MathWorks
AUTOMOTIVE CONFERENCE 2023
Europe

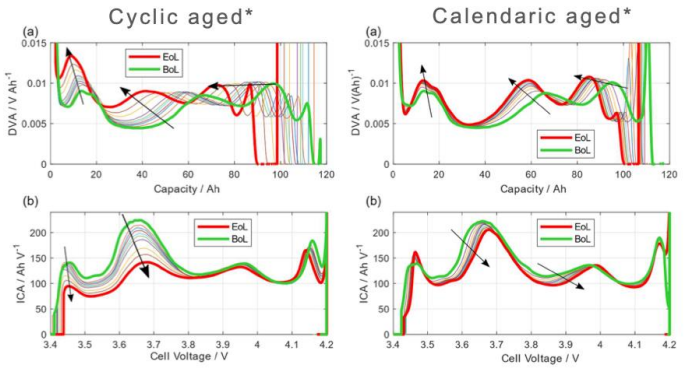
ChatGPT and Large Language Models with MATLAB

Deborah Ferreira, PhD, MathWorks

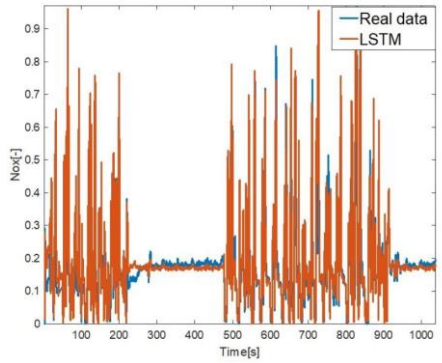
MathWorks®

AI is making its way into all stages of bringing vehicles to market

R&D

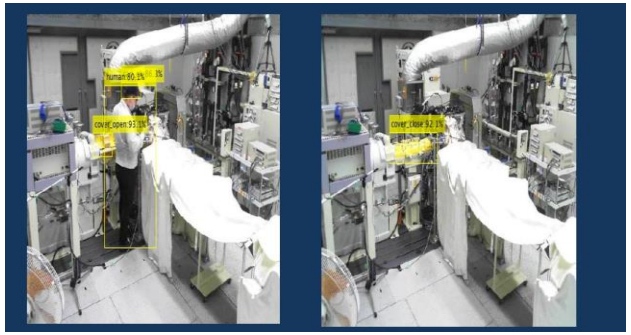


Gotion
Battery SOH Estimation



Renault
Virtual sensor for NOx estimation

Manufacturing



Toyota
Predictive maintenance of engine bench

Daihatsu
Engine knock detection

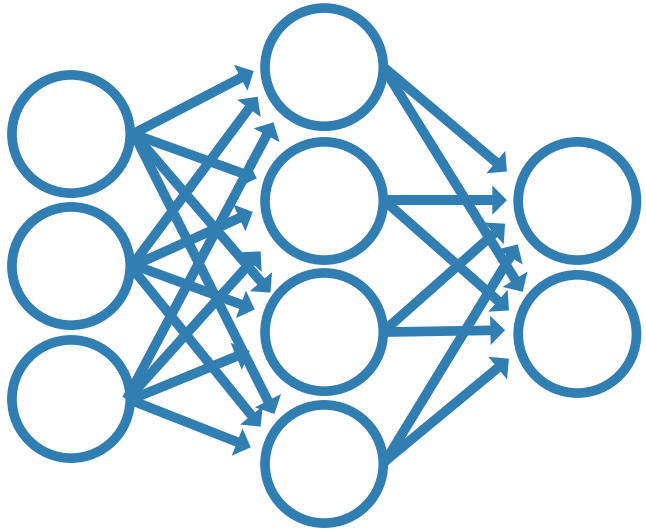
Disruptive AI technology: ChatGPT
Is it also going to disrupt the automotive industry?

Example:
Hey Mercedes integration with ChatGPT

But wait, what is it?

- It’s an AI chatbot, developed by OpenAI
- Built on GPT-3.5 and GPT-4, focused on human alignment
- GPTs are generative pre-trained transformers, a type of *large language models*

What are Large Language Models?

