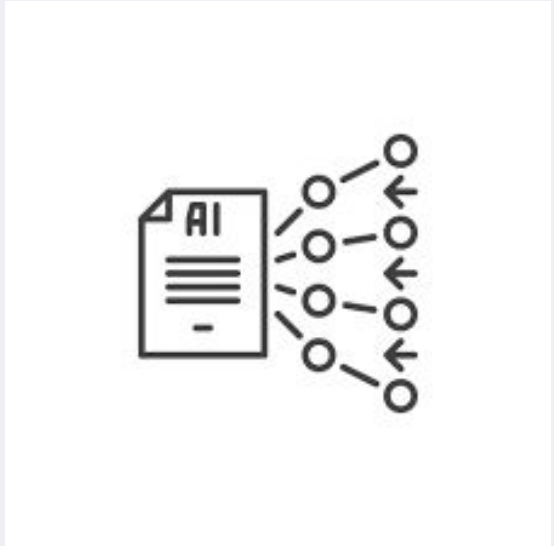
Text Generation

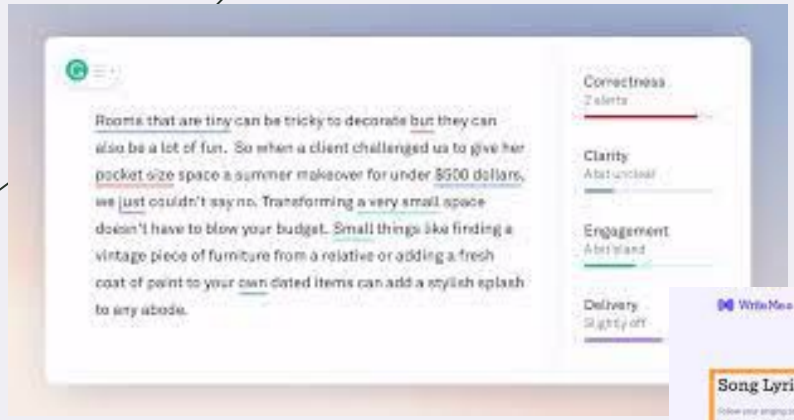
Teaching our robot to create sentences

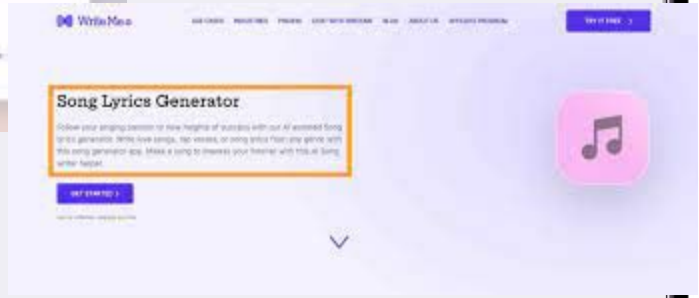
What is text generation?

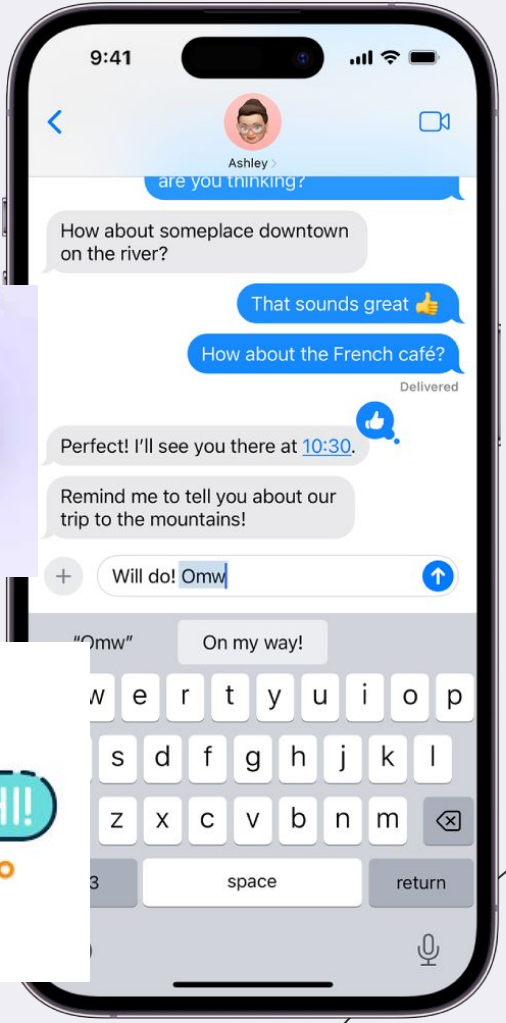
Text generation is having a machine create new text – this can happen in multiple ways.

