Beyond Basic Chatbots: Enhancing Website Search with LLMs

Mark Lee

Director of Web User Experience & Digital Strategy Colorado College

github.com/leemark/beyond-chatbots



tl;dr

• Enhance, Don't Replace

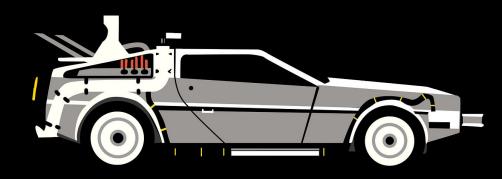
LLMs: A Do-Anything Multitool

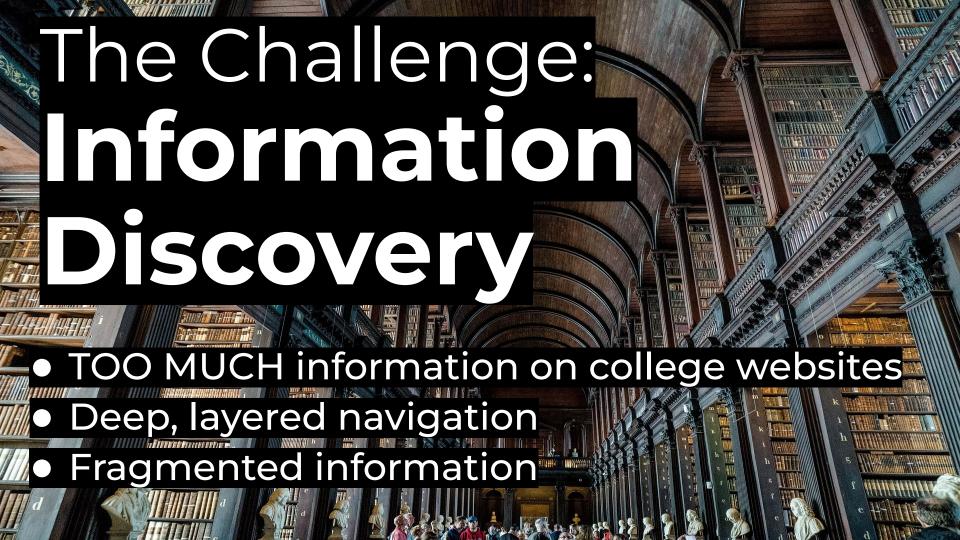
Embrace the Unexpected

Where we're going, we don't need roads!

(but it's nice to have a roadmap anyway)

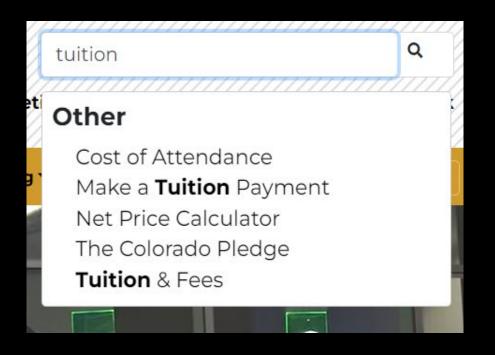
- 1) Challenges with web search
- 2) LLM chatbots and their limitations
- 3) Basic to advanced prompting
- 4) Advanced LLM usage
- 5) A RAG use case



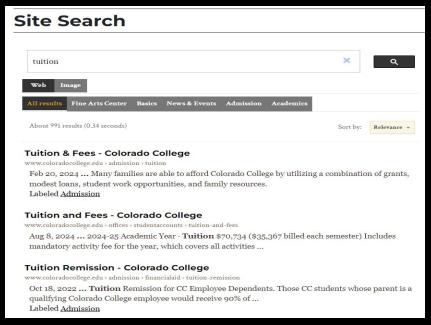


Current Search Tools

QuickSearch:Simple keyword search



Google Custom Search: Algorithmic; 10 blue links



Large Language Models (LLMs)

- Al that understands and generates human-like text
- Trained on massive datasets
- Uses deep learning and neural networks.
 Generative pre-trained transformers.

