

LIBER Learning Hub

LIBER DSLib & DSDCH

2025-07-10

Table of contents

Preface	3
1 Introduction	4
2 Understanding Large Language Models through an introduction to ChatGPT	5
2.1 Overview	5
2.2 What are Large Language Models (LLMs)?	5
2.3 What is ChatGPT?	6
2.3.1 What ChatGPT is NOT:	7
2.3.2 How ChatGPT Generates Responses:	7
2.3.3 When best to use ChatGPT:	8
2.5 Large Language Models (LLM) and ChatGPT Uses in Libraries Today	9
2.6 Hands On Activity	9
2.6.1 Writing prompts	9
2.7 Exploratory activities:	10
2.8 Recommended Reading/Further Learning	18
2.8.1 **Bas	18
2.8.2 **Gui	19
References	20

Preface

This is a Quarto book.

To learn more about Quarto books visit <https://quarto.org/docs/books>.

1 Introduction

This is a book created from markdown and executable code.

See Knuth (1984) for additional discussion of literate programming.

2 Understanding Large Language Models through an introduction to ChatGPT

A brief introduction to Large Language models (LLMs) by way of ChatGPT, this guide will demystify both. Through hands-on activities and case studies, you'll explore potential challenges and opportunities for the use of LLMs and ChatGPT in current and future library services.

2.1 Overview

2.2 What are Large Language Models (LLMs)?

A large language model (LLM) is a type of **machine learning model** that can perform **natural language processing** tasks such as generating and classifying text, answering questions, and translating text.

If your eyes just crossed reading that sentence, no wonder, there's a lot of terminology and jargon to unpack there in that one little sentence. Let's go back a few steps and take it a little slower!

It helps to first understand a little bit about what machine learning is. **Machine learning (ML)** is a core technology that underpins a lot of work in artificial intelligence work today. Artificial intelligence is actually a broad field that encompasses the development of systems and machines capable of performing tasks that typically require human intelligence, such as reasoning, problem-solving, learning, and perception, and...chatting!

Natural Language Processing (NLP) is a specialised area within artificial intelligence (AI) that looks specifically at the challenges of processing and understanding human language (another is Computer Vision (CV) which is, on the other hand, concerned with helping computers process and understand imagery in the vein of human vision). NLP plays a vital role in enhancing the capabilities of AI systems by working to make them more capable of natural and effective interaction with humans.

NLP involves the development of algorithms, models, and systems for analysing, understanding and extracting meaningful information from textual data representing human language. The field relies heavily on the use of machine learning to make sense of our complex textual data,

which might come from literature, historical documents, articles, twitter, wikipedia, you name it!

Machine learning involves the development of algorithms and models that allow computers to learn patterns and relationships from data and make predictions on new data. Instead of being explicitly programmed for specific tasks, ML algorithms use data to learn and improve their performance over time.

A **machine learning algorithm** is designed to recognize patterns, relationships, and trends in data and will adjust their internal parameters based on examples to improve their performance over time.

A **machine learning model** represents what was learned by a machine learning algorithm. The model that's created in the process and saved after running a machine learning algorithm on training data and represents the rules, numbers, and any other algorithm-specific data structures required to make predictions on new data.

So back to our definition:

A large language model (LLM) is a type of **machine learning model** that can perform **natural language processing** tasks such as generating and classifying text, answering questions, and translating text.

Language models are a type of machine learning model designed to predict the likelihood of a sequence of text, which means that they can be set up to predict the most likely way to continue a conversation. The models behind ChatGPT are large, complex neural networks that have been exposed to an enormous amount of text from books, articles, websites, and more. These models have learned something about patterns in grammar and word meaning, including the way that meaning arises contextually across multiple sentences and multiple turns in a conversation. When you ask ChatGPT a question, you are presenting the model with new information it tries to make a prediction on, in this case, it tries to generate a response that matches the pattern of conversation. Large language models are trained on massive amounts of text data and are essentially the backbone of NLP today. There are and have been for many years, large language models out there actually but ChatGPT has currently caught the popular imagination because of its publicly available interface and remarkable performance.

2.3 What is ChatGPT?

ChatGPT is a platform providing an interface specifically to the large language models developed by OpenAI. The GPT stands for [Generative Pre-Trained Transformers](#) which is just a fancy way of saying that the models you are chatting with have already been trained on a lot of data.

