

# Web Scraper Report

fed180001

## 1) Knowledge Base Creation

To create my knowledge base about Jeffrey Bezos, I first selected my top 10 terms: Jeff, Amazon, worth, billion, Blue, company, Washington, climate, CEO, announced. I felt that these ten terms would give me a variety of relevant information that a Jeffrey Bezos chatbot could answer. I know my information is recent because the programs filters out links that are from before 2020. I then created an “empty” dictionary called kb (aka knowledge base) where all my selected terms would act as keys, with empty strings acting as their current definitions. The code then iterates through all of my text files containing sentences I scraped from each site. It looks at each sentence and if one of the words was present in the sentences, the sentence is added to the dictionary value of that word. If multiple words are present it chooses the first word in the dictionary. Knowing this, I ordered the dictionary from least frequent to most frequent, so that more popular terms wouldn’t dominate the knowledge base. Finally to limit the size of my knowledge base for practicality, I limited each term's number of sentences to 75. To make the searchability of the knowledgebase easier, I ordered each term's sentences by length.

I allowed this knowledge base to be accessible from two different formats: a pickle file and an actual sqlite database. Both are formatted the same way as the original dictionary I created, with the database table “knowledge\_base” having two attributes: term and facts, which easily correlates from the dictionary mentioned above. Here is the structure pictured below

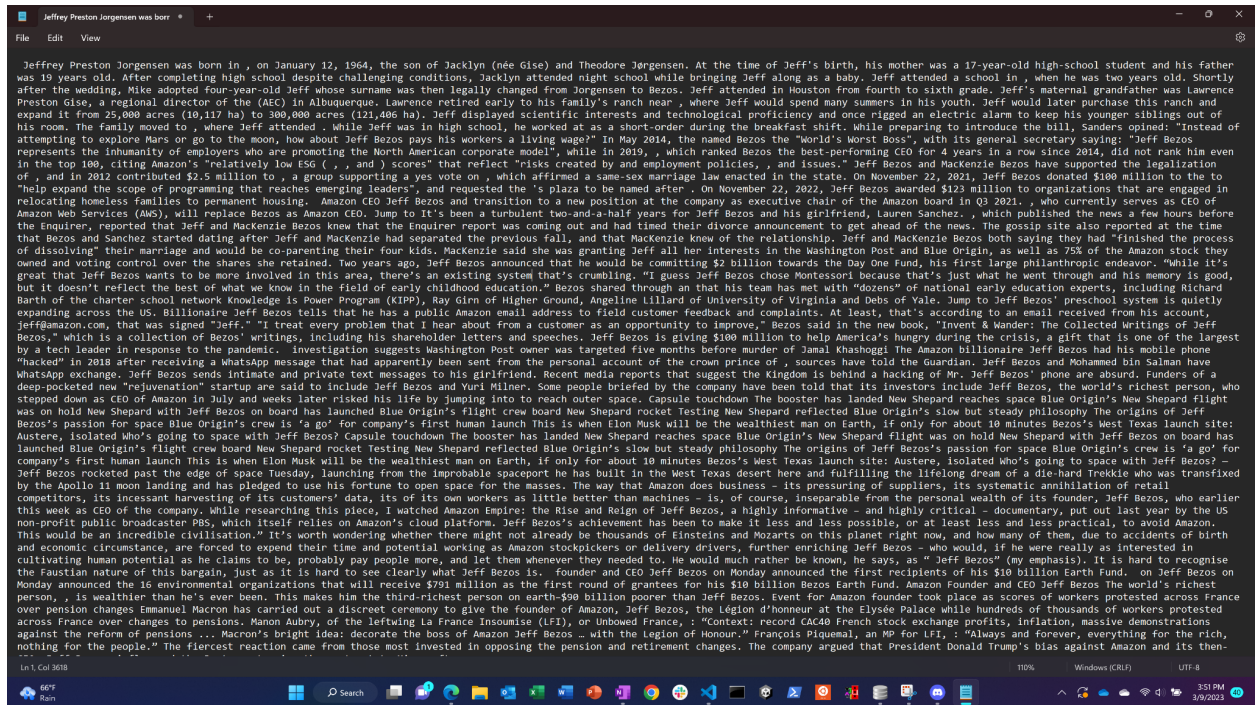
Database StructureBrowse DataEdit PragmasExecute SQL

Table: knowledge\_base

Filter in any column

	term	facts
	Filter	Filter
1	announced	On October 2, 2018, Bezos announced a company-wide wage increase, which Sanders applauded. In Februar...
2	Washington	He and his then-wife, , left their jobs at D. E. Shaw and founded Amazon in a rented garage in , Washington ...
3	CEO	Bezos owns 11% stake in Amazon, which has seen surge in demand with households on virtual lockdown The...
4	Blue	Blue Origin's vehicle in 2015 and afterwards successfully landed back on Earth; he flew into space on in 2021....
5	climate	Other recipients included pass-through organizations that make grants to smaller nonprofits working against ...
6	worth	According to the , Bezos is one of the few billionaires to have seen an increase to his net worth since the ...
7	Jeff	Jeffrey Preston Jorgensen was born in , on January 12, 1964, the son of Jacklyn (née Gise) and Theodore ...
8	company	Bezos owns an 11% stake in the company and has been the world’s richest person since 2017. The man, an ...
9	billion	According to the , Bezos is one of the few billionaires to have seen an increase to his net worth since the ...
10	Amazon	Bezos owns 11% stake in Amazon, which has seen surge in demand with households on virtual lockdown The...

Along with an example of all the sentences stored in the term “Jeff”



## 2) Sample Dialogue

Initial Dialog 1.

Hello, this is Jeffrey Bezos, jeff to my friends, its always day one! What’s your name amazonian?

I am Fanny

Nice to meet you Fanny, is there anything you would like to ask me, people always have tons of questions?

Yes actually, is it true you announced that you are no longer to pledge \$10 billion to charity ?

Yes, with the first \$791 million of it going to environmental agencies.

Why dont you donate more, you have an absurdly large net worth?

That’s true, I could always do more, but have you seen that I [recent event here]

Initial dialog 2

Hello, this is Jeffrey Bezos, jeff to my friends, its always day one! What’s your name amazonian?

I’m Fanny

Nice to meet you Fanny, is there anything you would like to ask me, people always have tons of questions?

Not really

Then I have one for you! Do you like Amazon as a company?

Not really, you guys treat your warehouse employees pretty poorly.

That may or may not be true, but check out all these cool things we did | random amazon tech facts|, does that not change your mind?