Can you name some ways that AI generates text?




Rooms that are tiny can be tricky to decorate but they can also be a lot of fun. So when a client challenged us to give her pocket size space a summer makeover for under $500 dollars, we just couldn't say no. Transforming a very small space doesn't have to blow your budget. Small things like finding a vintage piece of furniture from a relative or adding a fresh coat of paint to your own dated items can add a stylish splash to any abode.
Correctness
Clarity
Engagement
Delivery


Song Lyrics Generator


9:41
Ashley
are you thinking?
How about someplace downtown on the river?
That sounds great
How about the French café?
Delivered
Perfect! I'll see you there at 10:30.
Remind me to tell you about our trip to the mountains!
Will do! Omw
"Omw"
On my way!
space
return


May be an image of 1 person, standing, and Machu Picchu
May be an image of 5 people, including Jay Youmens, people playing musical instruments, people standing, 2 hats, and 5 drums

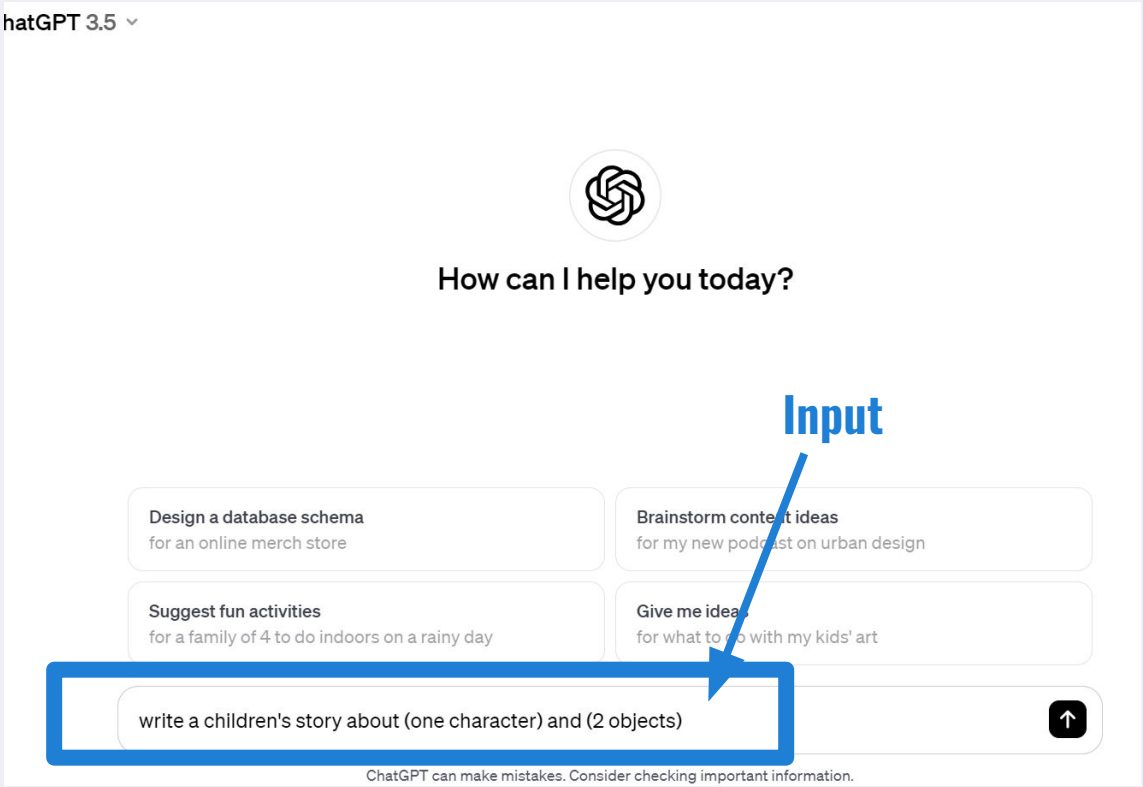

HI!

