

Copilot for R

- > what it does
- > how to use it, and
- > how it works

David Smith (@revodavid)
Principal Cloud Advocate
Microsoft

Talk Notes:
github.com/revodavid/copilot-for-r



GitHub Copilot

Your AI pair programmer

**Uses the context you've provided
and synthesizes code to match**

Convert comments to code

Autofill for repetitive code

Autosuggest tests

Show alternatives

TS sentiment.ts -GO write_sql.go parse_expense

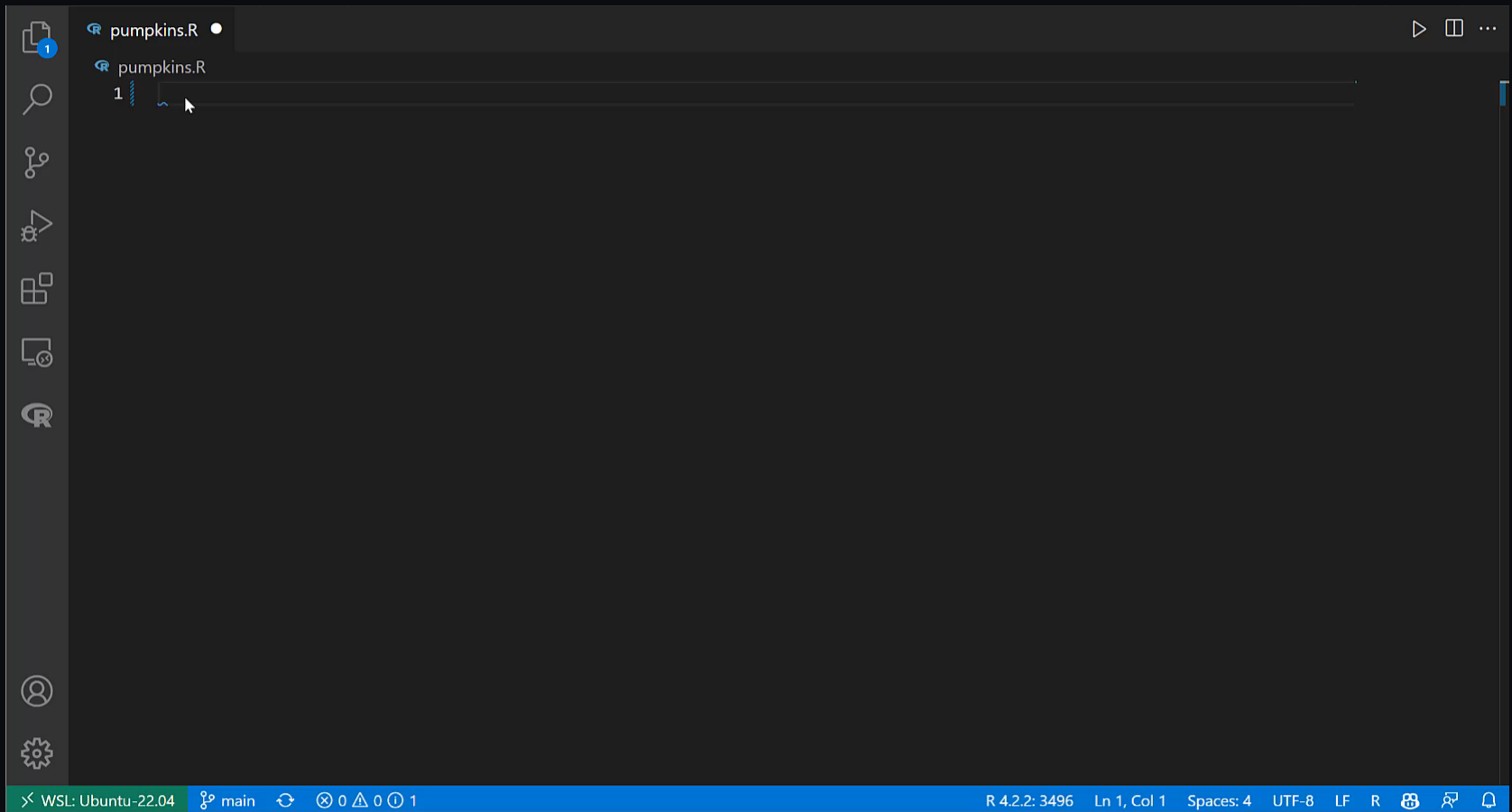
```
1 #!/usr/bin/env ts-node
2
3 import { fetch } from "fetch-h2";
4
5 // Determine whether the sentiment of
6 // Use a web service
7 async function isPositive(text: string) {
8   const response = await fetch(`http://
9     method: "POST",
10    body: `text=${text}`,
11    headers: {
12      "Content-Type": "application/x-w
13    },
14  });
15   const json = await response.json();
16   return json.label === "pos";
17 }
```



