# Natural Language Processing of ChatGPT Conversation about Space Exploration

# with Interactive Python Visualizations

Brandon L Morrow November 2, 2023

# Introduction

In recent years there has been a significant rise in Artificial Intelligence and Machine Learning algorithms that use natural language processing to interact with the user. This paper will analyze a discussion with ChatGPT regarding space exploration. Our analysis will perform natural language processing on textual data created by a conversation with ChatGPT regarding Space Exploration. We will analyze this conversation to see what insight we can extract from the data and will then create interactive data visualizations that will best display this data. The analysis we are conducting will be in response to the following questions that ChatGPT was asked:

1. In regard to Space Exploration what major discoveries have there been in the past few years?
2. Which country or private company do you believe is leading the world in space exploration?
3. What space exploration missions have there been?
4. What types of transportation systems are theoretically possible that could allow humans to expand these missions to travel to nearby solar systems? Who is working on these systems?
5. When would you predict humans will be able to travel to Mars based on current research speeds?
6. What are the biggest challenges to getting to Mars?

# Analysis

**Part of Speech Tagging**

Part-of-speech tagging (POS tagging) is a process in computational linguistics where words in a text (corpus) are marked as corresponding to particular parts of speech, based on both their definition and context. This tagging can be complex due to the ambiguity of language, where words can represent different parts of speech in different contexts. This is because certain words could be a noun or verb depending on their context. These tags are commonly taught parts of speech like nouns, verbs, adjectives, and adverbs, and there are many more categories and

sub-categories, with some tagging systems distinguishing up to 150 separate parts of speech for English (Wikipedia contributors, 2023). Our analysis sought to view the distribution of different speech counts to see if there is any noticeable pattern created by the Large Language Model ChatGPT. The table below demonstrates that we do see a similar pattern across all six questions with only moderate shifting. We will discuss further in our visualization portion, and we utilized the python package NLTK which utilized a pre trained model to classify the tokenized text.

# Table 1

*ChatGPT Part of Speech Counts*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | PRON | VERB | ADV | NOUN | PRT | ADJ | ADP | NUM | DET | CONJ |
| Question1 | 13 | 47 | 10 | 145 | 16 | 47 | 47 | 6 | 33 | 18 |
| Question2 | 8 | 49 | 8 | 147 | 14 | 41 | 34 | 1 | 23 | 20 |
| Question3 | 8 | 63 | 4 | 168 | 12 | 47 | 41 | 11 | 41 | 27 |
| Question4 | 11 | 87 | 22 | 161 | 17 | 88 | 48 | 1 | 39 | 29 |
| Question5 | 2 | 30 | 5 | 79 | 15 | 24 | 25 | 2 | 19 | 15 |
| Question6 | 3 | 83 | 10 | 152 | 19 | 69 | 39 | 4 | 41 | 34 |

# Sentiment Analysis

Sentiment Analysis is the process of computationally identifying and categorizing opinions expressed in a piece of text in order to determine whether the writer’s attitude towards a particular topic is positive, negative or neutral (Oxford Languages, 2023). During this analysis we wanted to see if ChatGPT’s responses to our Space Exploration conveyed positive, negative, or neutral sentiment. We hypothesized that there would be a mostly neutral opinion on the topic from the language model, and we calculated these results across our six questions using TextBlob, which is trained on movie review data to determine sentiment:

# Table 2

*ChatGPT Conversation Sentiment Scores*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Question 1 | Question 2 | Question 3 | Question 4 | Question 5 | Question 6 |
| Polarity | 0.125 | 0.159 | 0.125 | 0.082 | 0.110 | 0.012 |
| Subjectivity | 0.468 | 0.435 | 0.399 | 0.424 | 0.370 | 0.482 |

To interpret this polarity ranges from [-1, 1] with -1 being negative sentiment and 1 being positive sentiment. We see that ChatGPT displays a mostly neutral, but slightly positive tone when discussing Space Exploration with only a slight shift based on the question. This is exactly in line with what we expected to see regarding polarity.

The subjectivity score ranges from [0,1] with zero being very objective with facts, and 1 being very subjective and opinionated. ChatGPT scores are between 0.37 and 0.48 which means its responses are right on the line between objective facts and subjective opinion. Interestingly when we asked what seemed to be the most subjective question of “When would you predict

humans will be able to travel to Mars based on current research speeds?” we got the lowest subjective response with a rating of only 0.37.