It's a chatbot essentially designed to generate human-like text responses based on the input it receives. Or some say to think of it as a fancy predictive text generator. You ask questions or

give prompts, and it provides responses in natural language, or rather, estimates what should come next in a conversation.

2.3.1 What ChatGPT is NOT:

- Google Search Box: Unlike Google, ChatGPT doesn't search the internet for information. Instead, it generates responses based on a massive dataset it was trained on. This training dataset includes large segments of text sampled from the internet at a specific point in time, and so can reflect language describing content on the internet, but it doesn't have real-time access to the internet. When ChatGPT answers a question, it isn't actually looking up information and then composing that information into a response; it's just making an estimation of a response based on patterns it has seen. So, when you ask it factual questions, especially ones with common answers or phrases, it might give you an answer that sounds right but remember this is because it's mimicking what it has seen in its training data. Think of it perhaps as a very smart parrot, sometimes 'repeating' things it has seen before.
- Human Mind: ChatGPT is not a human; it's a machine learning model. It doesn't understand or reason like a human. It generates evidently linguistic responses based on patterns it learned from text data.

2.3.2 How ChatGPT Generates Responses:

- Language Modelling: So ChatGPT models are members of this class of models we're talking about: "large language models" (LLMs). A machine learning model represents what was learned by a machine learning algorithm. The model is created and saved after running a machine learning algorithm on training data and represents the rules, numbers, and any other algorithm-specific data structures required to make predictions on new data. Language models are designed to predict the likelihood of a sequence of text, which means that they can be set up to predict the most likely way to continue a conversation. The models behind ChatGPT are large, complex neural networks that have been exposed to an enormous amount of text from books, articles, websites, and more. These models have learned something about patterns in grammar and word meaning, including the way that meaning arises contextually across multiple sentences and multiple turns in a conversation. When you ask ChatGPT a question, it tries to generate a response that matches these patterns.
- Contextual Understanding: ChatGPT also looks at the context of the question you ask, including what's happened so far in your conversation and any instructions or details you offer about how it should respond. It tries to understand the context and generate a response that is contextually relevant, which can make its answers seem quite accurate.

- Existing Knowledge: It can recall and blend information it has observed in its training data. So, when you ask a factual question, it can sometimes generate a response that seems factual because it has encountered similar information during training. But, it can also make mistakes or provide information that is outdated or biased because it doesn't have the ability to verify the accuracy of what it says, and it is influenced by any views inherent in its training data, whether they are right or wrong. It's like the parrot repeating something it heard: it can't check if what it's saying is true or up-to-date. That's why it's important to double-check information output by ChatGPT with reliable sources, especially when using ChatGPT for research or critical tasks.
- User Feedback Loop: ChatGPT models like the latest [GPT-4 \(openai.com\)](https://openai.com) have been tuned and improved over time with user feedback, as part of a model training process known as “reinforcement learning”. So for example, when users request things like python code, and try it out, users provide feedback on whether or not it worked. This feedback helps the model get better at generating code that is more likely to work.

2.3.3 When best to use ChatGPT:

- Idea Generation: ChatGPT can help brainstorm ideas for projects, essays, or creative endeavours. For example, you can ask it for ideas on new business ventures, book plots, or marketing campaigns.
- Thesaurus and Rephrasing: It's excellent for finding synonyms, rephrasing sentences, or improving the clarity of your writing. You can ask it to provide alternative words or rewrite a paragraph for better readability.
- Creative Writing: ChatGPT can assist with creative writing tasks, including writing short stories, poems, or generating dialogues for characters.
- Some Learning and Explanation: It can provide explanations and answers to questions on a wide range of very general topics, making it useful for learning new concepts or understanding complex ideas. But again, always best to double check against other trusted resources.
- Language Translation: ChatGPT can offer some basic translation assistance, such as translating phrases or short sentences between languages.
- Generating code snippets: It may not produce highly complex or critical code reliably but you can use it as a starting point for some common coding tasks.
- Querying structured data: ChatGPT can provide outputs in the form of data structures including for instance SQL queries and JSON. So for instance if we have an SQL database with information about the content of a corpus, we can prompt ChatGPT to output queries that would apply to that database. We can then begin to imagine an interface where a non-expert user could make such a query using their own natural language,

and ChatGPT could provide a connection with the structured data. This is sometimes referred to as “function calling” in the ChatGPT literature.