How do we get a computer to automatically generate text?



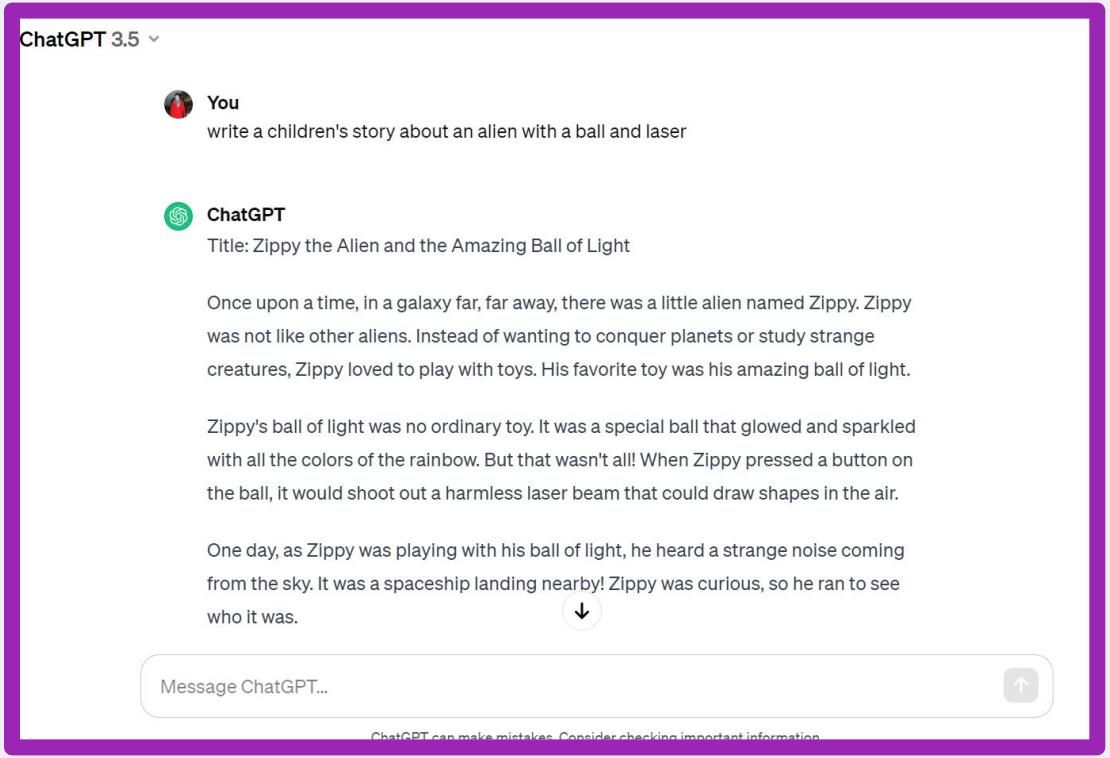
We still need some sort of data to train our model!!

Open Chat GPT



What do you think your trained “AGENT” will do with the input?

Output is a story!



Now you try.

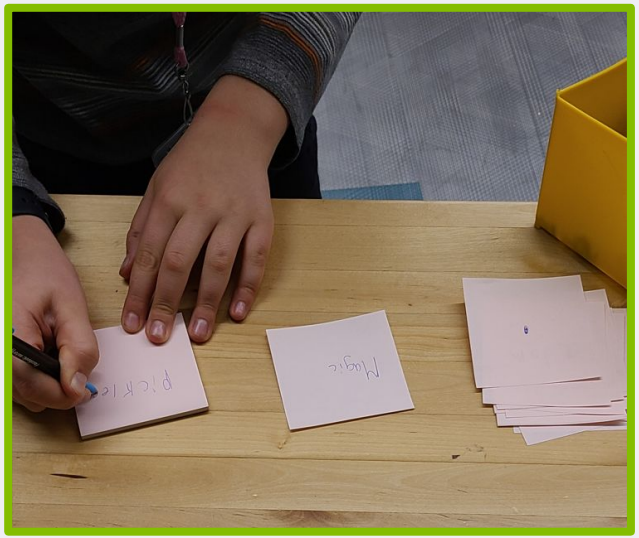
(Then we can talk about training data.)