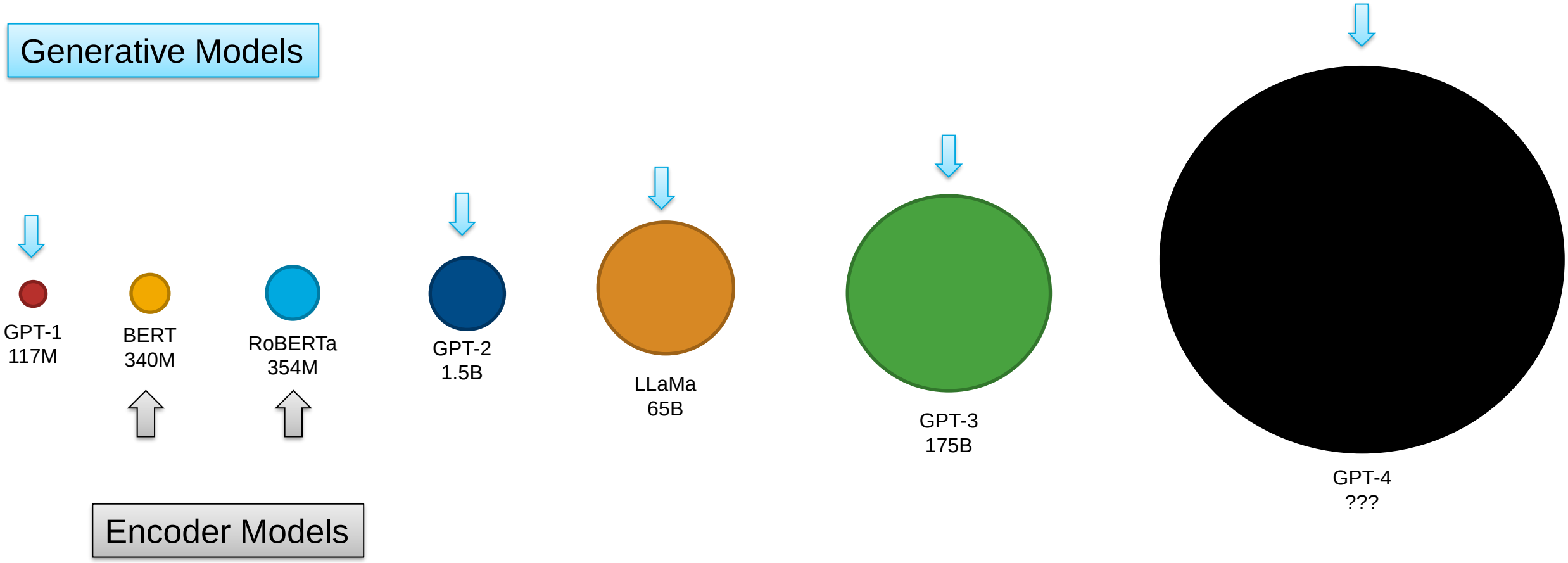


Large language models (or LLMs) are a type of artificial intelligence model

Trained on huge amounts of data

Networks with millions to trillions of parameters

Large Language Models Landscape



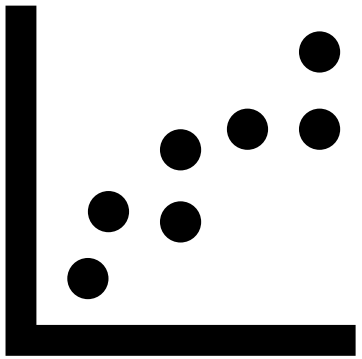
LLMs generate output word by word

What is the next word?

