

Team Skynet Presents:

chatGPT Sentiment Analysis

Multi-Class Classification

Booz | Allen | Hamilton



Booz | Allen | Hamilton

BAH Machine Learning Engineer Presentation



ChatGPT

Natural Language Processing Project that examines one month of tweets to determine sentiment regarding ChatGPT.

Meet the Team

Jules Morris,
Machine Learning
Engineer



Alyssa Soderlund,
Data Scientist



Rashod Qaim,
Data Engineer



Booz | Allen | Hamilton

Project

Sentiment analysis of Tweets featuring ChatGPT as the topic.

- Introduction
- Project Goals and Business Objectives
- Project Overview
- Architecture Design Diagram
- Tools and Software Environment Used
- Challenges
- Demo
- Question and Answers
- Conclusion



Project Goals and Business Objectives

- Create a NLP model that accurately captures the sentiment of tweets discussing ChatGPT with an accuracy above 80%.
- Final Model: 92% Accuracy | 220,000 tweets | 1.25 hours
- Increase model performance from baseline by 50%
- Our goal is to create a predictive algorithm that accurately classifies sentiment, which can eventually be scaled and generalized to capture sentiments about new technology.

Model Iteration	Manual/Auto	Amount of Data (Rows)	Accuracy	GPU/CPU	Cleansing	Stopwords	Time
1	Manual	2,000	61%	CPU	No	No	4 min
2	Manual	20,000	83%	CPU	No	No	1.5 hr
3	Manual	20,000	84%	CPU	Yes	No	1.5 hr
4	Manual	20,000	84%	CPU	Yes	Yes	1.5 hr
5	Manual	60,000	88%	CPU	Yes	Yes	4 hr
6	Manual	220,000	91%	GPU	Yes	Yes	1.25 hr
7	Manual	220,000	92%	GPU	Yes	No	1.25 hr
8	Auto	220,000	44%	Auto	No	No	20 min
9	Auto	220,000	87%	Auto	No	No	>24 hrs

Business Case Introduction

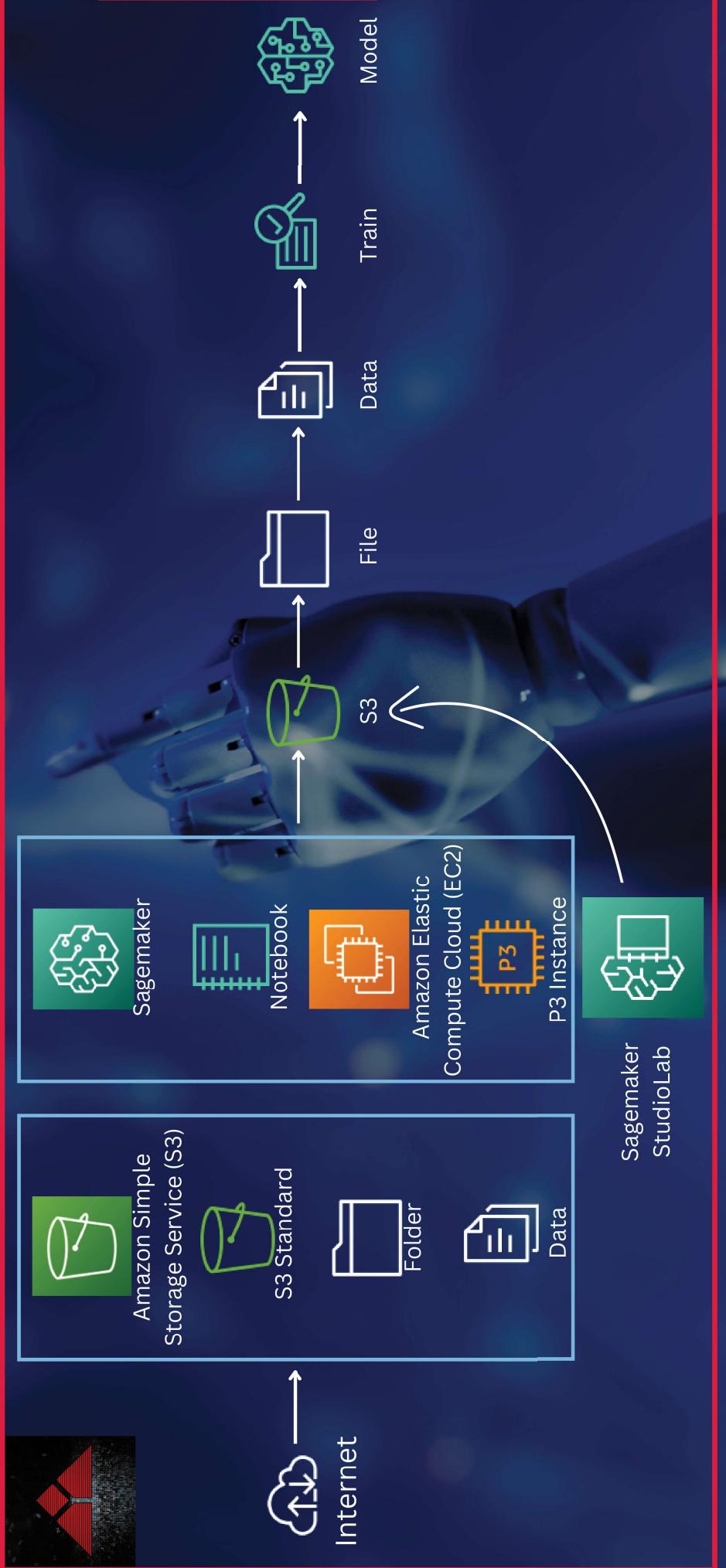
"As clients gain greater awareness of the power and potential of AI/ML to help drive their missions and business objectives, we will see more interest in bringing this capability into their organizations." - Booz Allen Hamilton Risk Mitigation Guidance