aka.ms/get-copilot



GitHub Copilot



How does it work?



Generative AI Models



Generative AI

GPT-3

Prompt:

Write a tagline for an ice cream shop.

Response:

We serve up smiles with every scoop!

Codex

Prompt:

Table customers, columns =
[CustomerId, FirstName,
LastName, Company, Address,
City, State, Country,
PostalCode]

Create a SQL query for all
customers in Texas named Jane
query =

Response:

```
SELECT *  
FROM customers  
WHERE State = 'TX' AND FirstName  
= 'Jane'
```

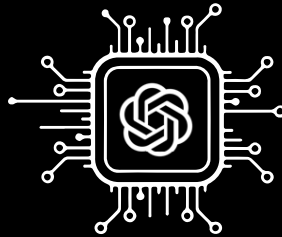
DALL·E

Prompt: A white Siamese cat

Response:



Training Data



Prompt

Model

Completion



Generative AI can:



Generate text, images and code

Different models are trained on different corpuses, depending on the application.



Generate “humanlike” output

What is a likely continuation of the prompt, given the training data?



Extract information

The continuation is likely to be similar to text frequently represented in the training data.



Create novel content

Text, images and code not contained in its training set. Translations. “Creative” works.



Generative AI is not:



Intelligent

It's just a predictive system, designed to give a likely continuation of the prompt given the training data.



Deterministic

Run the same prompt. Get back a different response (probably).



Trustworthy

It can “hallucinate” facts and confidently assert them to be true.

TODO



Generative AI does not:



Learn

The model is fixed at the time of its training.



Contain all of the information of its training set

Think: a [blurry jpeg](#) of its training data.



Include verbatim copies of its training data

But it can generate stuff that looks like it.



Generative AI doesn't:



Understand language

It's just a predictive engine. It doesn't understand math, either.



Understand facts

Many predictions echo information in the training set, but this is **not** guaranteed.



Understand manners, emotion or ethics

Also: avoid anthropomorphizing it.



Understand anything

It's just a prediction engine!

■ Prompt Engineering (very briefly)

Extract the mailing address from this email:

Hi John Doe,

It was great to meet up at Build earlier this week. I thought the AI platform talk was great and I really enjoyed it.

I appreciate the offer for the book. If you are OK, you can mail it to me at home, or 123 Microsoft Way, Bellevue WA 92004.

Regards,
Chris Hoder

Prompt – Text input that provides some context to the engine on what is expecting.

Completion – Output that GPT-3 generates based on the prompt and the trained model.



*Ensure that artificial
general intelligence (AGI)
benefits humanity.*



*Empower every person and
organization on the planet
to achieve more*

GPT-3

Generate and Understand
Text

Codex

Generate and Understand Code

DALL·E

Generate images from text
prompts

OpenAI Codex Model



Public code and text
on the internet

GitHub

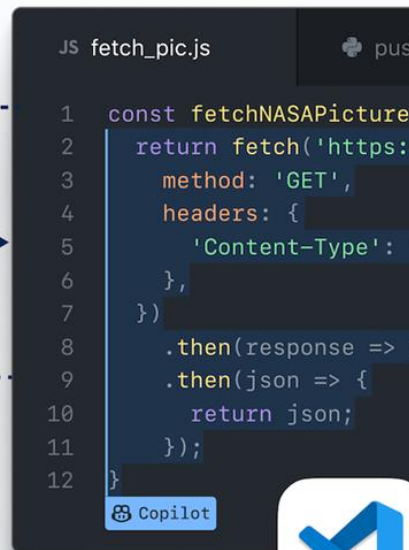


Provide Editor context

Provide Suggestions

Improve Suggestions

Private Code





Demo: Azure OpenAI Service



GitHub Copilot

Once enabled...