Accelerating the pace of engineering and science

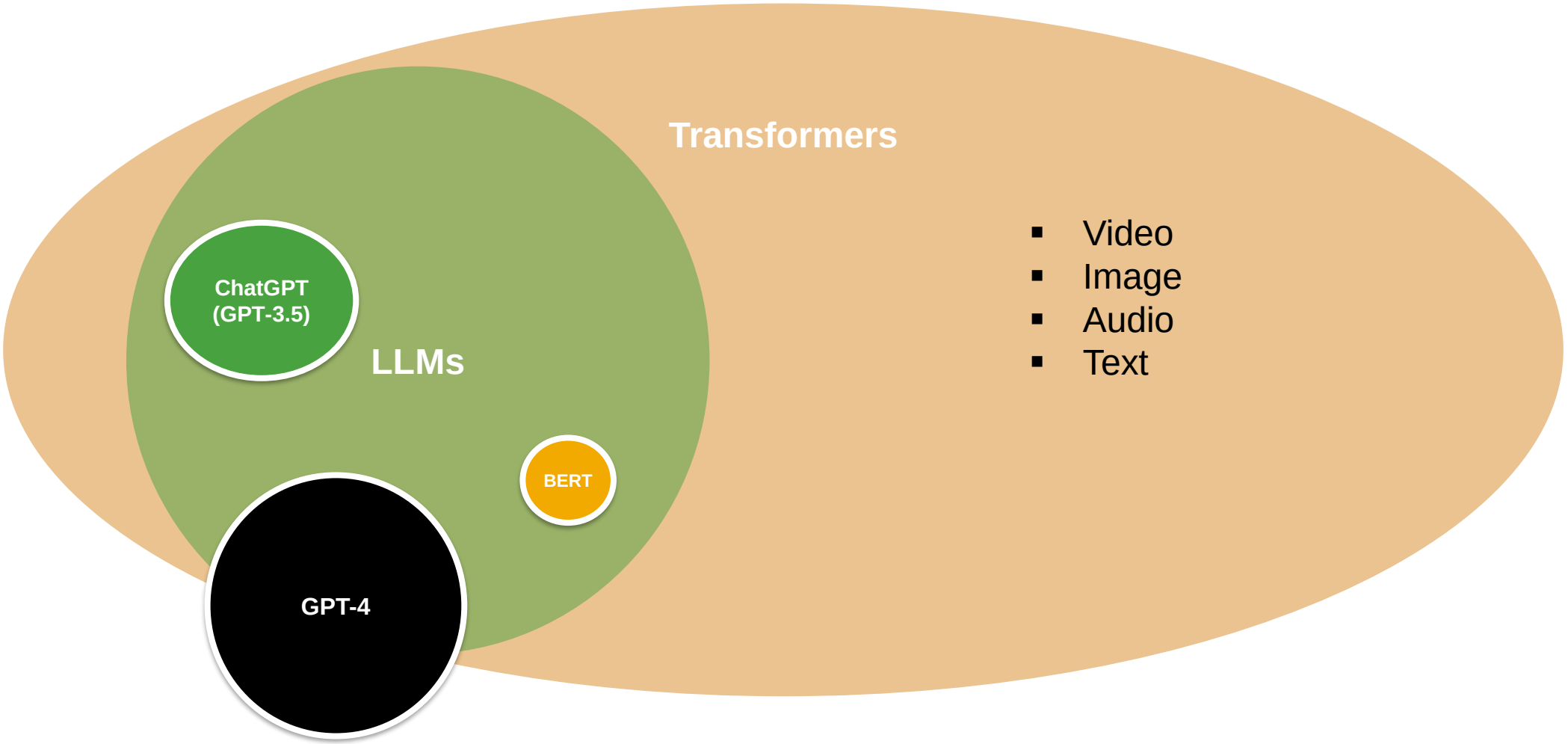
with tech plus research
of innovation and knowledge
then engineering or science

They do not understand text in the way humans do.



They recognize patterns in the data they were trained on.

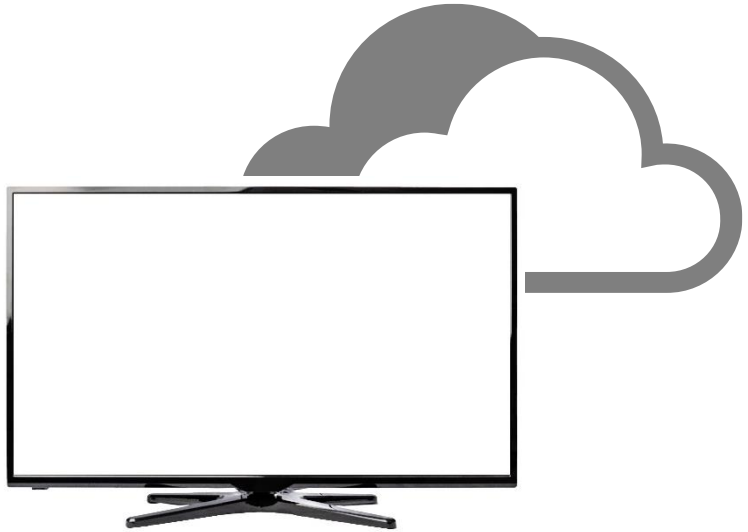
Large Language Models are Transformer Models