Today we will focus on the **text generation** task of writing a children's story

Text Generation



Sticky Notes Unplugged



- On your turn, move to the table and write the next word that seems logical
- You will only be able to see the last two words
- During the downtime please finish your “Finish the sentence” sheet

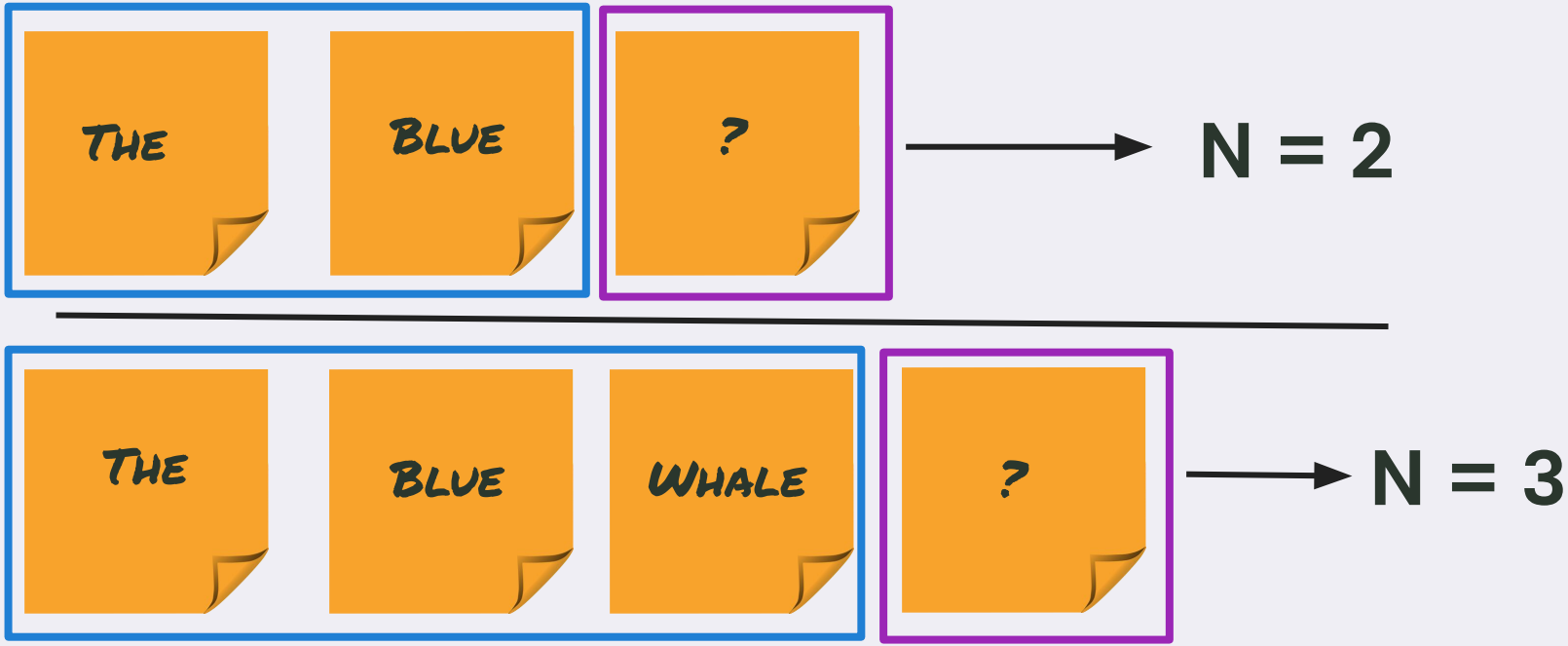
Discussion

- Why did you choose the word that you did?
- Did the final story make sense?
- What would help it make more sense?
- What “data set” did we have?



Ngrams

Ngrams are used in applications like language models, spelling corrections, semantic corrections and more. They are a set of numbers or symbols occurring in an order.



Training with Ngrams?

