growths which can be cancerous. While these images may appear alarming at first glance due their stark contrasts in colors representing healthy versus affected tissue, they are essential diagnostic tools for medical professionals assessing neurological conditions or planning treatments.

Source: https://huggingface.co/nisten/bakllava-14b-2xMoE-alpha-build

The Extract/Validate/Extrapolate Loop



Extract

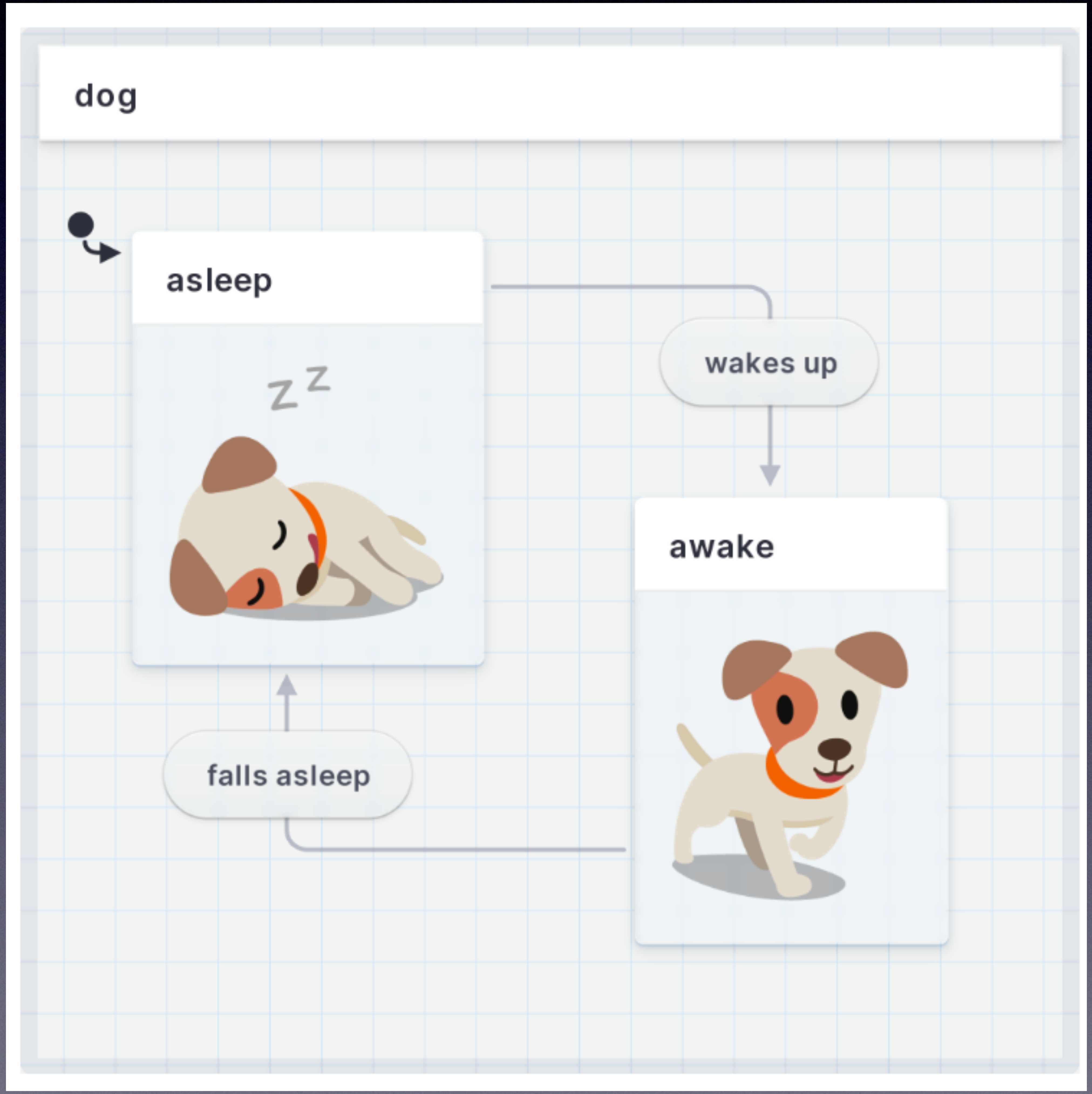
Validate

Extrapolate

Extrapolate

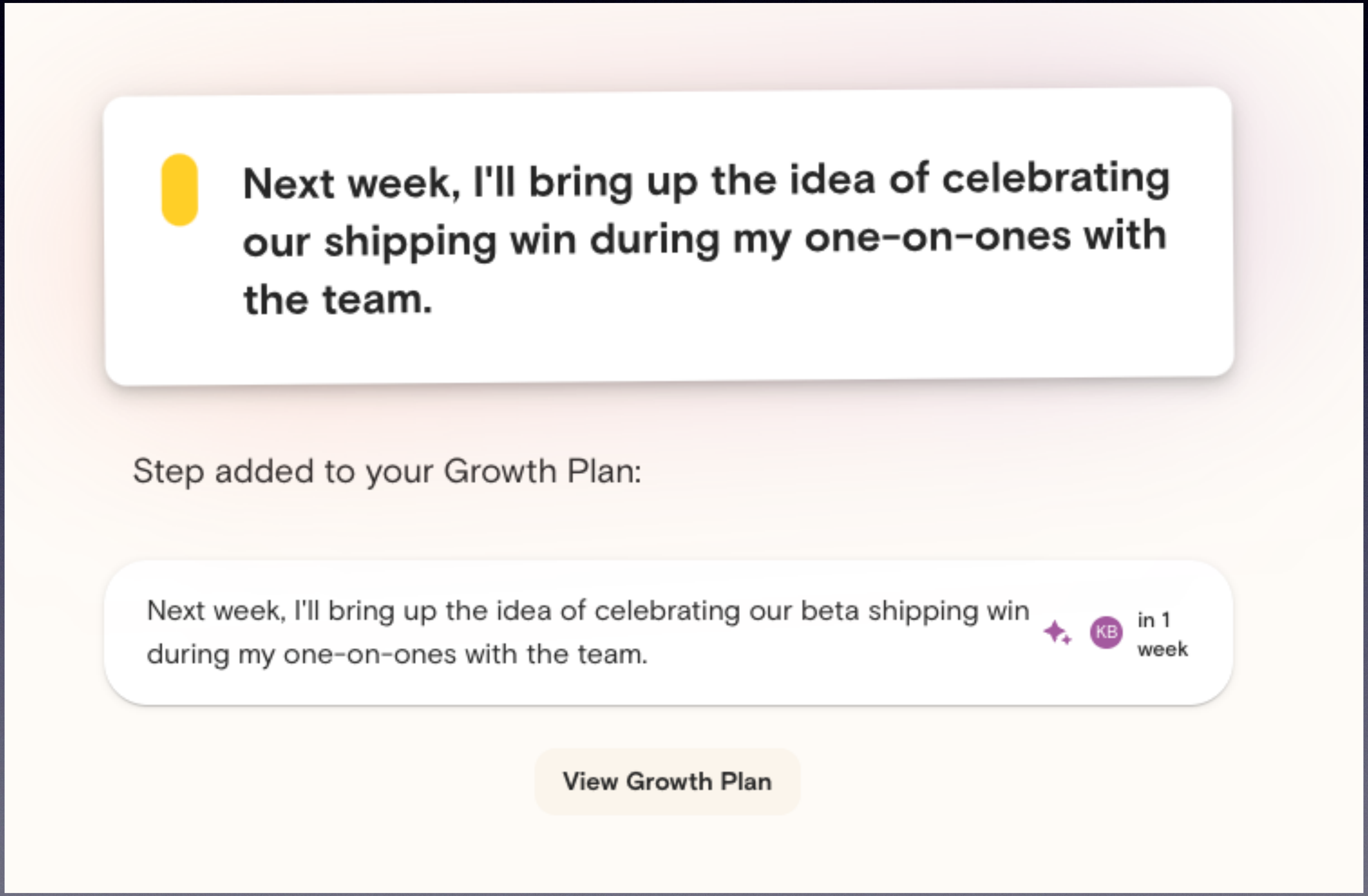
- Take action on confirmed information
- Reason or perform logic about what to do next
- Optionally move to a new Extract phase