Different ways of interacting with this type of model

Proprietary/Closed Models,
Access through API

ChatGPT



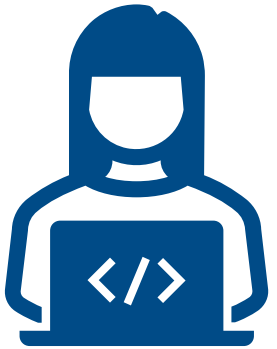
Open models,
user owned

Falcon, LLaMA

What Large Language Models are great at…

Natural Language Understanding

Generating text and code



Large Language Models are not great at…

Absence of Source Citation

Hallucination

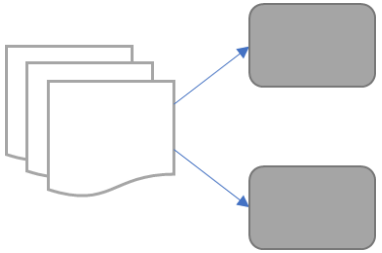
LLMs can help you solve specific problems

Generate Text

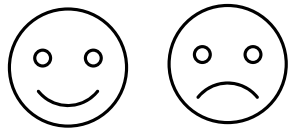
Write Code

Document

Classify Text

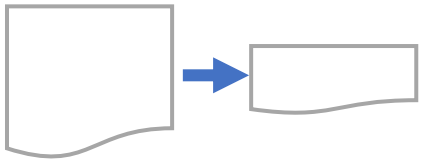


Classify Maintenance Documents

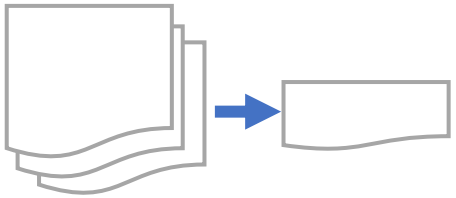


Analyze Sentiment

Summarize



One document



Multiple documents

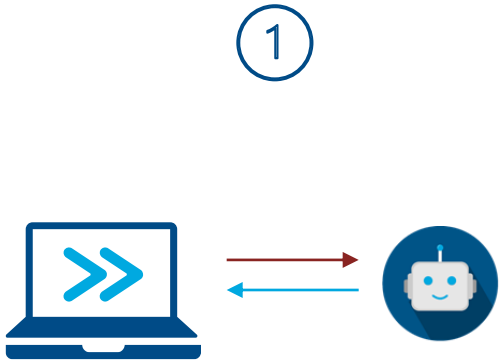
Retrieve Information

Scientific Discovery

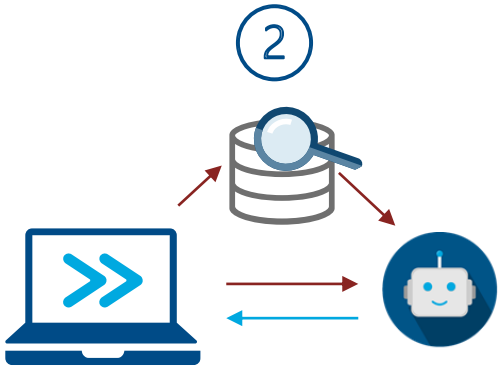
Requirements for design & engineering

Knowledge from internal documents

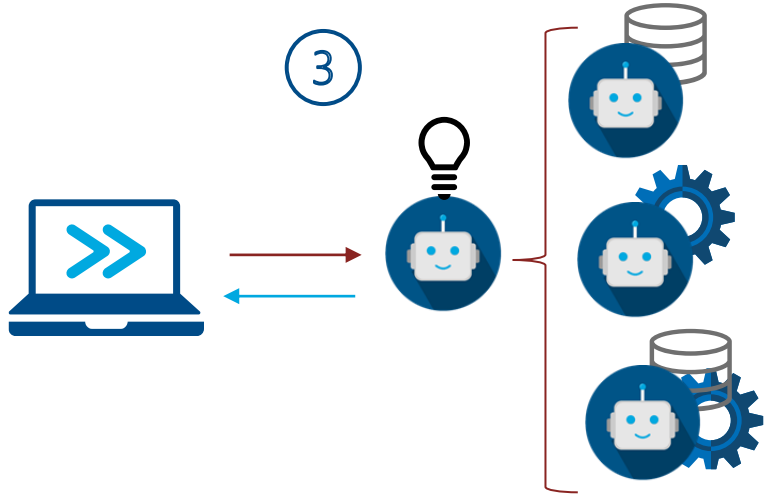
Approaches for leveraging LLMs



Prompt Engineering



Retrieval-Augmented Generation

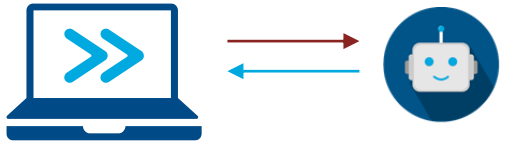


Fine-tuning

For some things it's just easier to use the chat interface…

Generate MATLAB code to compute the sine of values between 1 and 30. The code should be optimized and use vectorization. Let's think step-by-step.

Prompt engineering is the process of crafting and refining inputs to an AI model to guide its output.



Sure! Here's the MATLAB code that computes the sine of values between 1 and 30 using vectorization:

