



LONDON

# Explaining how GenAI works ( in a noisy pub )

Simon Aubury

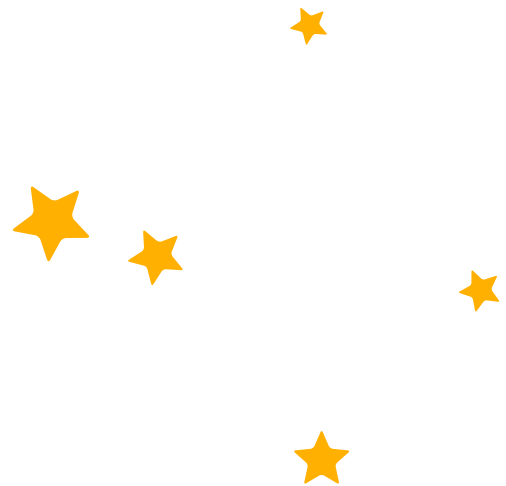


@SimonAubury



# GenAI

You may have heard of it?



# Large language models



## Data



A bunch of text

## Architecture



The transformer

## Training



Pre training & fine tuning

# 1. Data

Let's compress the internet

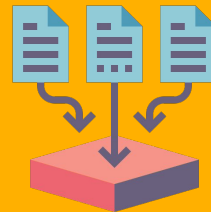
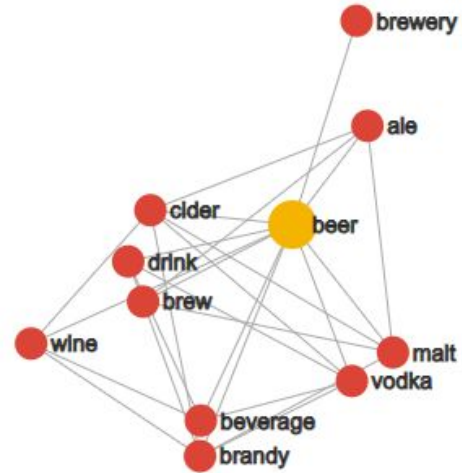




Image by [Republica](#) from [Pixabay](#)

# "Beer"



[WebVectors Online](#) for "beer"





Image by [Christian Birkholz](#) from [Pixabay](#)

## 2. Architecture

The transformer





Image by [delo](#) from [Pixabay](#)





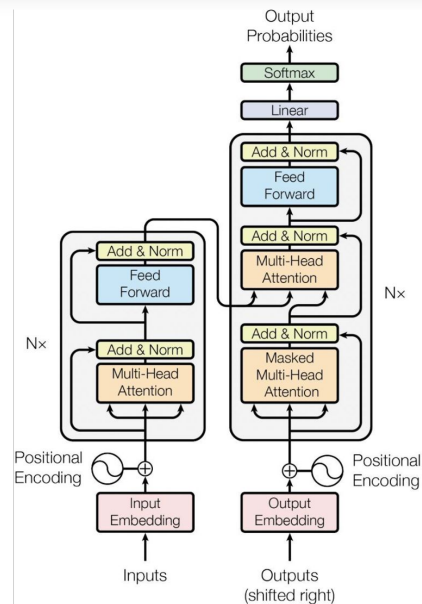
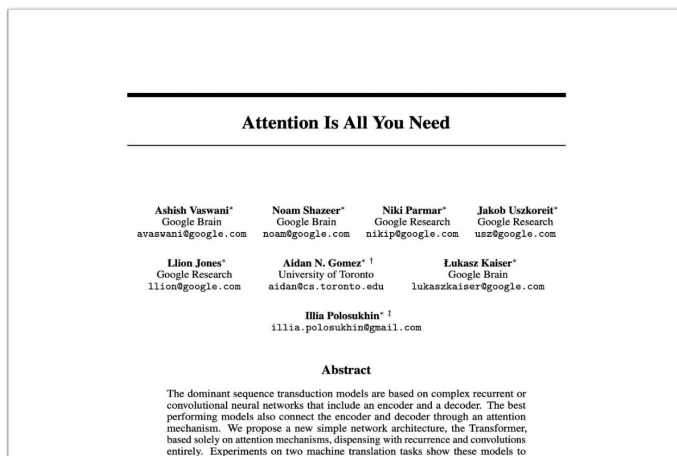
# Understanding language can be challenging

context

The **group of friends** bought cold beer at the local pub to **celebrate their reunion**.