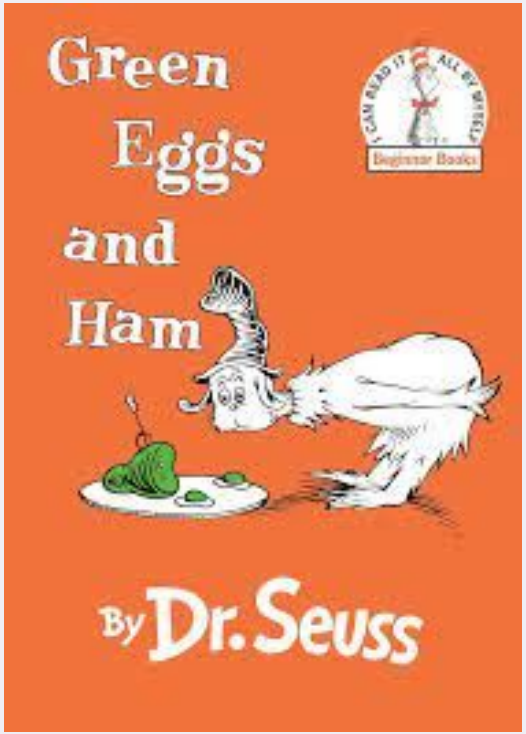
- A sailor went to _____.
- _____ went over the mountain.
- This little _____ of mine.



Sometimes we see sets of words together so often that we are “trained” as to what word comes next.

Task: Use the Green Eggs and Ham dataset to generate a new story.

Instead of all of the words and sentences – we are only going to use Green Eggs and Ham as our data set. A dataset is a collection of data taken from a single source or intended for a single project



We are going to TRAIN our model.

I am Sam
Sam I am

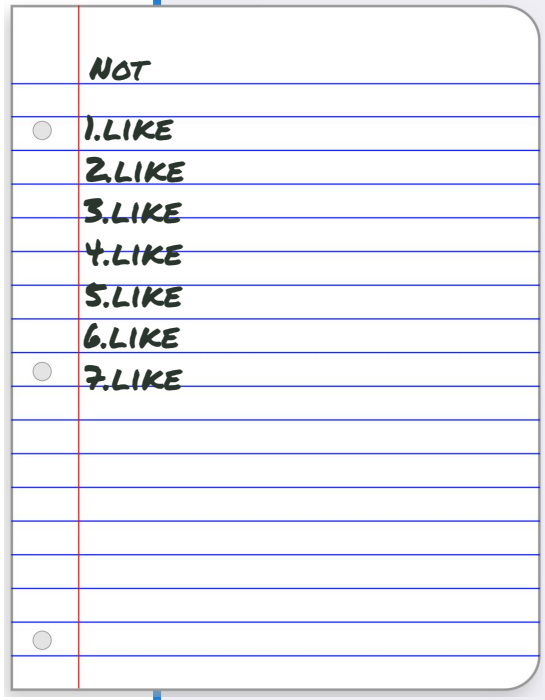
That Sam-I-am!
That Sam-I-am!
I do not like that Sam-I-am!

Do you like
green eggs and ham?
I do not like them, Sam-I-am.
I do not like
green eggs and ham.

Would you like them
here or there?

I would not like them
here or there.
I would not like them anywhere.

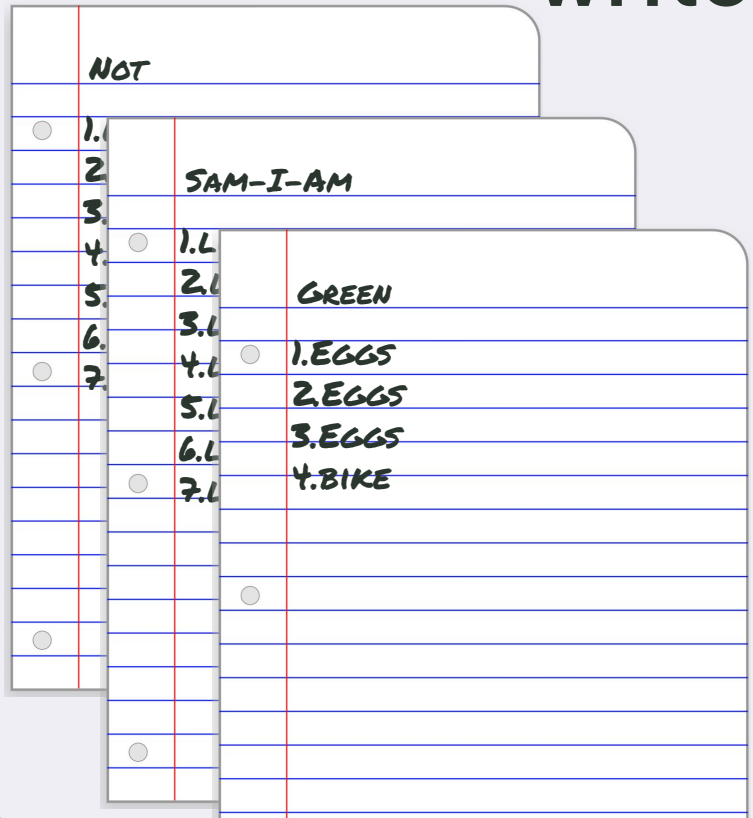
I do not like
green eggs and ham.
I do not like them, Sam-I-am.



