

In our knowledge base, we produced a searchable dictionary of our 10 chosen words out of the 30 most recurring novel words. We decided upon: Lamborghini, sport, engine, vehicle, Volkswagen, manufacturer, business, Subsidiary, Ducati, and model.

We started off on Lamborghini's Wikipedia page. To begin forming our knowledge base, we scraped the text of that page for links and filtered out any irrelevant links, like image links, quick links to the home page, Wikipedia's protection policy, and so on. Once all of the relevant links have been gotten, the method `save_to_file` stores the URLs into their own file with an appropriate file name. Next, we scrape through each link file with the `scrape_list_of_urls` method. Thanks to the helper method `scrape`, our URLs should be clean of all HTML scopes and newlines from the HTML code. To get rid of newlines and other whitespaces present in our raw text files, the `create_clean_files` method creates new files of the processed texts. We then combine the content from all the cleaned files with the method `combine_lines_from_files` and send the list of lines to the `filter_text` method to convert the sentences to lowercase and remove all stop words and punctuation. To narrow down our term frequency, we lemmatize the text to obtain only nouns and verbs with the `lemmatize_tokens` method. In `print_term_frequency`, we create a dictionary where the lemmatized terms are the keys and the values are the count of how often they appear in the corpus. Once we have looped through every term, we divide their value by the total length of the lemmatized list to normalize it. We then sort the dictionary in descending order and print out the first 30 terms. We hand-picked 10 terms from those 30 top terms to serve as the bases of our knowledge base. The `create_related_sentences_dict` method goes through every line in the corpus and adds it to a dictionary where each lemma has a list of the sentences which contain that word, which is our knowledge base. The knowledge base dictionary is then saved to multiple files, one for each lemma.

```
1 |They are a common feature of modern business life, and most multinational corporations organize their operations in this way.
2 |These, and others, organize their businesses into national and functional subsidiaries, often with multiple levels of subsidiaries.
3 |For this reason, they differ from divisions, which are businesses fully integrated within the main company, and not legally or otherwise distinct from it.
4 |The parent and the subsidiary do not necessarily have to operate in the same locations or operate the same businesses.
5 |Consolidation (business)
6 |Division (business)
7 |Categories : Legal entities Business terms Business models Types of business entity Corporate subsidiaries Hidden categories: Articles with short description Short description matches Wikidata All articles with unsourced statements
8 |The Porsche automotive business is fully owned by the Volkswagen Group.
9 |Born to grape farmers in Renazzo, from the comune of Cento in the Emilia-Romagna region, his mechanical know-how led him to enter the business of tractor manufacturing in 1948, when he founded Lamborghini Trattori, which qui
10 |In 1959, he opened an oil burner factory, Lamborghini Bruciatori, which later entered the business of producing air conditioning equipment.
11 |By 1974, Ferruccio had become disenchanted with his car business.
12 |After departing the automobile manufacturing business, Lamborghini continued his business activities in other areas, including his heating and air conditioning company, Lamborghini Calor.
13 |Ferruccio retired to his estate, "La Fiorita" at Lake Trasimeno, continuing to manage his other businesses while pursuing personal interests.
14 |Lamborghini even designed his own golf course, all while continuing to manage several business interests.
15 |Whether Lamborghini's spite was enough of a reason to enter the auto manufacturing business, or if the exchange inspired further business interests or merely served as an anecdote, is unknown.
16 |Sackey writes that the economic motives for entering the high-profit-margin sports car business far outweighed Lamborghini's personal sentiments.
17 |Measure of the profitability of a business venture
18 |Selected accounts Assets Cash Cost of goods sold Depreciation BS / Amortization (business) Equity Expenses Goodwill Liabilities Profit Revenue
19 |In business and accounting, net income (also total comprehensive income, net earnings, net profit, bottom line, sales profit, or credit sales ) is an entity's income minus cost of goods sold, expenses, depreciation an
20 |Selected accounts Assets Cash Cost of goods sold Depreciation BS / Amortization (business) Equity Expenses Goodwill Liabilities Profit Revenue
21 |Equity can apply to a single asset, such as a car or house, or to an entire business.
22 |A business that needs to start up or expand its operations can sell its equity in order to raise cash that does not have to be repaid on a set schedule.
23 |A business entity has a more complicated debt structure than a single asset.
24 |While some liabilities may be secured by specific assets of the business, others may be guaranteed by the assets of the entire business.
25 |If the business becomes bankrupt, it can be required to raise money by selling assets.
26 |Yet the equity of the business, like the equity of an asset, approximately measures the amount of the assets that belongs to the owners of the business.
27 |Financial accounting defines the equity of a business as the net balance of its assets reduced by its liabilities.
28 |For a business as a whole, this value is sometimes referred to as total equity, [2] to distinguish it from the equity of a single asset.
29 |Various types of equity can appear on a balance sheet, depending on the form and purpose of the business entity.
30 |Preferred stock, share capital (or capital stock) and capital surplus (or additional paid-in capital) reflect original contributions to the business from its investors or organizers.
31 |Treasury stock appears as a contra-equity balance (an offset to equity) that reflects the amount that the business has paid to repurchase stock from shareholders.
32 |Retained earnings (or accumulated deficit) is the running total of the business's net income and losses, excluding any dividends.
33 |Equity investing is the business of purchasing stock in companies, either directly or from another investor, on the expectation that the stock will earn dividends or can be resold with a capital gain.
34 |Investors in a newly established firm must contribute an initial amount of capital to it so that it can begin to transact business.
35 |Under the model of a private limited company, the firm may keep contributed capital as long as it remains in business.
36 |Other relevant factors include the prospects and risks of its business, its access to necessary credit, and the difficulty of locating a buyer.
37 |[85] In October 2005, Porsche acquired an 18.5%BS stake in the business, and in July 2006, Porsche increased that ownership to more than 20%BS percent.
38 |* "Volkswagen GroupBS Brands and business fields".
39 |===== "Wikipedia:WikiProject:Volkswagen:Sub- and DataPage"=====
```

