

## **Star Wars Chatbot**

Julia Cuellar

Data Science, Bellevue University

DSC 680: Applied Data Science

Professor Catherine Williams

August 14, 2021

## Star Wars Chatbot

### **ABSTRACT**

The Star Wars universe is a science fiction phenomenon created by the famous director, producer, screenwriter, and entrepreneur, George Lucas (see Appendix A). The Star Wars universe is also chock full of characters, planets, species, starships, and vehicles featured in both canon and non-canon (or legends) work (see Appendix B). A chatbot or a variant with API manipulation will be built and created from the Star Wars data collected over 40 years through the canon universe.

### **BACKGROUND**



The Star Wars universe has developed into a worldwide marvel featuring: 6 canonical movies (of the most recent sequel trilogy being labeled non-canon), 2 spin-off movies, 1 currently running live-action Disney+ series (of the most current news, 2-4 more upcoming shows, 1 being a spin-off, 1-2 being an original, and 1 possibly being an original from a spin-off movie), 3 canonical animated show (2 being completed in their programming and 1 currently running), variant specials/earlier animated shows/movies/specials during the rising of the Star Wars franchise, as well as various video games on different platforms, plus numerous

toys/merchandise/memorabilia/novelizations/soundtracks/etc. of the galaxy far, far away (see Appendix C).

## **BUSINESS UNDERSTANDING**

The Star Wars universe is an ever-growing international icon that features various characters, planets, species, starships, and vehicles told through canon and non-canonical (or legends) work. Hence, a chatbot or a variant with API manipulation will be built and created from the Star Wars data to learn more about the Star Wars universe for fans and potential fans in a simple user-friendly format.

## **DATA UNDERSTANDING**

**Table 1**

*characters.csv dataset*

A dataframe with 87 observations and 10 variables.

	name	height	mass	hair_color	...	birth_year	gender	homeworld	species
0	Luke Skywalker	172.0	77	blond	...	19BBY	male	Tatooine	Human
1	C-3PO	167.0	75	NaN	...	112BBY	NaN	Tatooine	Droid
2	R2-D2	96.0	32	NaN	...	33BBY	NaN	Naboo	Droid
3	Darth Vader	202.0	136	none	...	41.9BBY	male	Tatooine	Human
4	Leia Organa	150.0	49	brown	...	19BBY	female	Alderaan	Human

**Table 2**

*planets.csv dataset*

A dataframe with 61 observations and 9 variables.

	name	rotation_period	...	surface_water	population
0	Alderaan	24.0	...	40.0	2.000000e+09
1	Yavin IV	24.0	...	8.0	1.000000e+03
2	Hoth	23.0	...	100.0	NaN
3	Dagobah	23.0	...	8.0	NaN
4	Bespin	12.0	...	0.0	6.000000e+06

**Table 3**

*species.csv dataset*

A dataframe with 37 observations and 10 variables.

	name	classification	...	language	homeworld
0	Hutt	gastropod	...	Huttese	Nal Hutta
1	Yoda's species	mammal	...	Galactic basic	NaN
2	Trandosha	reptile	...	Dosh	Trandosha
3	Mon Calamari	amphibian	...	Mon Calamarian	Mon Cala
4	Ewok	mammal	...	Ewokese	Endor

**Table 4**

*starships.csv dataset*

A dataframe with 37 observations and 13 variables.

	name ...	starship_class
0	Executor ...	Star dreadnought
1	Sentinel-class landing craft ...	landing craft
2	Death Star ...	Deep Space Mobile Battlestation
3	Millennium Falcon ...	Light freighter
4	Y-wing ...	assault starfighter

**Table 5**

*vehicles.csv dataset*

A dataframe with 39 observations and 11 variables.

	name ...	vehicle_class
0	Sand Crawler ...	wheeled
1	T-16 skyhopper ...	repulsorcraft
2	X-34 landspeeder ...	repulsorcraft
3	TIE/LN starfighter ...	starfighter
4	Snowspeeder ...	airspeeder

## **DATA PREPARATION**

Due to the fact that these datasets are supplied from Kaggle, there is no Nulls, but there is NA's, however, there is no data manipulation when remembering the research question due to some characters, planets, species, starships, and vehicles being unapplicable for various statistics/characteristics.

### Business Question

- 1) How can we learn about the Star Wars universe through a chatbot?

### Assumption

- I. We would be able to create a chatbot or multiple chatbots based off the CSV files of the Star Wars data to learn or better understand the Star Wars universe from canonical work.

### **MODELING/METHODS**

To perform this machine learning technique known as chatbots, harnessing the necessary libraries within the programming language, Python, as well as implementing the CSV files to build and create a chatbot is essential to learn about the Star Wars universe. However, the variant with API manipulation is the same and includes wielding user input.