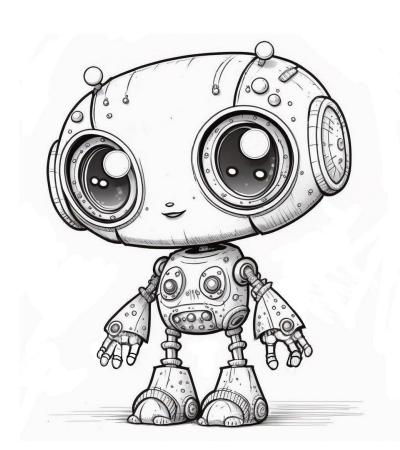
What Can LLMs Do Well?

- Translation
- Summarization

- Pattern recognition
- Question-answering

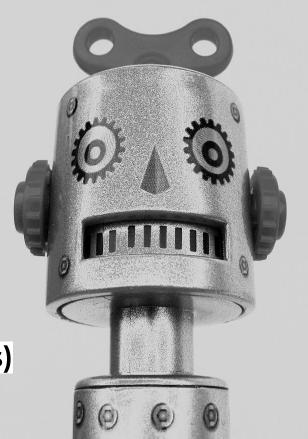
Some Advantages of LLM Chatbots

- Natural language queries
- Diverse language tasks
- Human-like responses
- Continuous improvement



Some **Limitations** of LLM Chatbots

- Potential for "hallucinations"
- Bias in training data
- Security and privacy concerns
- Limited to training data cutoff
- Limited Context window (sometimes)



LLMs in Action

How much is tuition at Colorado College?

You're giving feedback on a new version of ChatGPT.

Which response do you prefer? Responses may take a moment to load.



Response 1

For the 2023-2024 academic year, the tuition at Colorado College is \$68,634. This amount does not include additional costs such as room and board, fees, books, and personal expenses, which can add several thousand dollars to the total cost of attendance.



Response 2

As of the 2023-2024 academic year, the tuition at Colorado College is \$70,692 per year for undergraduate students. This does not include additional costs such as room and board, fees, books, and personal expenses, which can bring the total annual cost to approximately



Tuition & Fees - Colorado College

www.coloradocollege.edu > admission > tuition

Feb 20, 2024 ... Many families are able to afford Colorado College by utilizing a combination of grants, modest loans, student work opportunities, and family resources. Labeled Admission

Tuition and Fees - Colorado College

www.coloradocollege.edu > offices > studentaccounts > tuition-and-fees

Aug 8, 2024 ... 2024-25 Academic Year · Tuition \$70,734 (\$35,367 billed each semester) Includes mandatory activity fee for the year, which covers all activities ...

Tuition and Fees - Colorado College

www.coloradocollege.edu > offices > CC-students > tuition_and_fees

Apr 29, 2024 ... Block C courses will be entirely remote, and limited student housing will be available. Housing and meal plans rates are subject to change.

Cost of Attendance - Colorado College

www.coloradocollege.edu > admission > handbook > cost-of-attendance

Feb 20, 2024 ... CC Financial Aid & You: Partnership for Success. Many families are able to afford Colorado College by utilizing a combination of grants, modest ... Labeled Admission

Tuition and Fees - Colorado College

www.coloradocollege.edu > pre-college > precollegetuitionindex

Need-Based Scholarship. As part of your online course application, you will need to submit a 500 word essay on "The Importance of a Liberal Arts Education in

How much is tuition at Colorado College?

Sources

Colorado College - Net Price, Tuition, Cost to Attend, Financial Aid ... collegesimply · 1

The Colorado College -College Confidential

Tuition & Fees -Colorado College **企** 台 ==

collegeconfide... · 2

cc coloradocollege · 3

View 5 more



The tuition at Colorado College for the 2024-25 academic year is \$70,224 3 5. This amount covers full-time enrollment for the entire academic year, which consists of two semesters.

Some key points about Colorado College's tuition and costs:

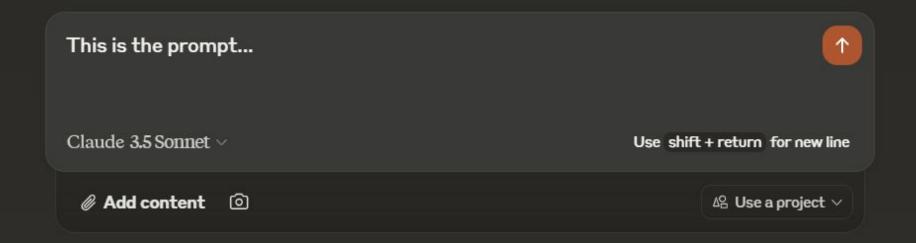
- The \$70,224 tuition includes a mandatory student activity fee of \$510 3.
- Tuition is the same for both in-state and out-of-state students, as Colorado College is a private institution 2.
- The total estimated cost of attendance for 2024-25, including tuition, fees, room and board, books, and other expenses, is \$91,154 3.