# Named Entity Recognition

NER is a subtask of information extraction that involves identifying and classifying named entities within unstructured text into predefined categories such as names of persons, organizations, locations, medical codes, quantities, monetary values, etc.. It typically involves two phases: detection of names and classification of these names by the type of entity they refer to. (Wikipedia contributors, 2023).

We are able to clearly see from our visualization and our analysis that the organizations and people that we expected to are listed frequently within their respective categories. For example, with the list of organizations we see that NASA is the topmost mentioned organization, and for people we see that Elon Musk is one of the most mentioned. We also see Mars as the most mentioned location since we were asking about possible travel and exploration there. We see that the Spacy Python package English language model we used to classify did have one misclassification with moon being labeled a person.

# Word Frequency

Word frequency is a relatively straightforward concept and relates to the frequency in which an individual word is seen in a text (Wikipedia contributors, 2023). Our top ten words are as expected and are space(62), missions(36), mars(36), exploration(23), NASA(18), propulsion(13), travel(13), life(12), interstellar(12), and moon(12). This immediately lets us know that the topic in this discussion was about space missions and exploration. This type of analysis allows for initial observations of the topics that may be covered, and what was most talked about. We can expand upon this analysis with topic modeling.

# Topic Modeling

Refers to a type of statistical model used in statistics and natural language processing for discovering the abstract topics that occur in a collection of documents or in our case text. Topic modeling is a text-mining tool used for the discovery of hidden semantic structures in a text body, identifying clusters of similar words and understanding the balance of topics within documents. Topic models, also known as probabilistic topic models, help organize and offer insights for understanding large collections of unstructured text by discovering semantic structures in the text (Wikipedia contributors, 2023). We fit the sklearn Latent Dirichlet Allocation machine learning algorithm to our text data for this analysis. The values we got from this classification are:

# Table 3

*ChatGPT Conversation Top Topics*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Topic 1 | Technology: 0.17 | Including: 0.17 | Travel: 0.17 | Challenging: 0.17 | Collaboration: 0.17 |
| Topic 2 | Space: 34.17 | Missions: 21.17 | Exploration: 14.17 | NASA: 10.17 | Mars: 9.17 |
| Topic 3 | Mars: 15.67 | Challenges: 8.17 | Missions: 6.81 | Life: 6.41 | Radiation: 6.16 |
| Topic 4 | Technology: 0.17 | Including: 0.17 | Travel: 0.17 | Challenging: 0.17 | Collaboration: 0.17 |
| Topic 5 | Travel: 10.17 | Interstellar: 10.17 | Theoretical: 10.17 | Systems: 7.17 | Propulsion: 7.17 |
| Topic 6 | Space: 18.05 | Mars: 10.66 | Missions: 7.52 | Discoveries: 6.17 | NASA: 6.17 |

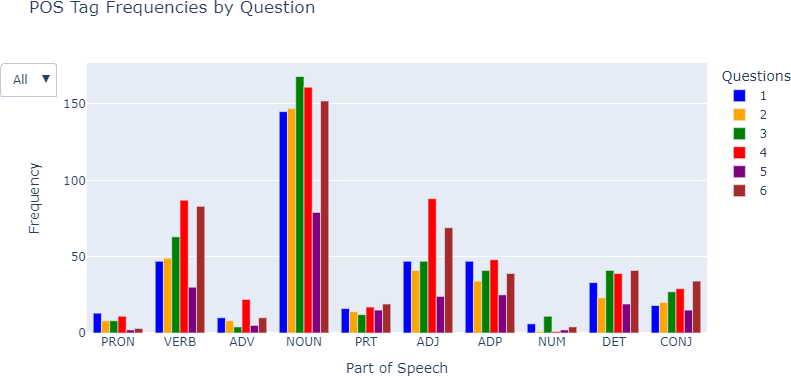
We see relatively low scores across the topics making these topics not as easily decipherable as we would hope. This algorithm only returns the topic and not a description of the topic so we must apply it. The fact that all of these top words are similar is unsurprising since we did not change the topic of the conversation much outside of space exploration. We can glean some insights from these, but not as much or as clearly as we would like.