2.4

2.5 Large Language Models (LLM) and ChatGPT Uses in Libraries Today

[Need to work on this section with written examples/more narrative around opportunities and challenges]

- [A Tech Librarian Explains How to Build AI Literacy - Choice 360](#)
- [ChatGPT: Implications for academic libraries | Cox | College & Research Libraries News \(acrl.org\)](#)
- [Using GPT on Library Collections](#) (Matt Miller)
- <https://programminghistorian.org/en/lessons/interrogating-national-narrative-gpt>

2.6 Hands On Activity

For this activity we'll primarily use the freely available [ChatGPT \(openai.com\)](#) interface. It's also possible for those who may have more advanced skills to investigate using the API directly though it's not free.

2.6.1 Writing prompts

To get a useful response from ChatGPT, “prompting” is key. If you only ask a simple question, you may not be happy with the results and decide to dismiss the technology too quickly, but today's purpose is to have a deeper play in order to develop our critical thinking and information evaluation skills, allowing us to make informed decisions about utilising tools like ChatGPT in our endeavours.

Some best practices for writing a good prompt (helpful hints/tips gleaned from the readings above):

- Give the context. The more you explain the ins and outs of what you want ChatGPT to do, the more relevant your answer will be. For example, “You are an expert in [fill in the blank for your topic].”

- Be very specific and concrete about what you want it to do (and equally, what you don't want it to do). For instance you may want to specify word count, how many items you'd like in a list, tone of voice, etc.
- Give clear instructions on what the output should look like, and be simple and concise in your language.
- Provide examples. Much like a human, ChatGPT often understands best when you provide an example for it to work with, especially if you want it to produce something that closely resembles something you've seen online.
- Ask ChatGPT to "act like" someone: Sometimes, it can help to start your prompt by asking ChatGPT to "act like" a certain type of person that would be particularly well-placed to respond to the query you're posing.

If you're not satisfied with any part of the answer, you can keep building on it, asking ChatGPT to revise it in any way you like as it will continue on from previous enquiries in a session. (If you'd rather it not do that you can add "Please ignore all previous instructions" to start a prompt fresh). To learn more about prompting, there is a free online course: [Learn Prompting](#).

2.7 Exploratory activities:

- Have a play generating responses to some of the below questions (or come up with your own)!
- Critically evaluate the responses you receive from ChatGPT.
 - Is the information/response credible?
 - Are there any biases in the responses?
 - Does the information align with what you know from other sources?
- Consider your findings, the strengths and weaknesses, ethical considerations and challenges of using AI tools such as this.

Try this....	Example Prompts (please share if you have come up with a useful one!)	Further considerations/Notes
Ask ChatGPT to create a title for a new exhibition		

Try this....	Example Prompts (please share if you have come up with a useful one!)	Further considerations/Notes
Ask ChatGPT to create an exhibition label for items in an exhibition	“write an exhibition label for this archived website http://moorfootrunners.blogspot.co.uk/”	
Ask ChatGPT to create a blog post on a topic for which you are very familiar		
Ask ChatGPT to help you come up with ideas for teaching	<p>Try something like this:</p> <p><i>“You are an academic librarian focusing on instruction. Please make an outline of important topics to cover for a classroom lecture on ChatGPT. This will be part of a series of instruction sessions about information literacy. You have already worked with this class on the basic concepts of information literacy. Create the outline as a bulleted list and follow it with a classroom activity based on one of the topics you’ve discussed about ChatGPT.”</i></p> <p>From: A Tech Librarian Explains How to Build AI Literacy - Choice 360</p>	<p>When you get the results, you can then ask it to give you an additional activity or set of activities on the same topic. Have a play around with different training subject areas.</p>
Ask ChatGPT to write a bio for yourself, providing it with details for instance as a list of facts about you		

Try this....	Example Prompts (please share if you have come up with a useful one!)	Further considerations/Notes
Ask ChatGPT to Act as a Translator	<p>Try something like this:</p> <p><i>"I want you to act as an English translator, spelling corrector and improver. I will speak to you in any language and you will detect the language, translate it and answer in the corrected and improved version of my text, in English. I want you to replace my simplified A0-level words and sentences with more beautiful and elegant, upper level English words and sentences. Keep the meaning same, but make them more literary. I want you to only reply the correction, the improvements and nothing else, do not write explanations. My first sentence is "istanbulu cok seviyom burada olmak cok guzel"</i></p> <p>From: Awesome ChatGPT Prompts This repo includes ChatGPT prompt curation to use ChatGPT better.</p>	
Ask ChatGPT to create an outline for a PowerPoint presentation		
Ask ChatGPT to extract metadata from a provided text		