Basic prompting

The art of crafting input text that elicits a specific response or behavior from the LLM.

***** Good evening, Mark

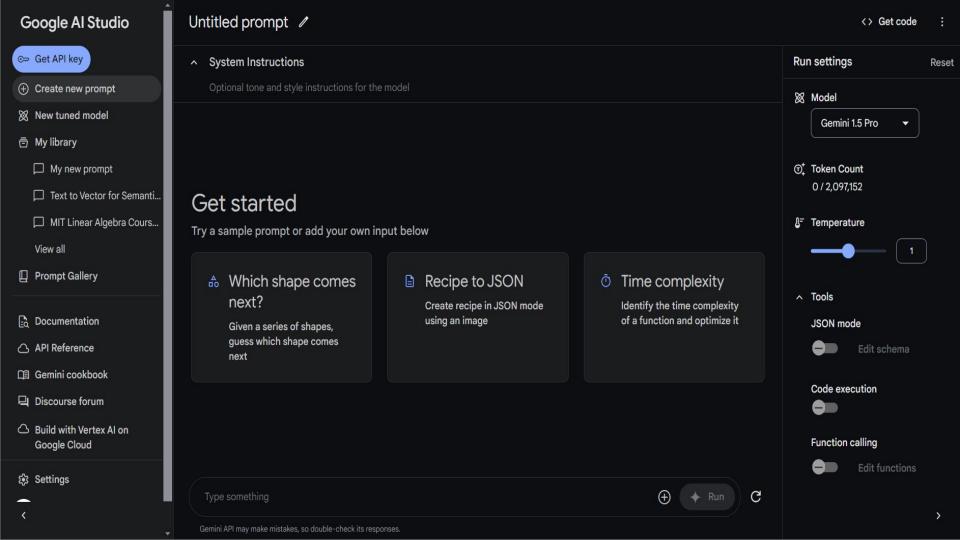


Further Improving LLM Performance

- Better prompting techniques
- "Think out loud" approach, let the LLM think
- Leverage speed and context windows
- Use the right tool for the job
- Let's start to think about how we can build a framework around the LLM

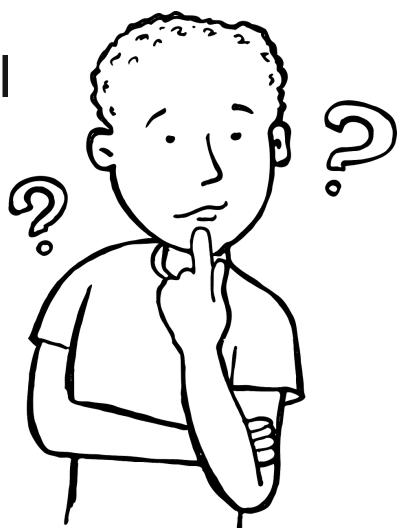
Let's look under the hood a bit more...

- Chat interface vs developer playground
- OpenAl, Anthropic, Google Al Studio, others
- Prompt library of various examples
- Choose different models
- Set Temperature
- Set your own System Prompt
- Test, iterate, then get the API code