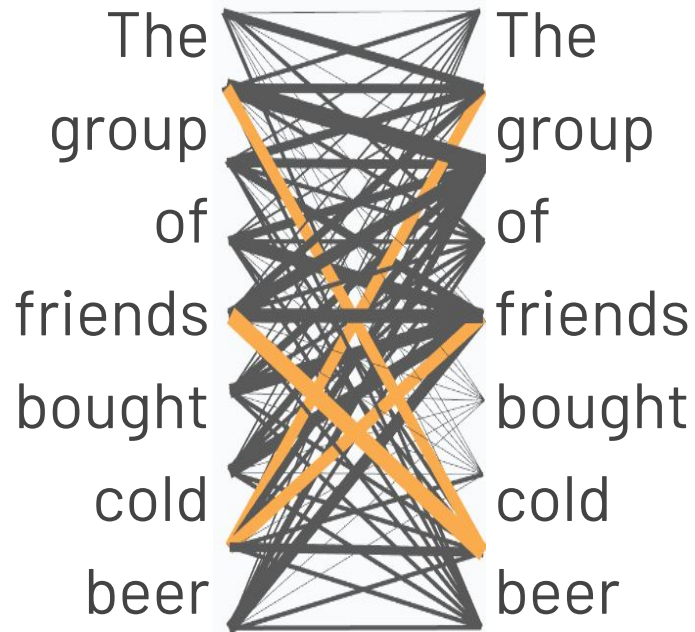
context

# Transformers



Source: Attention Is All You Need <https://arxiv.org/abs/1706.03762>

# Self-attention



## The base model



Context

Base model

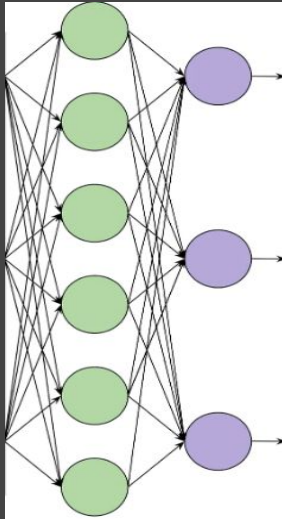
Completion

Beer

is

best

served



cold

with

friends

**A TRANSFORMER + DATA  
= A BASE MODEL**

**BASE MODEL + CONTEXT  
= COMPLETION**

# 3. Training

Pre training & fine-tuning





# How to train your LLM



## Pretraining (once)

1. PB's of text.
2. 1000's of GPUs.
3. Compress the text into a neural network,
4. Pay 💰 wait 📅



Obtain base model.

## Fine-tuning (recurring)

1. 1000's ideal Q&A responses (human)
2. Finetune base model on this data wait 📅
3. Obtain assistant model
4. Evaluate, deploy & monitor

# Reinforcement Learning from Human Feedback



RE

**You**

write a 4 line poem about beer



**ChatGPT**

In amber waves, the brew finds cheer,  
A hoppy dance, a malty leer.  
Under moon's glow or sun's bright sear,  
Beer's hearty laugh, we hold so dear.

RE

**You**

write a 4 line poem about beer



**ChatGPT**

In golden hues, beneath the moon's soft leer,  
A frothy whisper, tales of joy and cheer.  
In every gulp, a story does appear,  
A bond, a laugh, a moment held so dear.

RE

**You**

write a 4 line poem about beer



**ChatGPT**

In golden pours, a tale unfolds,  
Beneath the froth, a story told.  
In every sip, a moment's cheer,  
The hearty laugh, the warmth of beer.



It is faster to pick than to generate.

