Web Crawler

After gathering information about ChatGPT on the web, I determined that the most important terms from my search were:

```
["chatgpt", "openai", "conversation", "robotics", "limitations", "chat",
"math", "science", "math", "write"]
```

I chose to build a simple knowledge base. I built a Python dictionary with the key being an important term and the value being a list of any phrase that contains the important term. The Python dictionary was pickled into a file for convenience in the future. I included a screenshot of my knowledge base below…

```
{'chatgpt': ['chatgpt wikipedia jump content toggle sidebar search create accountlog personal tools create account log pages logged editors learn contributionstalk
navigation main pagecontentscurrent eventsrandom articleabout wikipediacontact usdonate contribute helplearn editcommunity portalrecent changesupload file tools links
hererelated changesupload filespecial pagespermanent linkpage informationcite pagewikidata item download pdfprintable version projects wikimedia commons languages
wikipedia language links top page across article title', 'contents move sidebar hide top limitations toggle features limitations subsection toggle service subsection
service service developer support toggle reception subsection toggle implications subsection cybersecurity academia concerns toggle ethical concerns subsection data
bias media also links toggle table contents toggle table contents chatgpt languages bokmålnorsk nynorsko'zbekcha simipусскийsimple uyghurchetiếng việt文言吴语粤語中文 edit
links articletalk english readview sourceview history readview sourceview history wikipedia free encyclopedia artificial intelligence chatbot developed openai
chatgptlogodeveloper openaiinitial releasenovember months ago stable releasefebruary days ago chatgpt artificial intelligence chatbot developed openai launched
november', 'chatgpt launched prototype november quickly garnered attention detailed responses articulate answers across many domains knowledge', 'following release
chatgpt openai valuation estimated us billion', 'training chatgpt member generative transformer gpt family language models', 'addition openai continues gather data
chatgpt users could used train chatgpt', 'users upvote downvote responses receive chatgpt fill text field additional feedback', 'features limitations features chatgpt
asked question jimmy wales killed tiananmen square protests', 'chatgpt correctly answers incorrectly gives wales age time instead', 'although core function chatbot
mimic human conversationalist chatgpt versatile', 'chatgpt training data includes man pages information internet phenomena programming languages bulletin board systems
python programming language', 'comparison predecessor instructgpt chatgpt attempts reduce harmful deceitful responses', 'one example whereas instructgpt accepts premise
prompt tell christopher columbus came truthful chatgpt acknowledges counterfactual nature question frames answer hypothetical consideration might happen columbus came
using information voyages christopher columbus facts modern world including modern perceptions columbus actions', 'unlike chatbots chatgpt remembers previous prompts
given conversation journalists suggested allow chatgpt used personalized therapist', 'prevent offensive outputs presented produced chatgpt queries filtered openai
```

With this knowledge base, I hope to create a chatbot that answers basic questions about ChatGPT.
A sample dialog is:

Q: What is ChatGPT?
A: ChatGPT is an artificial intelligence chatbot developed by the company OpenAI.