Choosing a model

- Open vs closed
- Cost vs intelligence
- Cost and speed
- Intelligence vs speed



Turning up the Temperature

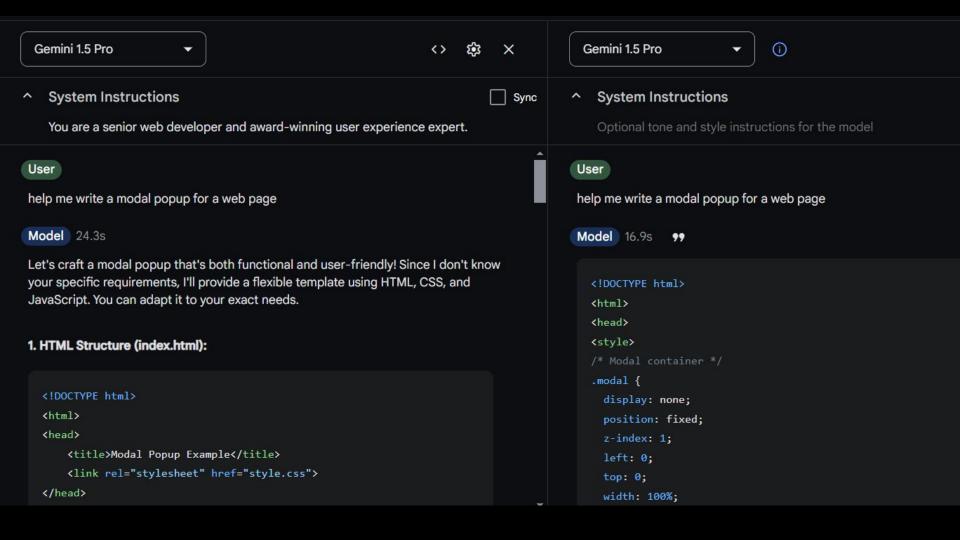
Temperature Controls Randomness

Higher values = more creative, varied responses

Lower values = more focused, deterministic outputs

System prompt vs. user prompt

"You are a helpful assistant..."



Advanced Prompting Techniques

- Chain-of-thought prompting
- Few-shot or many-shot (in-context) learning
- Pre-fill/partial response
- <scratchpad>
- Reflection

Chain-of-Thought Prompting

- Encourages step-by-step reasoning
- Improves accuracy for complex tasks
- Example: "Think through this problem step-by-step..."

Prompting Techniques: From Zero to Many

Zero-shot = No examples provided

Classify the sentiment of the following sentence: 'The product exceeded my expectations.'

Few-shot = 1-5 examples

```
Classify the sentiment of the following sentences as positive, neutral, or negative:
1. 'The customer service was terrible.' -> Negative
2. 'The meal was okay, nothing special.' -> Neutral
3. 'The product exceeded my expectations.' ->
```

Many-shot = 10+ examples

Pre-filled Response

- Start the response for the Al
- Guides the structure and content

Example:

The three main dining locations on campus are: (1)

The <scratchpad> Technique

Ask the AI to show its work, for example in <scratchpad> XML tags, before composing its answer. Improves transparency and accuracy.

Key points supporting AI achieving consciousness: 1. Increasing complexity of AI systems mimicking brain structure 2. Advancements in neural network architectures 3. Growing computational power approaching brain-level processing 4. Emergence theory - consciousness as an emergent property of complex systems 5. Gradual development of self-awareness in AI systems 6. Potential for AI to model its own thought processes 7. Information Integration Theory of consciousness 8. Functionalist arguments - consciousness arising from functional organization 9. Evolutionary perspective on consciousness development 10. Technological singularity leading to super intelligent AI 11. Quantum computing potentially enabling consciousness-like states 12. Philosophical arguments for multiple realizability of mental states

13. Advancements in understanding and replicating neurobiological processes

15. Machine learning systems showing signs of creativity and intuition

14. Potential for AI to develop subjective experiences or qualia

</scratchpad>

Reflection: Self-Improvement for Al



A proposed methodology for better Web search

R.A.G.

Retrieval-Augmented Generation

Why Use RAG?

Contextual, natural language answers

References and citations

- Follow-up questions for deepened engagement
- Combines LLM capabilities with up-to-date info

Ask Colorado College 🐯

Enter your question:

Where do first year students live on campus?

Search

Answer: First-year students at Colorado College are housed in the "Big 3" traditional halls: Mathias Hall, Loomis Hall, and South Hall [1].