40% of new code
written with Copilot

75% of devs felt more
fulfilled with their jobs

87% of devs said it helped
preserve mental effort

github.com/revodavid/copilot-for-r

TS sentiment.ts

-GO write_sql.go

parse_exp

```
1 #!/usr/bin/env ts-node
2
3 import { fetch } from "fetch-h2";
4
5 // Determine whether the sentiment
6 // Use a web service
7 async function isPositive(text: str
8   const response = await fetch('htt
9     method: "POST",
10    body: `text=${text}`,
11    headers: {
12      "Content-Type": "application/
13    },
14  });
15   const json = await response.json(
16   return json.label === "pos";
17 }
```

Copilot

Azure OpenAI Service

GPT-3

Codex

DALL·E *(preview)*



Deployed within your Azure subscription, secured by you, accessed only by you, and tied to your datasets and applications



Large, pretrained AI models to unlock new scenarios



Custom AI models fine-tuned with your data and hyperparameters



Built-in responsible AI to detect and mitigate harmful use



Enterprise-grade security with role-based access control (RBAC) and private networks



Don't fly solo

Developers all over the world use GitHub Copilot to code faster, focus on business logic over boilerplate, and do what matters most: building great software.

Which plan is right for you?

Copilot for Individuals

\$ **10** per month /
\$100 per year

[Start a free trial](#)

- > Plugs right into your editor
- > Turns natural language prompts into code
- > Offers multi-line function suggestions
- > Speeds up test generation
- > Blocks suggestions matching public code

Copilot for Business

\$ **19** per user
per month

[Contact sales](#)

- ← Everything included in Copilot for Individuals, plus...
- > Simple license management
- > Organization-wide policy management
- > Industry-leading privacy

Thank you!

aka.ms/get-copilot

github.com/revodavid/copilot-for-r

David Smith (@revodavid)
Principal Cloud Advocate, Microsoft

Bonus slides

For Q&A

Artificial Intelligence

Machine Learning

Deep Learning

Generative AI



Artificial Intelligence

the field of computer science that seeks to create intelligent machines that can replicate or exceed human intelligence



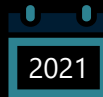
Machine Learning

subset of AI that enables machines to learn from existing data and improve upon that data to make decisions or predictions



Deep Learning

a machine learning technique in which layers of neural networks are used to process data and make decisions



Generative AI

Create new written, visual, and auditory content given prompts or existing data.

Microsoft's AI Principles



Select your preferences

You can change these at any time from your account settings.

Suggestions matching public code *

GitHub Copilot can allow or block suggestions matching public code. See [GitHub Copilot FAQ](#) to learn more.