Try this....	Example Prompts (please share if you have come up with a useful one!)	Further considerations/Notes
Ask ChatGPT to write a query for a suppositional SQL database that you describe	<p><i>“We have an SQL database with the following tables: 'books' and 'authors'. The 'books' table has columns 'title', 'date', and 'topic'. The 'authors' table has columns 'name', 'birth_date', 'death_date', and 'gender'. I'm going to write some questions about things in this database, and I would like you to respond with SQL queries that would find the information I'm asking about.”</i></p> <p><i>RESPONSE: “Of course! Feel free to ask any questions you have about the database, and I'll provide you with SQL queries to retrieve the information you need. Please go ahead and ask your questions.”</i></p> <p><i>INPUT: “I'd like to find books about the second world war written by women born after the war ended.”</i></p> <p><i>RESPONSE: “To find books about the Second World War written by women born after the war ended, you can use the following SQL query:”</i></p> <pre> SELECT b.title, b.date, a.name FROM books b JOIN authors a ON b.author_id = a.author_id WHERE b.topic = 'Second World War' AND YEAR(a.birth_date) > 1945 AND a.gender = 'female'; </pre>	

Try this....	Example Prompts (please share if you have come up with a useful one!)	Further considerations/Notes
Have a look at this project Matt Miller did experimenting with [Using GPT on Library Collections (thisismattmiller.com)] .	<p><i>“You are a helpful assistant that is summarizing and extracting data from a journal written by Susan B. Anthony in {RECORD_DATE_HERE}. You only answer using the text given to you. You do not make-up additional information, the answer has to be contained in the text provided to you. Each page is a diary entry or financial bookkeeping. You will structure your answer in valid JSON, if there are any quote marks in the content escape them as ”.”</i></p> <p><i>“If the following text contains multiple journal entries, extract each one into an array of valid JSON dictionaries. Each dictionary represents one of the entries, extract the date and the date again in the format yyyy-mm-dd and the city or state it was written in and other geographical locations mentioned that entry and people mentioned that entry and the complete full text of the entry and a one sentence summary of the text, using the JSON keys dateText, dateFormatted, cityOrState, geographicalLocations, people, fullText, summaryText: {FULL_TEXT_HERE}“</i></p>	Thinking about this project, could you think of a similar collection that you might be able to experiment with to parse data similarly? How might you like to re-write the prompt for your own project?

Try this....	Example Prompts (please share if you have come up with a useful one!)	Further considerations/Notes
Try to use it to create social media posts	<i>There is a new network called CHOSN for those working in Galleries-Libraries-Archives-Museums (GLAM) and cultural heritage organisations as research support and research-active staff. This network is a community of practice for research support and research-active staff who work in GLAM organisations interested in developing and sharing open scholarship knowledge and skills, organising events, and supporting each other in this area. Can you write social media post for 140 character to announce this network and call people to join in?</i>	

Try this....	Example Prompts (please share if you have come up with a useful one!)	Further considerations/Notes
PROMPT: Act as a librarian advising a history student in the UK doing research into the third plague pandemic in India during the early 20th century. What archival resources could they use?	<p><i>RESPONSE: Certainly! If you're a history student in the UK researching the Third Plague Pandemic in India during the early 20th century, there are several archival resources you can explore to enhance your research. These archives will provide you with primary sources, documents, and historical records that can help you gain a deeper understanding of this important historical event. Here are some key archival resources to consider:</i></p> <ol style="list-style-type: none"> <i>**The National Archives (TNA) of the United Kingdom**:</i> <ul style="list-style-type: none"> <i>- TNA holds records related to British colonial administration in India during the early 20th century. Look for documents related to public health, sanitation, and official correspondence about the plague.</i> <i>**The India Office Records**:</i> <ul style="list-style-type: none"> <i>- These records are part of the British Library's holdings and contain a wealth of information on India during the colonial period. You can find reports, letters, and administrative documents related to the Third Plague Pandemic.</i> <i>**The British Library**:</i> <ul style="list-style-type: none"> <i>- The British Library has a substantial collection of materials related to India, including newspapers, periodicals, and publications from that era. You can find newspaper articles, books, and reports that discuss the plague's impact in India.</i> <i>**The Wellcome Collection**:</i> 	