Related questions:

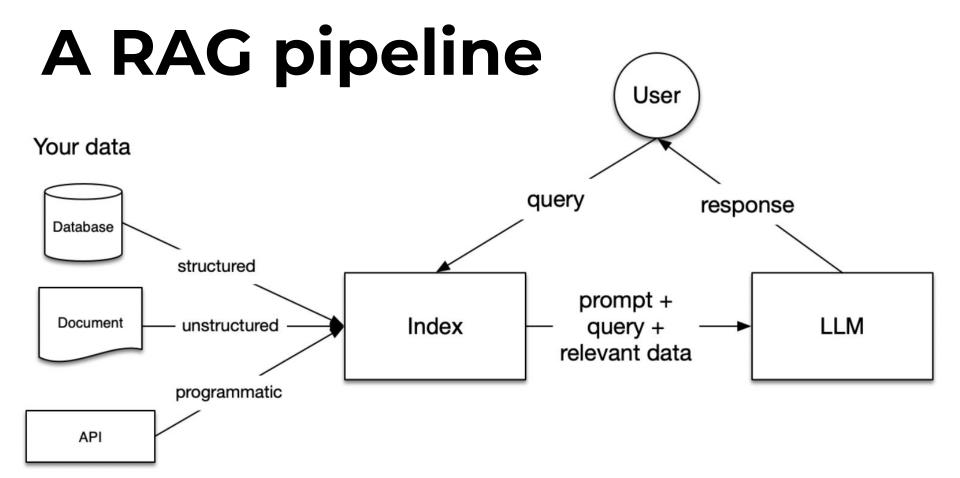
What are the amenities like in the Big 3 residence halls?

Are there single rooms available in the first-year residence halls?

What is the social atmosphere like in the first-year dorms?

Sources used:

[1] https://www.coloradocollege.edu/offices/housing-and-conferences/livingatcc.page/first-year.page.html



Use an LLM to evaluate & rewrite the user query (1 or *n* variations)

Search information sources (search API: Google, Brave, DDG)

Grab and parse the contents of resultant pages

Store results in vector database (optionally*, for semantic search)

Retrieve relevant info (optional, for semantic search)

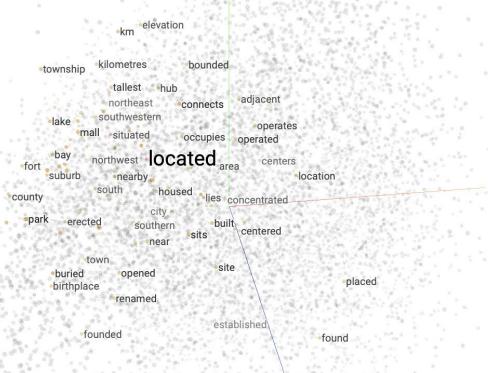
Pass original user query plus relevant context to LLM to generate the answer (don't forget to use our advanced prompting techniques)

Use LLM to suggest follow-up questions

Send all to user: generated answer, citations, and follow-up questions

Embedding Models & Vector Databases

- Embed content into high-dimensional space
- Semantic similarity search
- Efficient retrieval of relevant information



```
def main():
   # Step 1: Rephrase the user's query
                                                                  How to build
   rephrased query = rephrase query(USER PROMPT)
   print(f"Rephrased Query: {rephrased query}")
                                                                      rudimentary
   # Step 2: Search for relevant URLs
   urls = search urls(rephrased query)
                                                                      RAG-based search
   # Step 3: Scrape and parse content from the URLs
   contents, sources = scrape and parse(urls)
                                                                  in Python
   # Step 4: Embed the scraped content into the Chroma database
   embed into chroma db(contents, sources, "colorado college tuition")
   # Step 5: Query the Chroma database for relevant information
   relevant docs, relevant sources = query chroma db(rephrased query, "colorado college tuition")
   # Step 6: Prepare the context for the response generation
   context = "\n".join([f"Source: {source}\nContent: {doc}" for doc, source in zip(relevant docs, relevant sources)])
   # Step 7: Generate the final response
   response = generate response(context, rephrased query)
   # Step 8: Print the results
   print("Response:")
   print(response)
   print("\nRelevant Sources:")
   print("\n".join(relevant sources))
```

Rephrasing user queries

```
{"role": "system", "content": "You are a search assistant. Rephrase the following as a concise search query expertly crafted to retrieve relevant results from a search engine. Return only the query, no explanation:"},

{"role": "user", "content": user_query_goes_here}
```

Searching for relevant URLs

```
url = f"https://api.search.brave.com/res/v1/web/search?q={query}"
headers = {"X-Subscription-Token": BRAVE API KEY}
response = requests.get(url, headers=headers)
results = response.json()
```