Instructions:

- You will be given a word
- You will search that word in the Green Eggs and Ham Text using the word search tool
- You will write down:
  - How many times does that word exist
  - Write each word that comes after that word

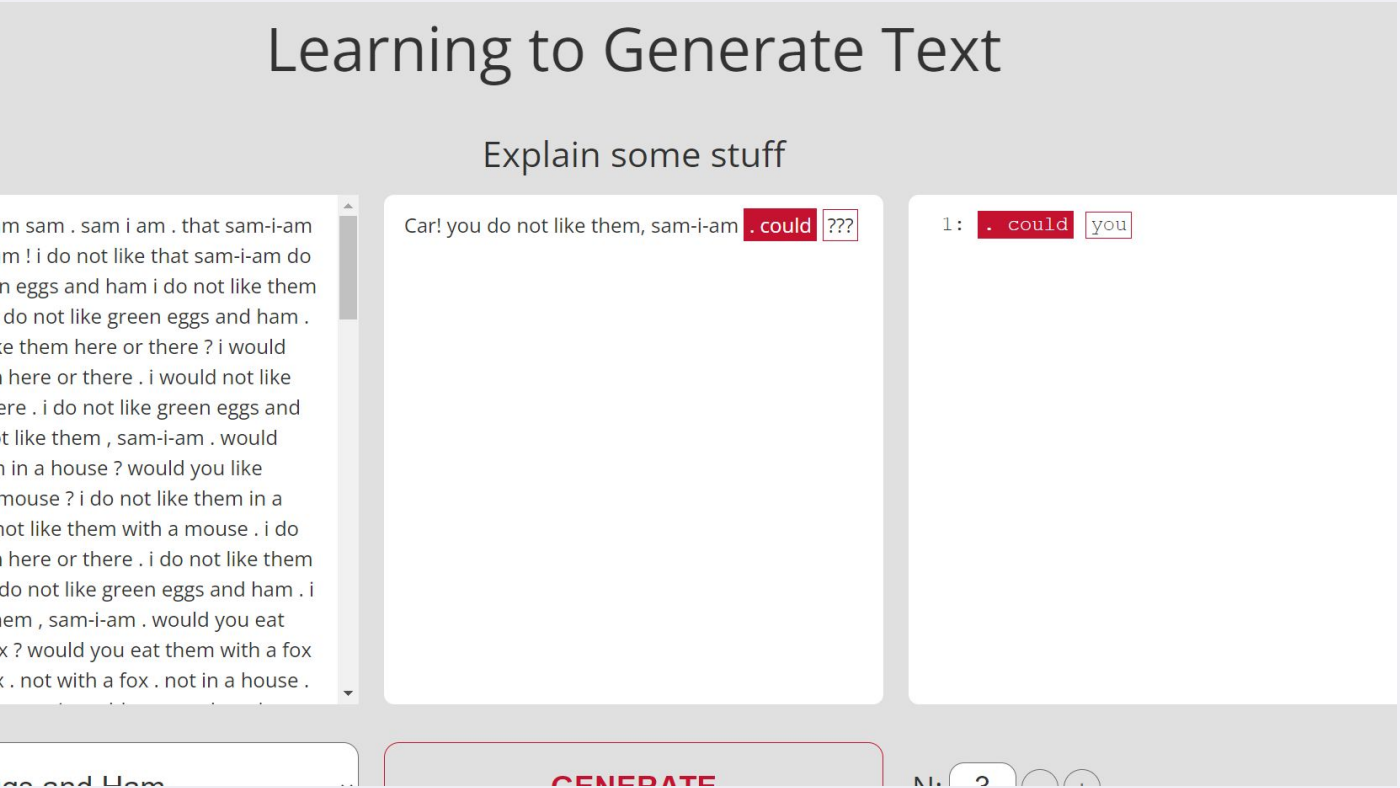
We will use the data you collect to write a story.