Extrapolation Example: State Machines



Source: https://stately.ai/docs/state-machines-and-statecharts

Extrapolation Example: Workflows



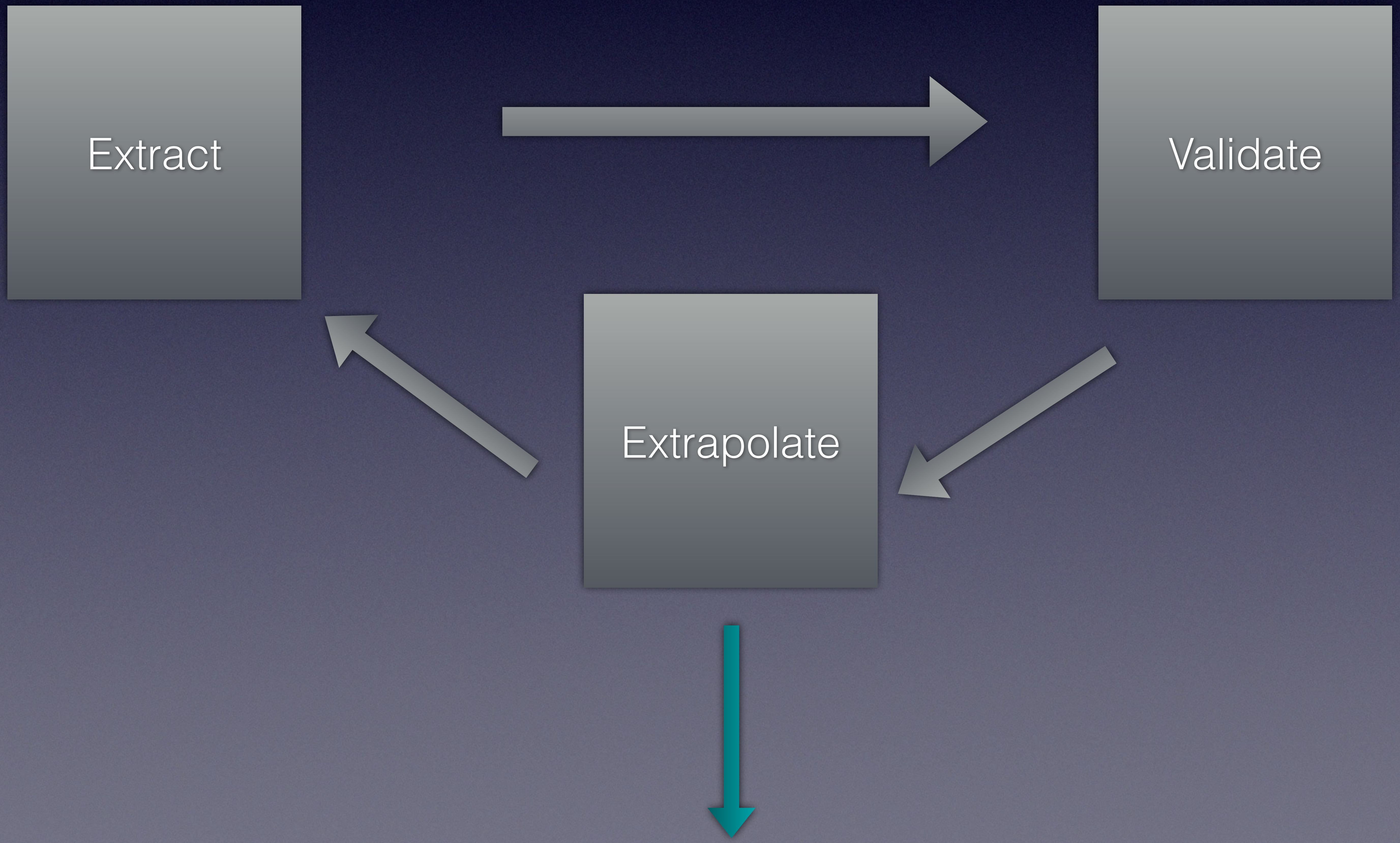
Next week, I'll bring up the idea of celebrating our shipping win during my one-on-ones with the team.

Step added to your Growth Plan:

Next week, I'll bring up the idea of celebrating our beta shipping win during my one-on-ones with the team.