Figure 1: The "business" knowledge base.

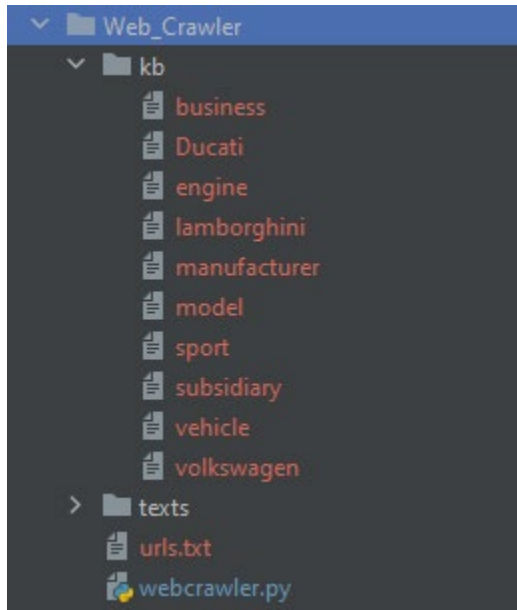


Figure 2: The directory of our program.

Sample Dialogue 1:

Hi, I'm LamboBot, a bot who knows everything about Lamborghini. What would you like to know?

*Tell me about Lamborghini's **engines**.*

Did you know that "using parts from military vehicle engines and differentials from ARAR centres, Lamborghini built the first of his "Carioca" tractors, themselves based on the six-cylinder petrol engines of Morris trucks"? Source: ...1

*That's crazy! Why was the founder of **Lamborghini** interested in tractors?*

About founder Ferruccio Lamborghini: "born to grape farmers in Renazzo, from the comune of Cento in the Emilia-Romagna region, his mechanical know-how led him to enter the business of tractor manufacturing in 1948, when he founded Lamborghini Trattori, which quickly became an important manufacturer of agricultural equipment in the midst of Italy's post-WWII economic boom". Source: ...2

*Could you direct me to more **sources** about Lamborghini?*

Sure! Try:

Lamborghini: The Spirit of the Bull (Hardcover ed.).

Lamborghini: Forty Years (Hardcover ed.).

Lamborghini: Italy's Raging Bull.

Thank you!

No problem!

Sample Dialogue 2:

Hi, I'm LamboBot, a bot who knows everything about Lamborghini. What would you like to know?

*Hello LamboBot, can you give me some **models** of Lamborghini?*

Some models are the Aventador, Gallardo, and Urus. There are other Lamborghini models in the Lamborghini museum. Source: ...1

*Can you tell me more about the **Lamborghini museum**?*

Founded in 1995 by Ferruccio's son Tonino, "the Lamborghini museum honors the Lamborghini's legacy. It was first located at Centro Studi e Ricerche Ferruccio Lamborghini in Dosso (Ferrara), which was moved to Argelato (Bologna) in 2014 with the new name Ferruccio Lamborghini Museum". Source: ...2

That's pretty cool.

I know, right?