The Good...

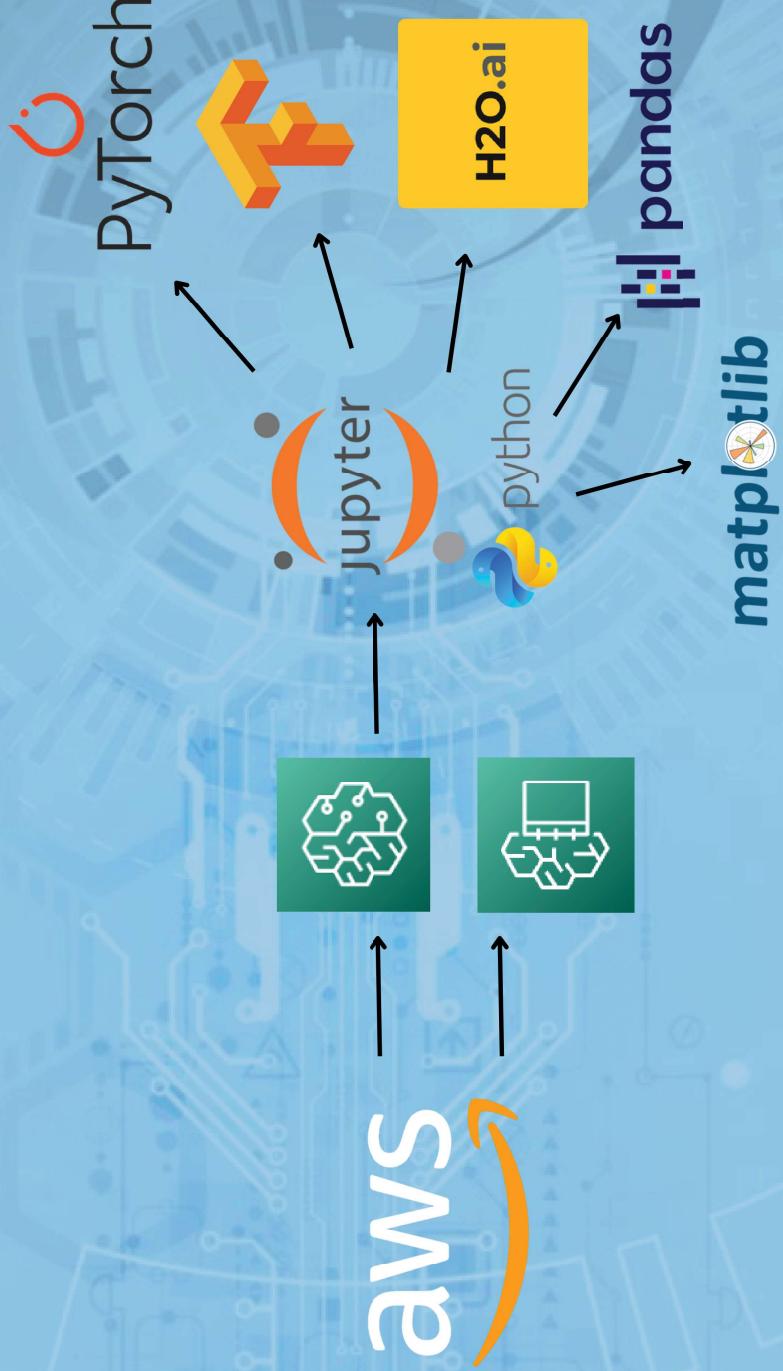
The Bad...