# Large language models



## Data



A bunch of text

## Architecture

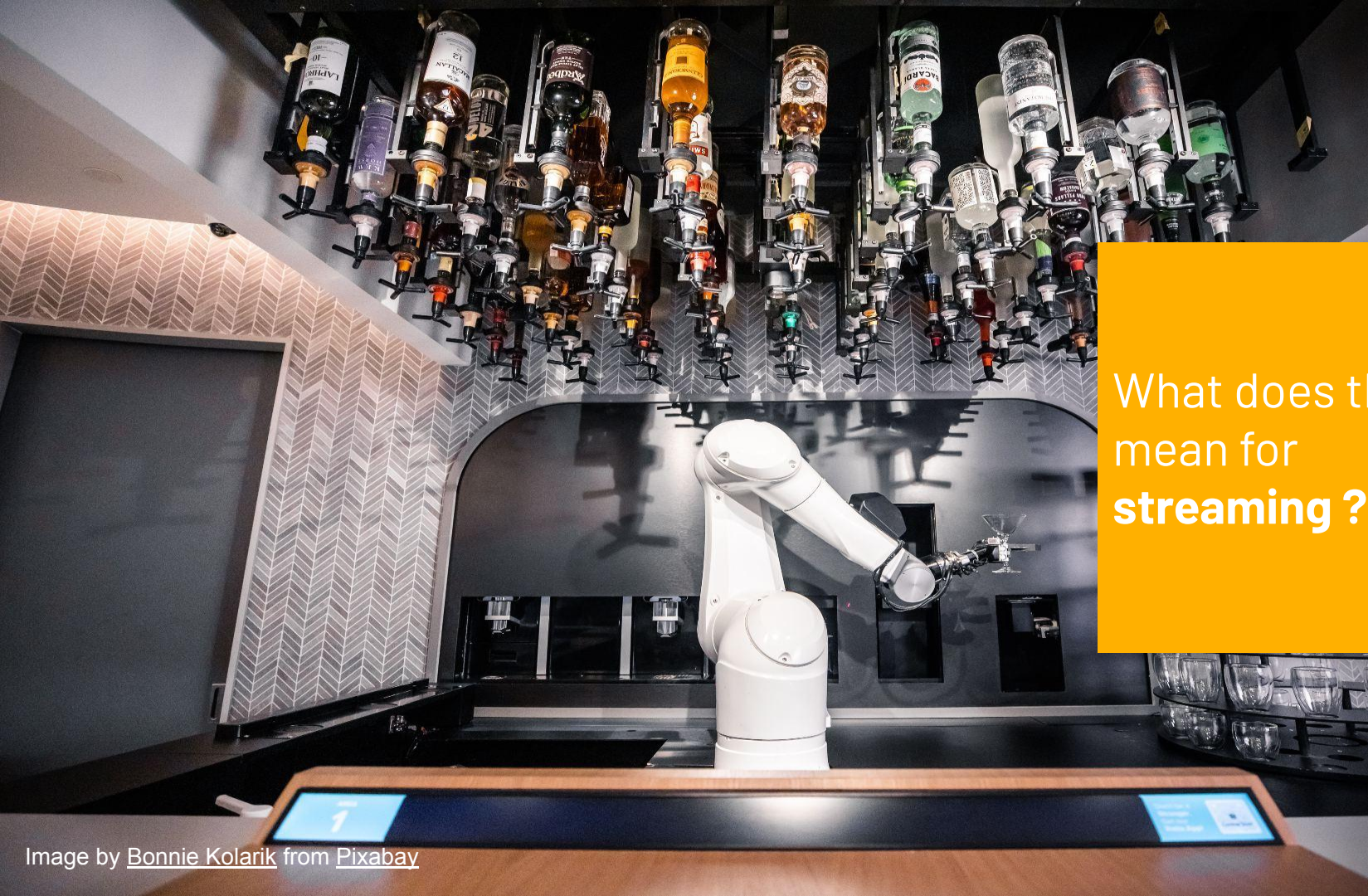


The transformer

## Training



Pre training & fine tuning



What does this  
mean for  
**streaming ?**



**The entire history of software engineering is one of rising levels of abstraction.**

**Grady Booch, IBM chief scientist**



# Thanks!



@SimonAubury