# Visualization of Text Analysis

To visualize these different metrics, we chose between many different visualization options to find interactive visualizations that would best model our data. We implemented many interactive features to help our visualizations be easier to interpret. I utilized Plotly for four of these five visualizations, and ipywidgets for one of the five (the word cloud). Plotly is an interactive JavaScript package in Python which renders well when exported to an html file. Let’s now go in order of our analysis and talk about each visualization:

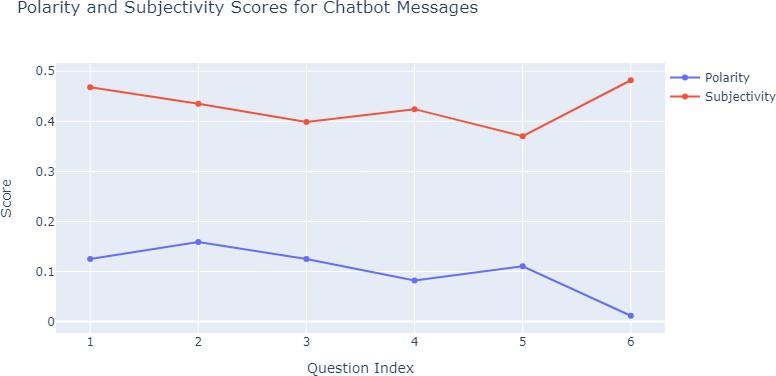
# Part of Speech Tagging

For our part of speech tagging graphic, we will utilize a grouped bar chart to be able to compare the different responses to see if there were any differences, or if the sentence structure remained relatively the same. This visualization allows us to easily compare each question and the word types. We have implemented a filter on this visualization to be able to filter by question for easier analysis of the individual questions. This chart is also interactive and hovering over each bar will display the value for that bar.



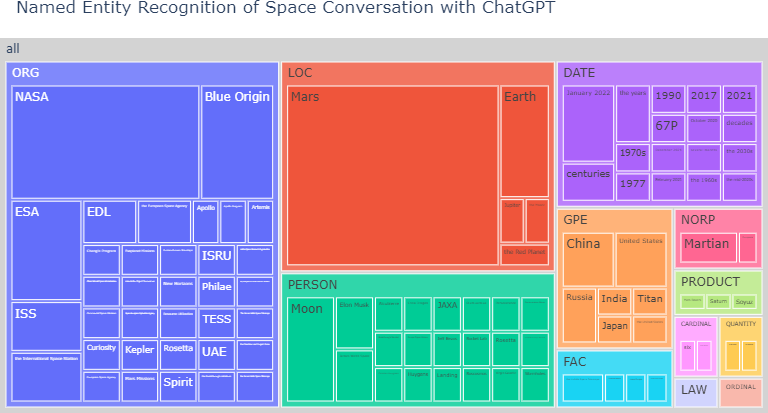
# Sentiment Analysis

We wish to interpret the sentiment over time as our conversation with ChatGPT evolved, and as a result this is best displayed with a line chart. This allows us to see the fluctuations in the polarity and subjectivity scored per question to see if anything we asked it caused a significant shift. This graphic is also interactive allowing users to hover over the points to see their values or to zoom in on a particular area if required.



