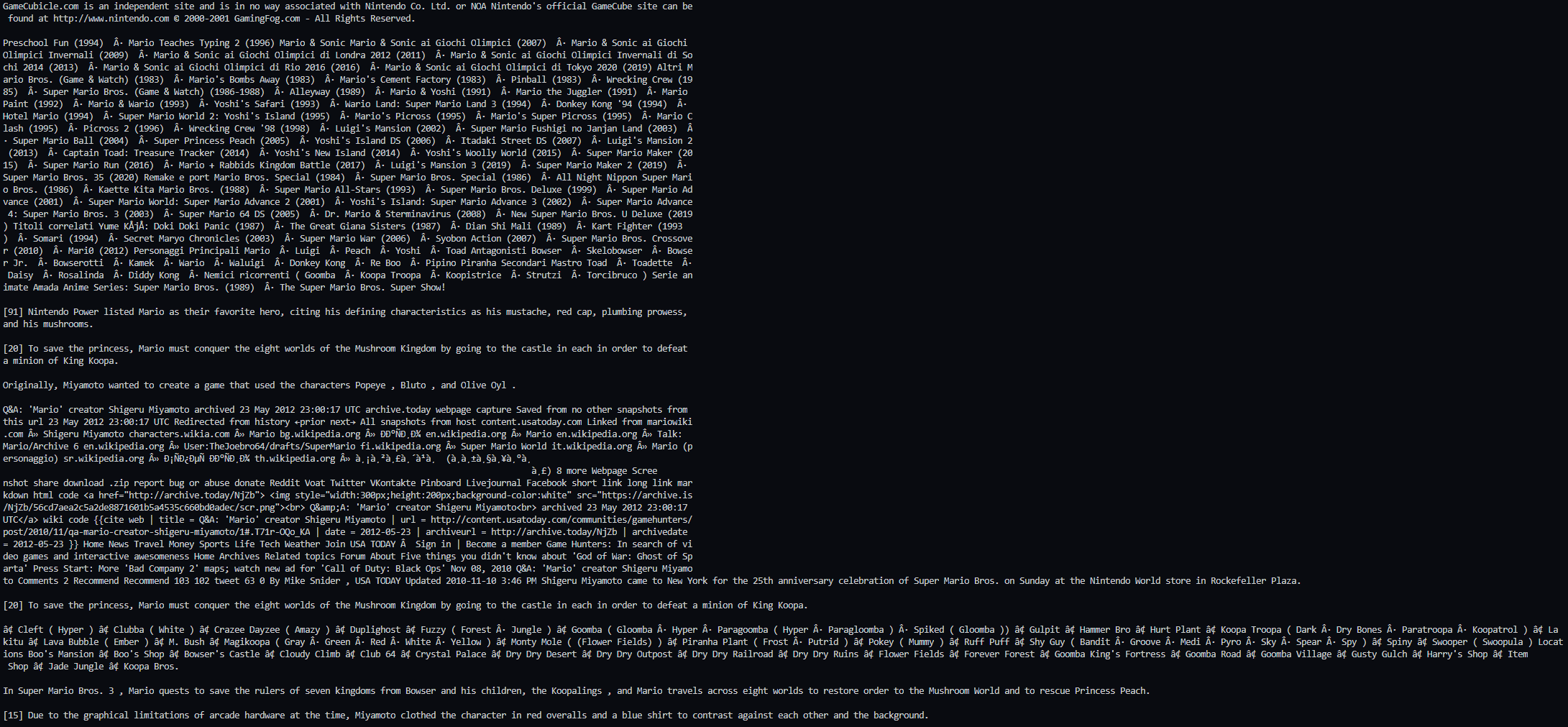
HLT – Fall 2020

Homework 5: Web Scraping

1. describe how you created your knowledge base, include screen shots of the knowledge base, and indicate your top 10 terms;

I created my knowledge base in a very simple way: for every term in my top 10 terms, and for every url, if there was a sentence containing the term, I would add that sentence to a dictionary such that the term was the key and the value was a list of sentences.

Knowledge base:



This screenshot includes item 0 in the list of sentences for each of the 10 terms. As you can see, some text is not legible; in future builds I may modify it to exclude any sentences with characters not found in standard written English (i.e. the character ‘Â’)

For my top 10 terms, I used the following:

top10 = ['nintendo', 'bowser', 'mushroom', 'princess', 'characters', 'miyamoto', 'castle', 'goomba', 'kingdom', 'arcade']  
If I had more terms available, I would add other terms such as ‘enemies’, ‘lakitu’, or ‘waluigi’ but these are generally covered in the top 10.

1. write up a sample dialog you would like to create with a chatbot based on your knowledge base

Dialogue1:

Hi, I’m Mariknow, a chatbot that knows about Mario. What would you like to know?   
 What do you know about mario?

I know about things like his franchise, characters in his games, and his creator.   
 Tell me about his creator.

Shigeru Miyamoto was Nintendo’s first artist. He created Mario in 1981.   
 Was he always called mario?

The playable character, initially "Jumpman", was eventually named for Mario Segale, the warehouse landlord.  
 That’s interesting.

Dialogue 2:

Hi, I’m Mariknow, a chatbot that knows about Mario. What would you like to know?   
 Who does Mario fight against?

In Super Mario Bros. 3 , Mario quests to save the rulers of seven kingdoms from Bowser and his children, the Koopalings , and Mario travels across eight worlds to restore order to the Mushroom World and to rescue Princess Peach.  
 What is a koopaling?

The Koopalings are a fictional group of seven Bowser Jr.'s clones in the Mario video game franchise by Nintendo. Have you played Mario?   
 Yes, I played the arcade game.

The arcade game! Due to the graphical limitations of arcade hardware at the time, Miyamoto clothed the character in red overalls and a blue shirt to contrast against each other and the background.  
 I didn’t know that.