matlab — Copy code

```matlab
% Define the range of values
startValue = 1;
endValue = 30;

% Generate the values within the range
x = startValue:endValue;

% Compute the sine of the values using vectorization
y = sin(x);

% Display the results
disp(y);
```

For other things, you might want to use ChatGPT programmatically

…

Design notes

Summarize…

1. Copy and paste each document individually into the chat interface.
2. Ask ChatGPT for a summary
3. Copy and paste the result into another file

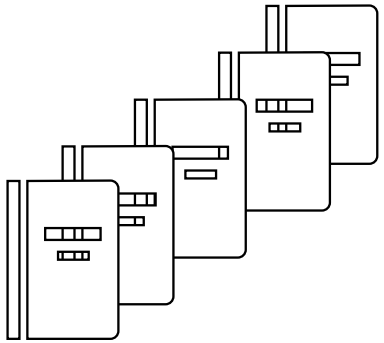
1. Read each file inside MATLAB in a loop
2. Request the ChatGPT API for a summary for each document
3. Automatically save each summary to a file

If there is repetition, there is potential for automation via MATLAB

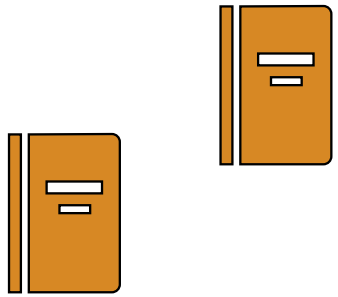
Retrieval-Augmented Generation (RAG) to ask questions about your documents

Text Analytics Toolbox

Documents
maintenance reports



Retrieve relevant
reports



ChatGPT

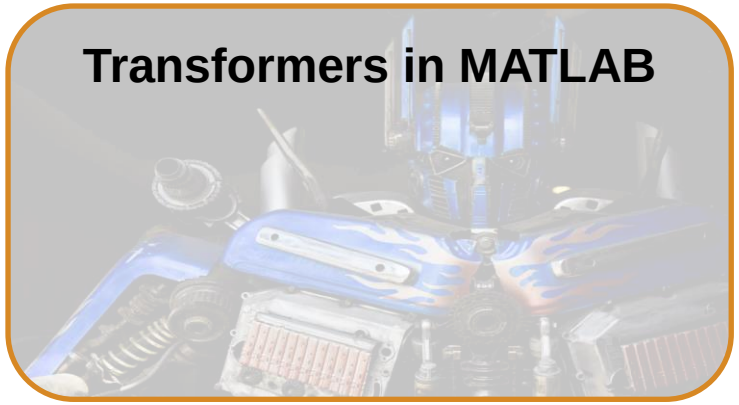
Focused
answer

Query: Model XYZ had a faulty spark plug, what should I do?

Use RAG when you want to tailor the model's output to your own data!

Fine-tuning Large Language Models

For domain-specific text, you might want to fine-tune (adapt) a model to your own data.



Fine-tuning a large model requires parameter efficient fine-tuning techniques.

Fine-tuning models

Models like BERT can be fine-tuned in low-resource settings