Scraping & parsing content

```
response = requests.get(url, headers=headers)
soup = BeautifulSoup(response.content, 'html.parser')
contents.append(soup.get_text())
```

sources.append(url)

Embedding content into a vector database

```
client = chromadb.PersistentClient()
collection = client.get_or_create_collection(collection name)
for content, source in zip(contents, sources):
     collection.add(
         documents=[content],
         metadatas=[{"source": source}],
         ids=[str(uuid.uuid4())]
```

Querying the database for relevant information

```
results = collection.query(
        query_texts=[query],
        n_results=top_k,
        include=["documents", "metadatas"]
)
```

Generating the response

prompt = f"""You are a helpful assistant that answers
questions about Colorado College only based on
information from the CC website, provided below.

Question: {question}

Context: {context}

Instructions:

- 1. If you have enough information to answer the question confidently and accurately, provide a direct answer.
- 2. If you don't have enough information to answer the question appropriately then reply with a brief ...

Ask Colorado College 🐯

Enter your question:

Where do first year students live on campus?

Search

Answer: First-year students at Colorado College are housed in the "Big 3" traditional halls: Mathias Hall, Loomis Hall, and South Hall [1].

Related questions:

What are the amenities like in the Big 3 residence halls?

Are there single rooms available in the first-year residence halls?

What is the social atmosphere like in the first-year dorms?

Sources used:

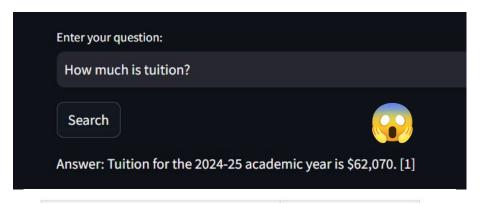
[1] https://www.coloradocollege.edu/offices/housing-and-conferences/livingatcc.page/first-year.page.html

Thoughts after building this.

The first 80% is easy, the last 5% of testing and refinement is where the real challenge and the real learning is.

Beyond search, it hints at new, methods for web content Quality Assurance.

Consider *also* having a very well-vetted information source in the pipeline/index.



Tuition and Fees:	\$62,070.00
Books and Supplies:	\$1,240.00
Food:	\$1,188.00

Ask Colorado College 😈

Enter your question:

¿Dónde puedo comer en el campus?

Search

Answer: En el campus, puedes comer en Rastall Dining Hall o Benjamin's, ambos ubicados en el Worner Campus Center. [1] Rastall ofrece un buffet con una variedad de opciones, incluyendo comida tradicional, vegana y vegetariana, mientras que Benjamin's tiene una parrilla con hamburguesas y la opción de crear tu propio plato de fideos.

Related questions:

¿Qué tipo de comida se sirve en Rastall Dining Hall?

¿Hay opciones sin gluten disponibles en el campus?

The Future (or present?) of Al-Powered Web Experiences

- Multimodal search (text, image, audio)
- Personalization
- Conversational search
- Integration with other campus systems
- Autonomous Agents

tl;drevisited

Enhance, Don't Replace

LLMs: Your Al Multitool

Embrace the Unexpected

Enhance, Don't Replace

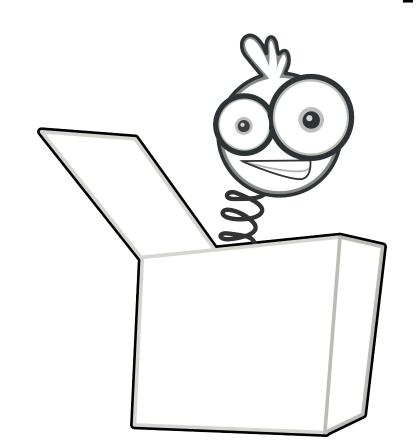
choose

Evolution over Revolution

A Do-Anything Multitool



Embrace the Unexpected



A few final thoughts

- LLMs CAN offer powerful information search capabilities
 - But ONLY with "fact checking" and accurate info
- RAG combines the power of LLMs with up-to-date info from traditional search
- Keep learning, keep exploring.
 Things are changing fast.

Thank You!

Questions?

Resources, code examples, slides, etc. github.com/leemark/beyond-chatbots