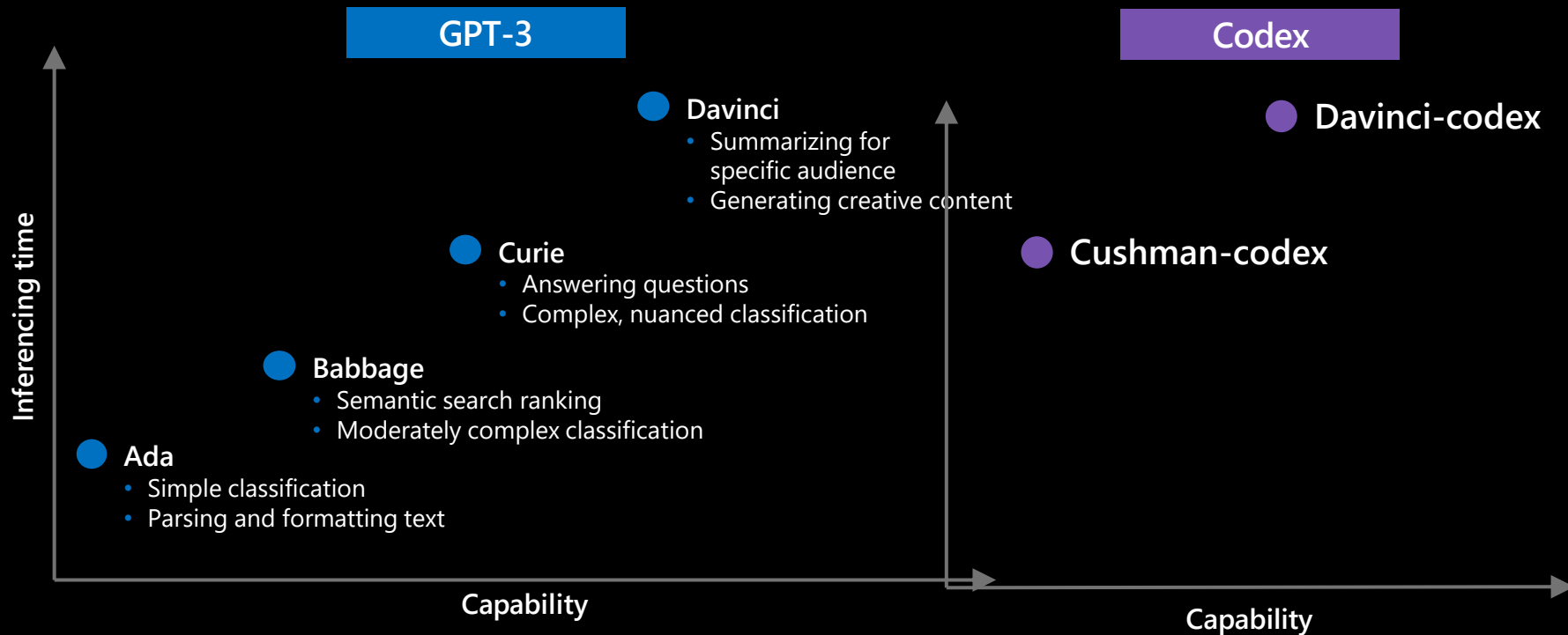
Select an option ▼

☒ **Allow GitHub to use my code snippets for product improvements ***

Allow GitHub and its affiliates to use my code snippets to research and improve GitHub Copilot suggestions, related models and product features. More information in [Privacy FAQ](#).

Save and get started

Azure OpenAI Service models



| Azure OpenAI | GPT-3 Family of Models

Model	Request	Description, performance, cost	Use cases
Davinci	4,000 tokens	Most capable GPT-3 model. Can do any task the other models can do, often with <i>higher quality, longer output</i> and <i>better instruction-following</i> .	Complex intent, cause and effect, summarization for audience
Curie	2048 tokens	Very capable , but <i>faster</i> and <i>lower cost</i> than Davinci.	Language translation, complex classification, text sentiment, summarization
Babbage	2048 tokens	Capable of straightforward tasks, <i>very fast</i> , and <i>lower cost</i> .	Moderate classification, semantic search classification
Ada	2048 tokens	Capable of very simple tasks, usually the <i>fastest</i> model in the GPT-3 series, and <u>lowest cost</u> .	Parsing text, simple classification, address correction, keywords

OpenAI Codex



Makes APIs more accessible



Accelerates software development



Widens who can code

OpenAI Codex Models

Derived from base models and trained on both NL and code (billions of Lines of Code)

Support multiple programming languages

Python, C#, SQL, Java, JavaScript, TypeScript, Go, Perl, PHP, Ruby, Swift, Shell (bash)

Multiple tasks:

Comment → Code

Autocomplete function or next line (in context)

Knowledge searching (API or Library call)

Documenting code (comments)

Refactoring

Use Cases

Natural Language to Code

Documenting code (comments)