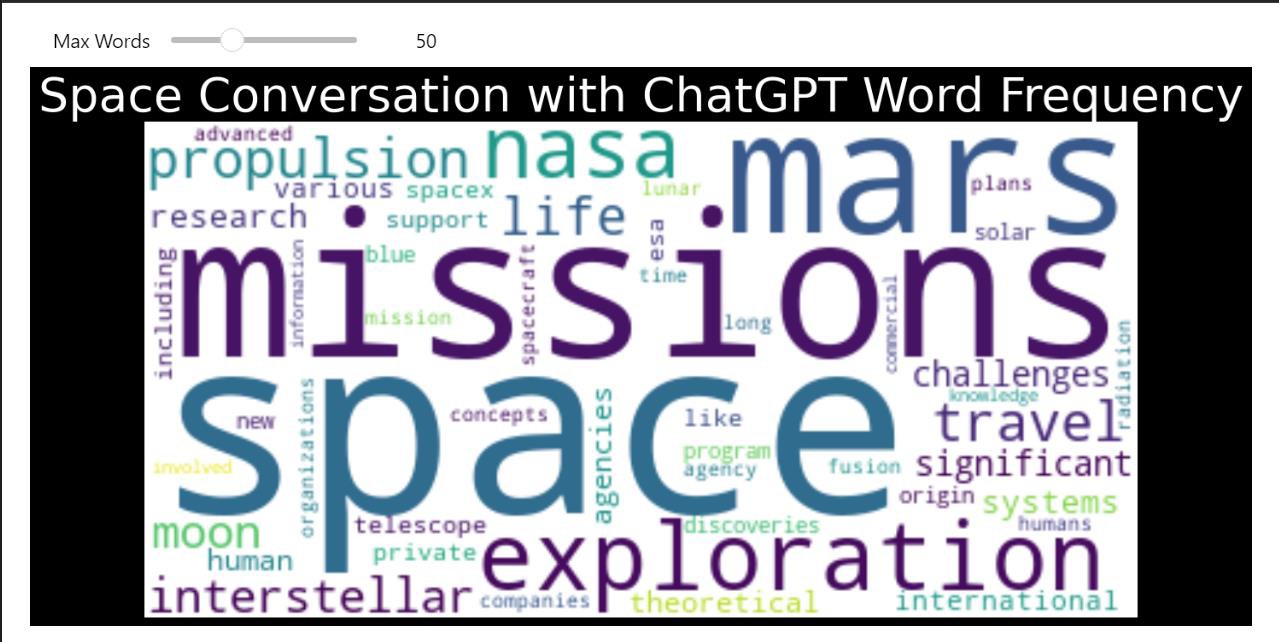
# Named Entity Recognition

Named Entity Recognition is extremely hierarchal in that it categorizes words from our text conversation into the categories’ organization, location, person, date, etc. Due to this nature, I decided that it would best be visualized by a TreeMap data visualization. For this TreeMap we counted the number of occurrences of each name and then display these names under their respective categories. This TreeMap is interactive which allows the viewer to hover over points to see the count, and to click on the smaller categories to zoom in and better see their distribution.



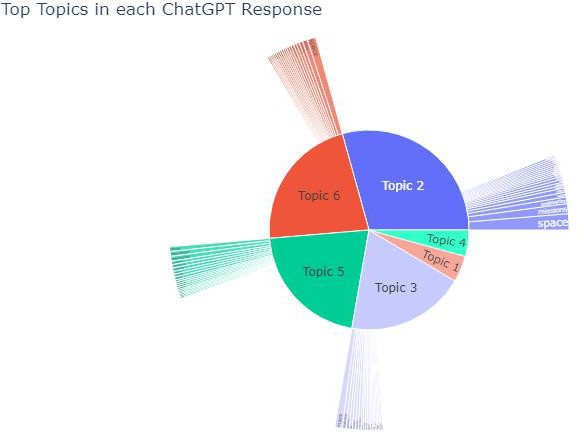
# Word Frequency

For our fourth visualization we want to visually see the top words that are used in the ChatGPT responses. This visualization is interactive and allows the user to determine how many words they want to see. To see only the ten most frequent words, you can move the slider to down to ten. This will control the number of words that are displayed on the word cloud diagram. I opted for this visualization due to its ease of interpretation, and the ability to easily see what the main topics were in the conversation. Here we clearly see words in line with our topic of “Space”, “Missions”, “Exploration”, and many more.



# Topic Modeling

For our final visualization we will visualize the Topic Modeling machine learning algorithm. This model returns hierarchal data with the words that comprise a topic, and for this reason we will use an interactive Sunburst diagram to best visualize this dataset. This will allow us to click on each of the individual topics and see what words comprise these topics. In our implementation we only show the top five subtopics but could expand this to all words comprising a topic.



# Conclusion

We have seen that there are a wide variety of ways to process natural language text. We have been able to implement many different Natural Language Processing Techniques, and we have displayed five different ways in which we can effectively model this data using interactive data visualizations. These interactive data visualizations allow us to glean much better insights than we would be able to otherwise.

# References

Oxford Languages. (2023). Sentiment analysis. Retrieved November 3, 2023, from <https://languages.oup.com/>

Wikipedia contributors. (n.d.). Part-of-speech tagging. In Wikipedia. Retrieved November 3, 2023, from <https://en.wikipedia.org/wiki/Part-of-speech_tagging>

Wikipedia contributors. (n.d.). Named-entity recognition. In Wikipedia. Retrieved November 3, 2023, from <https://en.wikipedia.org/wiki/Named-entity_recognition>

Wikipedia contributors. (n.d.). Word frequency effect. In Wikipedia. Retrieved November 3, 2023, from <https://en.wikipedia.org/wiki/Word_frequency_effect>

Wikipedia contributors. (n.d.). Topic model. In Wikipedia. Retrieved November 7, 2023, from <https://en.wikipedia.org/wiki/Topic_model>