Source: https://app.mento.co

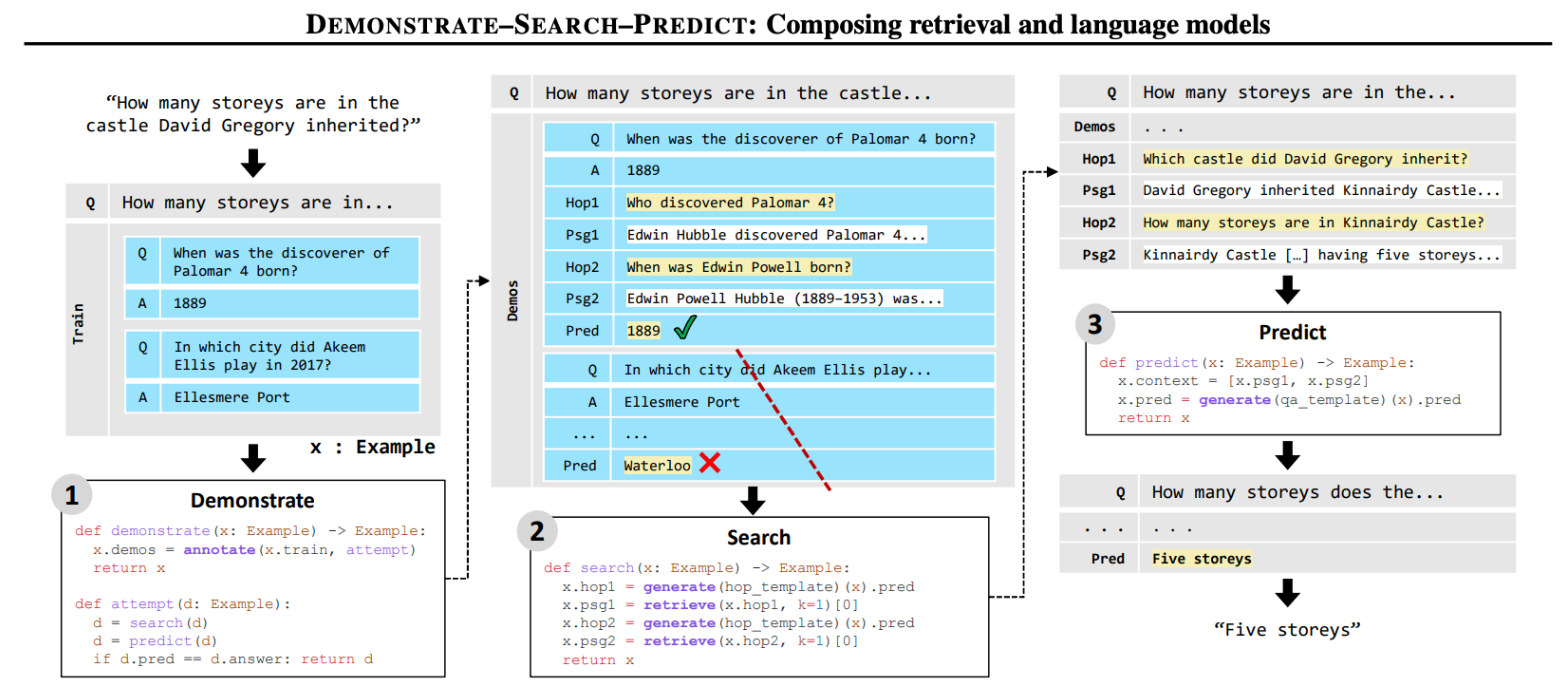
The Extract/Validate/Extrapolate Loop


Extract

Validate

Extrapolate

Examples of the full loop: DSPy



DEMONSTRATE–SEARCH–PREDICT: Composing retrieval and language models

"How many storeys are in the castle David Gregory inherited?"

Q	How many storeys are in...

Train

Q	When was the discoverer of Palomar 4 born?
A	1889
Q	In which city did Akeem Ellis play in 2017?
A	Ellesmere Port

x : Example

1 **Demonstrate**

```
def demonstrate(x: Example) -> Example:
  x.demos = annotate(x.train, attempt)
  return x

def attempt(d: Example):
  d = search(d)
  d = predict(d)
  if d.pred == d.answer: return d
```

Q	How many storeys are in the castle...

Demos

Q	When was the discoverer of Palomar 4 born?
A	1889
Hop1	Who discovered Palomar 4?
Psg1	Edwin Hubble discovered Palomar 4...
Hop2	When was Edwin Powell born?
Psg2	Edwin Powell Hubble (1889–1953) was...
Pred	1889 ✓
Q	In which city did Akeem Ellis play...
A	Ellesmere Port
...	...
Pred	Waterloo ✗

2 **Search**

```
def search(x: Example) -> Example:
  x.hop1 = generate(hop_template)(x).pred
  x.psg1 = retrieve(x.hop1, k=1)[0]
  x.hop2 = generate(hop_template)(x).pred
  x.psg2 = retrieve(x.hop2, k=1)[0]
  return x
```

Q	How many storeys are in the...
Demos	. . .
Hop1	Which castle did David Gregory inherit?
Psg1	David Gregory inherited Kinnairdy Castle...
Hop2	How many storeys are in Kinnairdy Castle?
Psg2	Kinnairdy Castle [...] having five storeys...