### **DEPLOYMENT/RESULTS**

After building and creating a chatbot or multiple chatbots, a performance check should be implemented but has no relevance in this machine learning technique. However, with the variant in API manipulation and wielding user input, a user is able to extract data about certain characters/planets/species/starships/vehicles.

### **SUMMARY & CONCLUSIONS**

Since the five datasets were structurally formatted in a way where there is no data preparation or wrangling that needs to be done, the more legwork will be building and creating a chatbot or multiple chatbots or an API variant with user input exploitation based off the business question proposed to the business problem of what we can learn or better understand from the Star Wars universe in the form of a simple chatbot or an API variant with user input extraction. Following the CRISP-DM process, after building and creating a chatbot or multiple chatbots, a performance check should be administered but this has no significance to this machine learning technique. Likewise for the API variant with user input wielding capabilities. Lastly, an understanding of how we are able to learn about the Star Wars universe through a chatbot or API variant with user input manipulation will be executed. Constructing a chatbot was quite difficult and fell out midway due to most of the Star Wars data being a mix of categorical and numerical data. A chatbot is more inclined to textual data only, ergo, the Star Wars data was not appropriate for this machine learning technique. Therefore, an API variant with user input manipulation was employed instead. However, even though the API was exercised correctly, extracting data from the API by user input wasn't able to perform properly, hence the user would just retrieve whatever data they wanted about characters/planets/species/starships/vehicles from clicking on the API url and locating that specific characteristic/statistic about the character/planet/species/starship/vehicle. Although, a more efficient method could have been utilizing a database (partially in the code) and then withdrawing from said database through SQL coding. In the end, there was a few different modeling/methods (chatbot/multiple chatbots or API variant with user input capabilities or a database through SQL coding) that can transpired upon the Star Wars data (characters/planets/species/starships/vehicles) for deployment/results of exploring the Star Wars universe for fans and potential fans of this science-fiction splendor.

## References

Canon. (n.d.). Retrieved from [Canon | Wookieepedia | Fandom](#)

George Lucas. (n.d.). Retrieved from [George Lucas | Biography, Movies, & Facts | Britannica](#)

Goyal, K. (2020 , July 27). *How to Make a Chatbot in Python Step By Step [Python Chatterbox Guide]*. [How to Make a Chatbot in Python Step By Step \[Python Chatterbox Guide\] | upGrad blog](#). upGrad Blog.

How to Build a Chatbot with Python in 4 Easy Steps. (2021, May 12). Retrieved from [How to Create a Chatbot in Python \[Step-by-Step\] | KeyUA](#)

Kharwal, A. (2020, November 1). *Chatbot with Machine Learning and Python*. [Chatbot with Machine Learning and Python \(thecleverprogrammer.com\)](#). The Clever Programmer.

Krishnan, S. (2018, December 6). *Chatbots are cool! A framework using Python*. [Chatbots are cool! A framework using Python | by Sundar Krishnan | Towards Data Science](#). Towards Data Science.

Navlani, A. (2020, August 30). *Building a Chatbot using Chatterbot in Python*. [Python Chatterbot: How to Make a Chatbot using Python - DataCamp](#). DataCamp.

Pickell, D. (2019, February 8). *What is a Chatbot? The Full Guide to Chatbots in 2020*. [What Is a Chatbot? The Full Guide to Chatbots in 2020 \(g2.com\)](#). G2.

Saba, G. (2018, September 28). *Python Chatbot – Build Your Own Chatbot With Python*. [Python Chatbot - Build Your Own Chatbot With Python \(simplifiedpython.net\)](#). Simplified Python.

Shahid, U. (2020, June 10). *Building an AI-based Chatbot in Python*. [Building an AI-based Chatbot in Python \(datasciencedojo.com\)](#). Data Science Dojo.



Star Wars. (n.d.). Retrieved from [Star Wars | Wookieepedia | Fandom](#)

Young, J. (2019). *Star Wars*. [Star Wars | Kaggle](#). Kaggle.

## **Appendices**

### **Appendix A**

#### **George Lucas**

The article describes George Lucas' bio including his early work before Star Wars, Star Wars, and the growth of Lucasfilm Ltd.

## **Appendix B**

### **Canon**

The article describes what canon, and the expanded universe is in the Star Wars franchise.

## **Appendix C**

### **Star Wars**

This article describes everything about Star Wars and its origins, including what is considered canon or legends (non-canon), the expanded universe, the several animated or live-action television shows, the numerous books (comics included), and multiple platform video games.