Refactoring

Code to Natural Language

Natural Language to SQL

| Azure OpenAI | GPT-3 Models

Powerful language
models accessible to
all skill levels



General purpose text-in/text-out
interface—flexibility



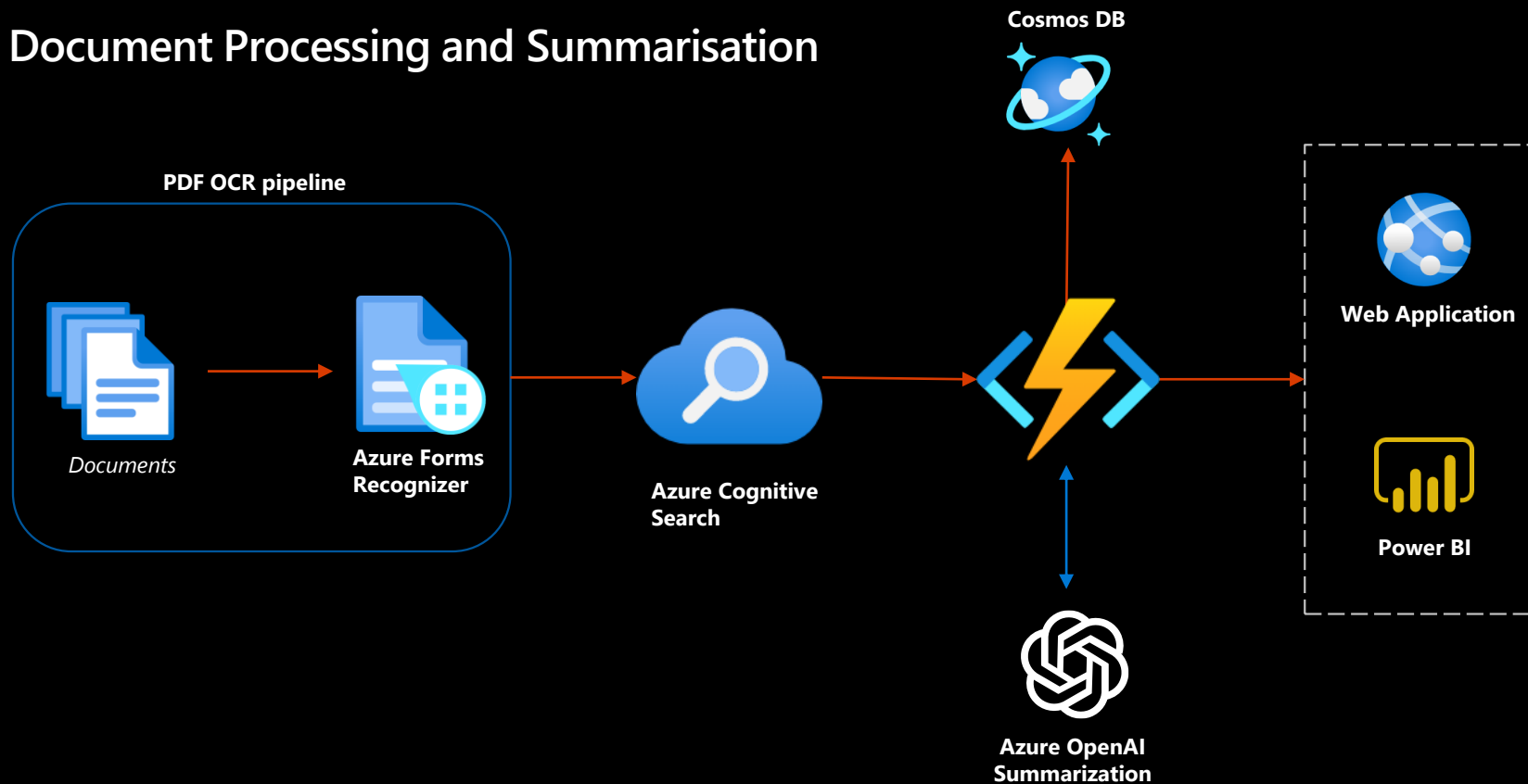
Simple UX—validate proof of concepts fast



Built in ML science intuition for everyone, with
deeper controls for ML practitioners

Azure OpenAI | GPT-3 Sample High Level Architecture

Document Processing and Summarisation



What's next?

githubnext.com