

ETC5523: Communication with Data

Assignment 1: Example breakdown of article

AUTHOR
Caoyu Shao

PUBLISHED
August 24, 2025

Piece 1 — List A(Print media)

Title: Mapping Australia's childcare blackspots(ABC News)

Links: https://www.abc.net.au/news/2022-03-22/mapping-australia-s-childcare-blackspots/100894808?utm_campaign=abc_news_web&utm_content=link&utm_medium=content_shared&utm_source=abc_news_web

Two-sentence summary: Australia has serious geographical inequality in access to childcare and over one-third of Australians live in "childcare deserts". The article starts with an family story and show interactive maps to show deserts concentrate on city fringes and regional areas while "oases" cluster in wealthier inner suburbs in large cities.

Lede: At the end of the article, there is a box and we could input our residential address and get the scarcity of child care resources nearby, which is very rare in the online articles. But it is becoming more and more frequent because it could enhance the readers' sense of participation.

The background information in the middle of the article:

The data from the Mitchell Institute for Education and Health Policy at Victoria University, supplied exclusively to the ABC, found stark geographical divides in access to child care. The project was partially funded by the Minderoo Foundation as part of the [Thrive by Five initiative](#).

Child care is critical to women's employment. This, in turn, drives economic growth, boosts financial security and improves social and health outcomes for women and families.

Ultimately, if people can't get the child care they need, they're not going to be able to work, Dr Hurley says.

Background: What the Mitchell Institute measured. It provides the proof of our main results——National maps and key numbers and determines the scope of statistical scope. Why childcare access matters for work, equity and early learning, it indicates the importance and necessity of the article.

The main result in the middle of the article:

"The key finding is that where you live matters when it comes to access to child care," says Peter Hurley, a Mitchell Institute education policy fellow and lead author of the report.

The heat maps about major cities in the article

Main result: National maps +key numbers. The maps have straightforward illustrated the inequality of children care access.

The summary in the middle of the article:

Sprawling deserts, scattered oases

Higher supply, higher fees

Summary: Re-anchors to inequality; policy takeaways.

Piece 2 — List B(Blog post)

Title: How does ChatGPT work? As explained by the ChatGPT team — The Pragmatic Engineer.

Link: <https://newsletter.pragmaticengineer.com/p/scaling-chatgpt?ref=blog.pragmaticengineer.com>

The summary in the middle of the article:

Input. We take your text from the text input.

Tokenization. We chunk it into tokens. A token roughly maps to a couple of unicode characters. You can think of it as a word.

Create embeddings. We turn each token into a vector of numbers. These are called embeddings.

Multiply embeddings by model weights. We then multiply these embeddings by hundreds of billions of model weights.

Sample a prediction. At the end of this multiplication, the vector of numbers represents the probability of the next most likely token. That next most likely token are the next few characters that spit out of ChatGPT.

Two sentence summary: This post is a Q&A with an OpenAI engineering explaining how ChatGPT works. It uses respond pipelines and five-step graphs to demystify large language models for the general people.

The background information in the middle of the article:

Sometimes the best explanations of how a technology solution works come from the software engineers who built it. To explain how ChatGPT (and other large language models) operate, I turned to the ChatGPT engineering team.

I asked this from Evan Morikawa at OpenAI. Evan joined OpenAI in 2020 – two years before ChatGPT launched – and has led the Applied engineering team as ChatGPT launched and scaled. His team was the one that created ChatGPT, and Evan has been there from the very beginning.

Background: Who the interviewee is & Why interview him, this could enhance the credibility of the post blog by stating the importance of interviewee. Context on the definitions, this could help the readers understand easily because it uses plain words to explain abstract concepts.

Lede: Author prepares the interview and launches it with the form of Q&A instead of article. Besides the Q&A, enumerate the steps plainly, like input, tokenisation, create embeddings, multiply embeddings by model weights, sample a prediction. As a reader without any relevant knowledge, we could simply know how LLM works.

This is the main result in the article:

Input. We take your text from the text input.

Tokenization. We chunk it into tokens. A token roughly maps to a couple of unicode characters. You can think of it as a word.

Create embeddings. We turn each token into a vector of numbers. These are called embeddings.

Multiply embeddings by model weights. We then multiply these embeddings by hundreds of billions of model weights.

Sample a prediction. At the end of this multiplication, the vector of numbers represents the probability of the next most likely token. That next most likely token are the next few characters that spit out of ChatGPT.

How do we generate this complex set of model weights, whose values encode most of human knowledge? We do it through a process called pretraining.

Once we have our model we can run inference on it, which is when we prompt the model with text. It makes this prediction based on past input, and it happens repeatedly, token by token, word by word, until it spits out your post!

Main result: Five steps of pipelines (input, tokenisations, embeddings, matrix, multisampling); Pretraining and inference.

This is the summary of the article:

How ChatGPT works isn't magic, and is worth understanding.

Summary: How ChatGPT works isn't magic, and is worth understanding

Piece 3 — List C(Video)

Title : Why I fell in love with monster prime numbers —Adam Spencer(TED)

Link: <https://www.youtube.com/watch?v=B4xOFsygwr4>

Two-sentence summary: A lively talk uses record-prime hunts to convey the wonder of mathematics and collaborative computation. Humor and visuals make a huge, abstract number feel tangible and exciting.

Background: What primes number is, it provide the mathematical foundation of audience so that they could get the talk show; Why the large prime number matter, it tells the audience the importance of the speech.

Lade: Highly academic content can be presented in the form of a talk show. In our impression, talk shows seem to be very relaxed, while academics are very serious. But Adam combined the two together. In fact, this kind of talk show could reduce our cognitive loads.

Main result: Stories/ visuals of record primes.

Summary: Inspire curiosity; invites ongoing discovery

Combined reflection

Difference

1. **Openings:** News initiated with an empathetic vignette; the blog starts with an authority hand-off; the talk uses performance energy.
2. **Summaries:** The news use a interactive map to illustrate the conclusion; the blog post uses five steps of pipelines to show how LLM works; the Ted talk use stage visuals and narratives in the stage.

Reasons for difference

1. News faces the mass and it needs vignette to attract the attention and maps to indicate the inequality in childcare access straightforwardly. Readers could get conclusions by reading the heat map. Lastly, at the end of the article we could input our post code, and then the scarcity of childcare resources will appear.
2. Blog post's readers are mostly followers, who are general people, not experts in LLM. So the introduction of the interviewee could improve its authority and schematics could separate the steps how LLM works. Most importantly, the post blog break the process into plain words. It's easier for readers to remember the process.
3. The Ted talk introduced many concepts and narratives to make the abstract concepts like prime number concrete. The stage could help impress the audience. Compared to the previous two forms, it reduced the audiences' cognitive burden.