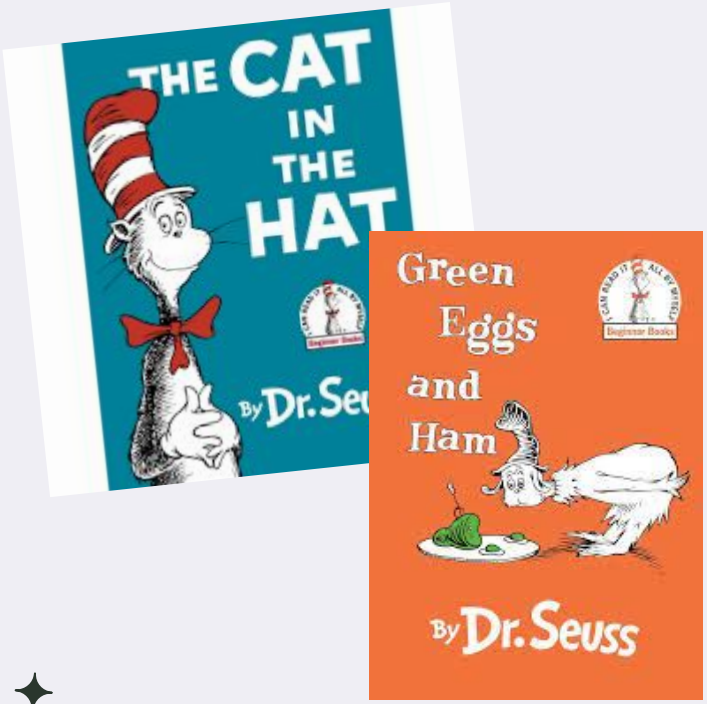
AND I WILL NOT, WILL NOT EAT GREEN EGGS AND HAM..... ?



Let’s have the AI do it instead.

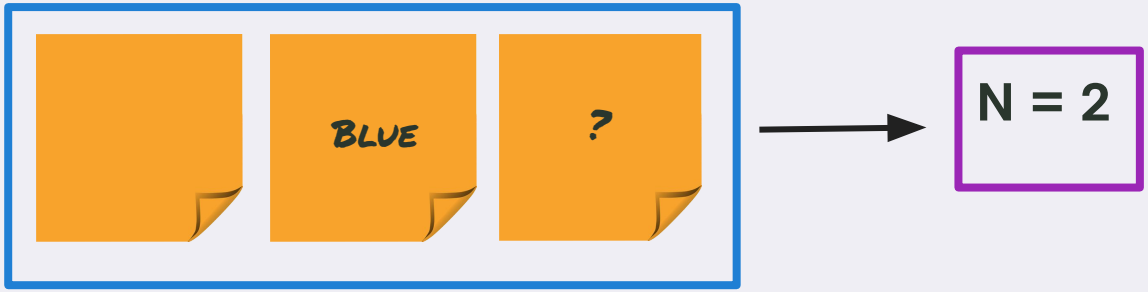


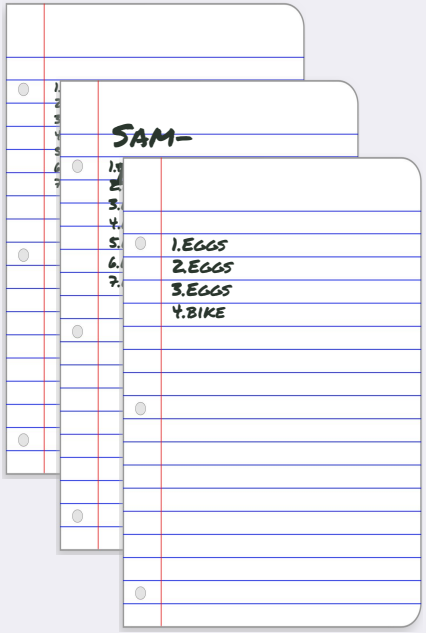
Discussion



- What was the hardest part about data collection?
- Was the computer able to make a good story?
- Was it a very unique story?
- What if we gave the computer more books to pull words from? What are the pros and cons?

Review





- Ngrams is a way for computers to group data
- We needs lots of data to create sensible text
- Data takes a long time to go through for humans