Try this....	Example Prompts (please share if you have come up with a useful one!)	Further considerations/Notes
PROMPT: Act as an archivist producing a catalogue record for an archival document. What subject, place and name authorities or terms would you choose for the following text? [gave it c50% of the full text transcription of IO R/F/4/1754/71645]	<p><i>RESPONSE: Here's a suggested catalog record for the provided archival document, along with subject, place, and name authorities/terms:</i></p> <p><i>Title: Development of Hemp Cultivation for Cordage Manufacturing</i></p> <p><i>Author: [Author's Name if available]</i></p> <p><i>Date: [Date of the Document, if available]</i></p> <p><i>Description:</i></p> <p><i>This archival document explores the historical development of hemp cultivation for cordage (rope) production, emphasizing its critical importance for mechanical purposes and ship rigging during the 19th century. It discusses the transition from skin and leather ropes to vegetable fiber, particularly hemp, and the widespread cultivation and commerce of hemp in various nations. The document highlights the vital role of hemp during wartime and its impact on British shipping during supply disruptions. Additionally, it underscores the attention given to hemp cultivation in India and its various applications, focusing on the Cannabis Sativa plant.</i></p> <p><i>Subjects:</i></p> <p><i>Cordage Industry</i></p> <p><i>Hemp Cultivation</i></p> <p><i>Rope Manufacturing</i></p> <p><i>Ship Rigging</i></p> <p><i>Hemp Fiber</i> 17</p> <p><i>Agriculture in India</i></p> <p><i>Colonial Agriculture</i></p> <p><i>International Commerce</i></p> <p><i>War Supply Chains</i></p> <p><i>Mechanical Engineering</i></p> <p><i>Fiber Plants</i></p> <p><i>Places:</i></p>	<p>COMMENT: The inclusion of Yersin as a suggested authority is interesting / worrying, as he had cropped up in an answer to a previous question, and I had told it to disregard all of the previous conversation data before posting the cataloguing prompt...</p> <p>In terms of indexing, it included a lot of incidental references in the text which a cataloguer would not include as subject or place index terms [should flag at this point that I did not write the BL's catalogue entry for this]]</p> <p>I should really give it the full text before judging the description...</p>

Try this....	Example Prompts (please share if you have come up with a useful one!)	Further considerations/Notes
--------------	---	------------------------------

2.8 Recommended Reading/Further Learning

- [40 Best ChatGPT Prompts You Simply Have To Try in 2023 \(tech.co\)](#)
- [A Tech Librarian Explains How to Build AI Literacy - Choice 360](#)
- [ChatGPT: Implications for academic libraries | Cox | College & Research Libraries News \(acrl.org\)](#)
- [Using GPT on Library Collections \(Matt Miller\)](#)
- <https://programminghistorian.org/en/lessons/interrogating-national-narrative-gpt>
- [Welcome | Learn Prompting: Your Guide to Communicating with AI](#)

Guide Element	Description
---------------	-------------

2.8.1 **Bas

ic Guide Metadata**

Topic title:

Understanding Large Language Models through an introduction to ChatGPT

Topic summary:

A brief introduction to Large Language models (LLMS) by way of ChatGPT, this guide will demystify both through hands-on activities, and explore potential challenges and opportunities for current and future library services.

Author(s):

Nora McGregor

Date first published:

16/10/2023

Date last

reviewed/Reviewed By:

Date last modified:

DOI:

How to cite:

Learner/Skill level:

Easy

Topics covered

Machine learning and AI in Libraries literacies (From [controlled category tags](#))

Guide Element	Description
2.8.2 **Gui	
de Content Blocks** Learning objectives	<p><i>This guide aims to:</i></p> <ul style="list-style-type: none"> • <i>Introduce [topic] and its importance in the context of library work</i> • <i>Outline how these technologies are being used in libraries</i> • <i>Showcase different tools & technologies relevant to [topic] through hands-on activities</i> • <i>Demonstrate how [topic] can be used for [outcome]</i>
Pre-requisites & set-up:	No previous knowledge is required. You'll need a browser and to create a login to chatgpt. ChatGPT (openai. com)
I: Introduction to the topic	
II: Relevance to the Library Sector	
III: Hands-on activity and other self-guided tutorial(s)	
IV: Recommended Reading & Further Learning	<p>This section should include reference to:</p> <ul style="list-style-type: none"> • Recommended reading • Networks/Organisations? • Advanced tutorials/courses/learning?

References

Knuth, Donald E. 1984. “Literate Programming.” *Comput. J.* 27 (2): 97–111. <https://doi.org/10.1093/comjnl/27.2.97>.