3 **Predict**

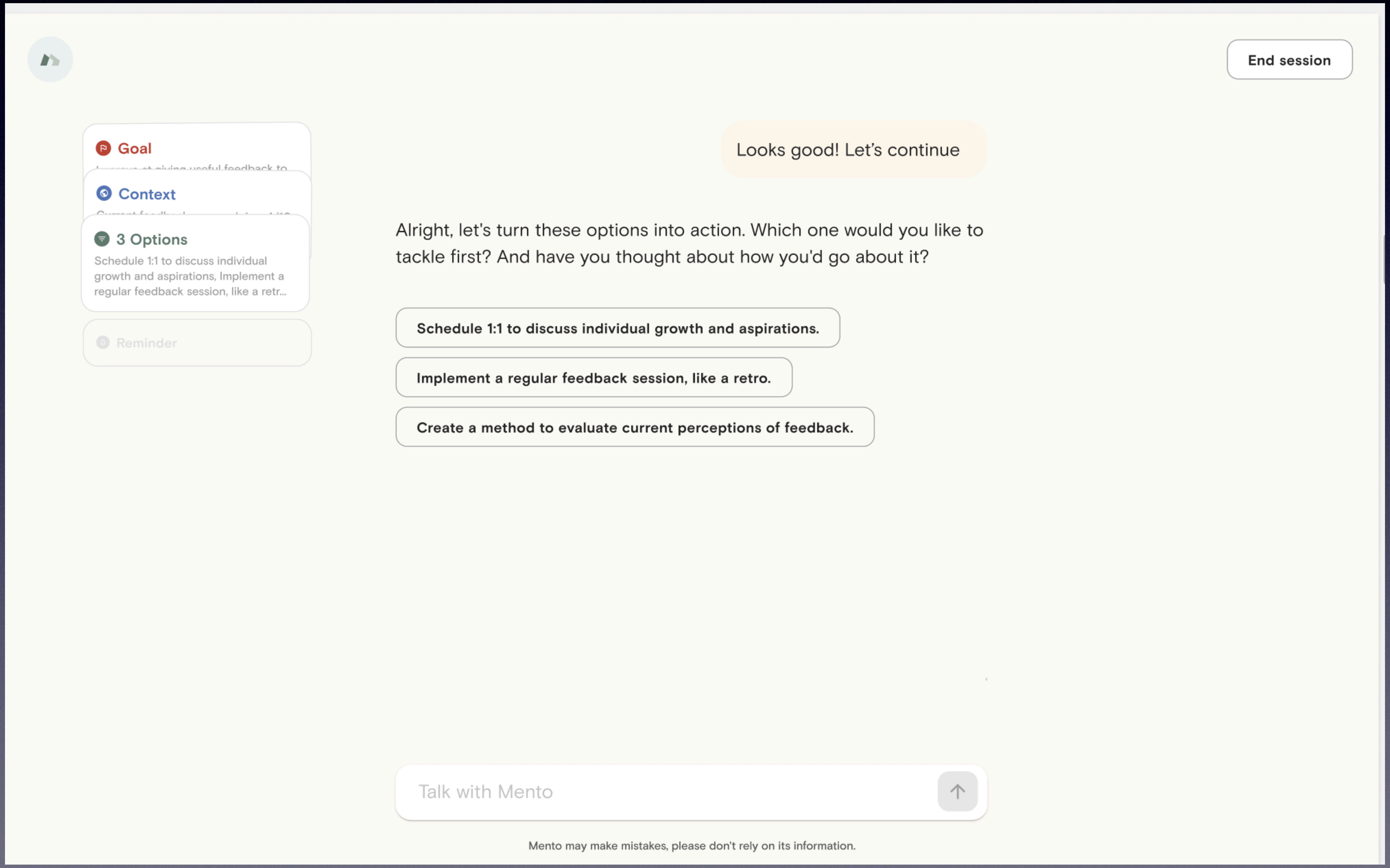
```
def predict(x: Example) -> Example:
  x.context = [x.psg1, x.psg2]
  x.pred = generate(qa_template)(x).pred
  return x
```

Q	How many storeys does the...
. . .	. . .
Pred	Five storeys

"Five storeys"

Source: https://arxiv.org/pdf/2212.14024

Examples of the full loop: Mento



Schedule 1:1 to discuss individual growth and aspirations, Implement a regular feedback session, like a retr...

Looks good! Let's continue

Alright, let's turn these options into action. Which one would you like to tackle first? And have you thought about how you'd go about it?

Schedule 1:1 to discuss individual growth and aspirations.

Implement a regular feedback session, like a retro.

Create a method to evaluate current perceptions of feedback.

Mento may make mistakes, please don't rely on its information.

Source: https://app.mento.co

Implications for app development

- Decompose LLM applications into composable blocks
- Look for opportunities to validate
- Elevate decision making to validatable logic or humans

What does this get us?

- Finer grained control of context
- Catch and course correct hallucinations as they happen
- Smaller, more debuggable & eval-friendly prompts

Thank you

@kbal11
kevin@mento.co