Architecture and Design

Skynet

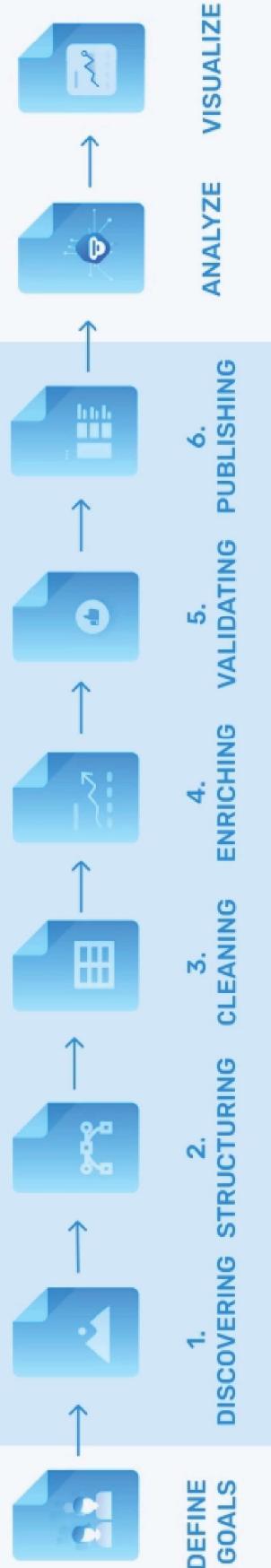


Frameworks and Modules

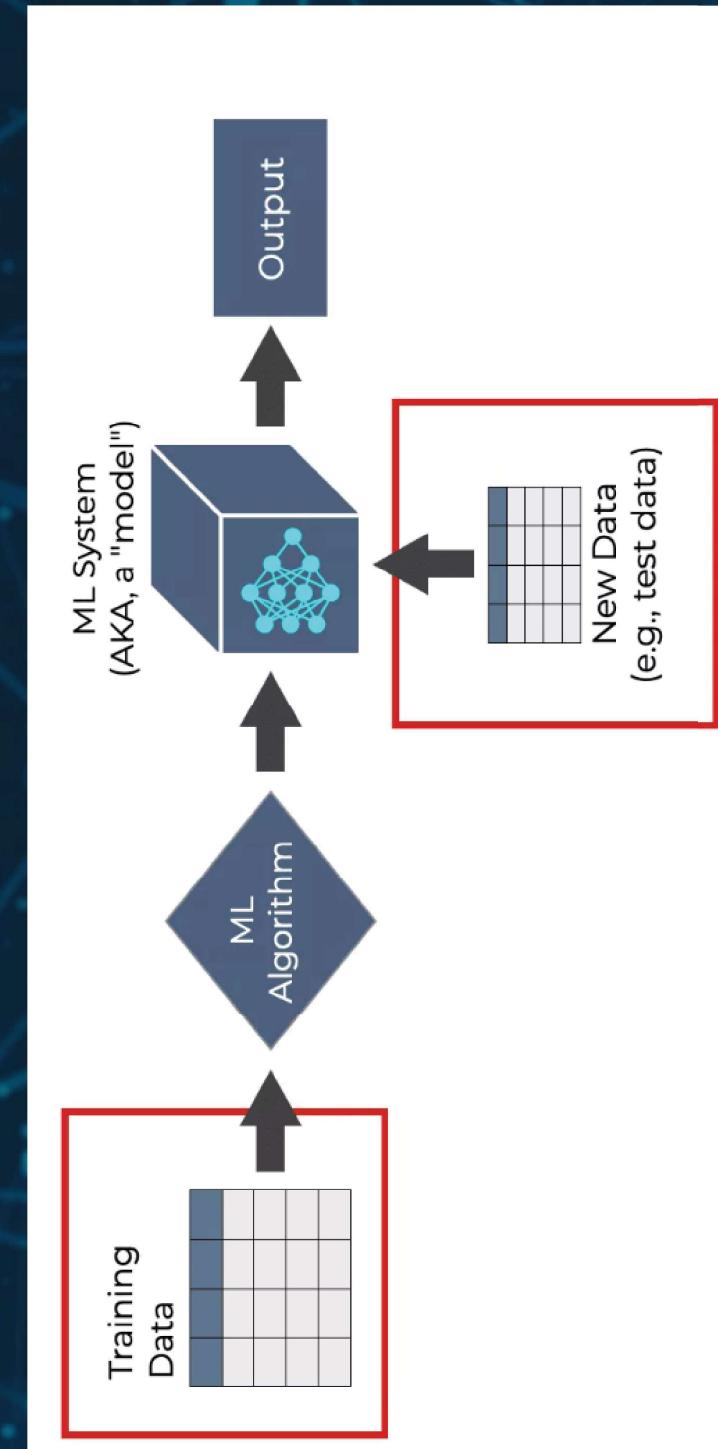


Data Preprocessing

Data Wrangling

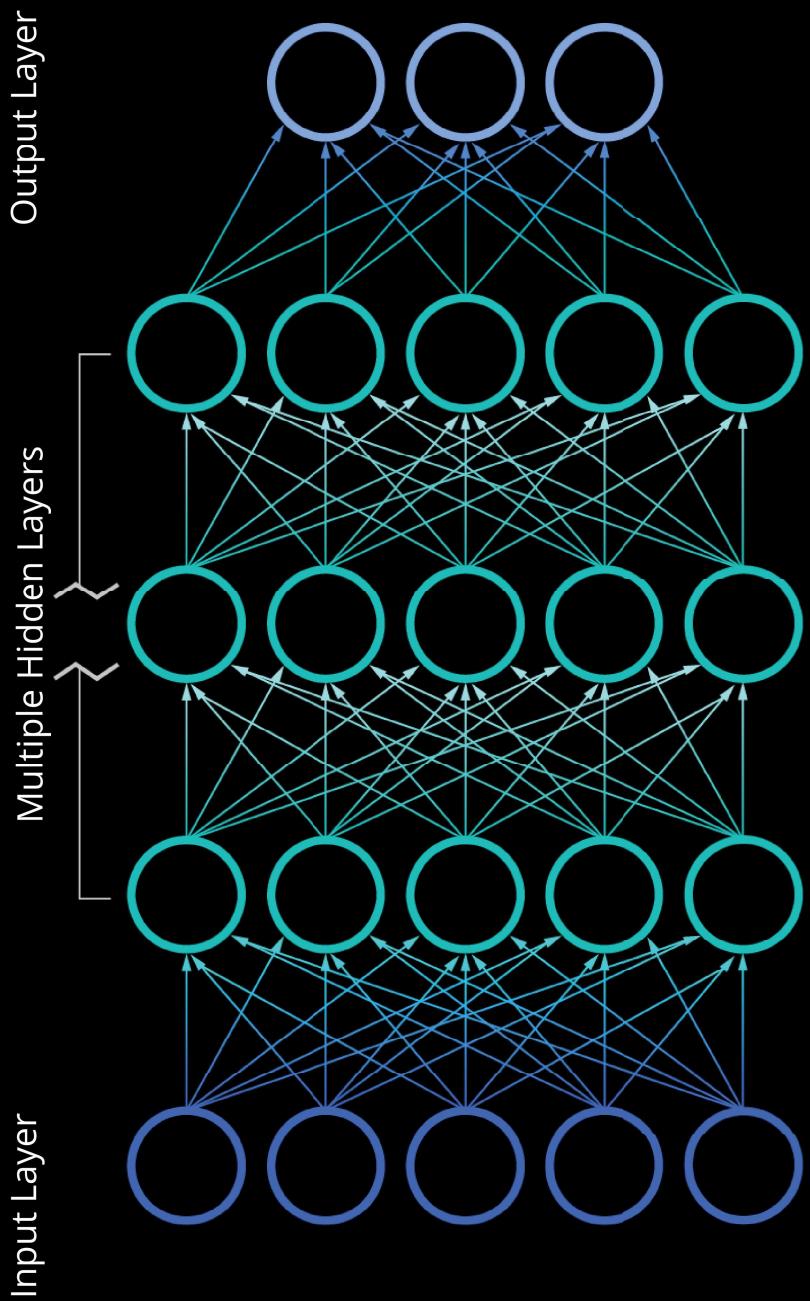


Train Test Split



Neural Networks

Deep Neural Network



BERT Model Design

- Perfect for sentiment analysis
- Computationally expensive
- Model training time: ~45 minutes
- Great metrics

AI

Challenges and Issues

- NLP model transformation process
- Data is only from November 2022
- Relyes on data from public dataset
- Sentiment can change over time
- Size of data neural network increased computational demands
- Model deployment

DEMOnSTRATION...

>Loading . . .



Future Innovations

WE ARE A BIG
COMPANY,
WITH BIGGER
IDEAS

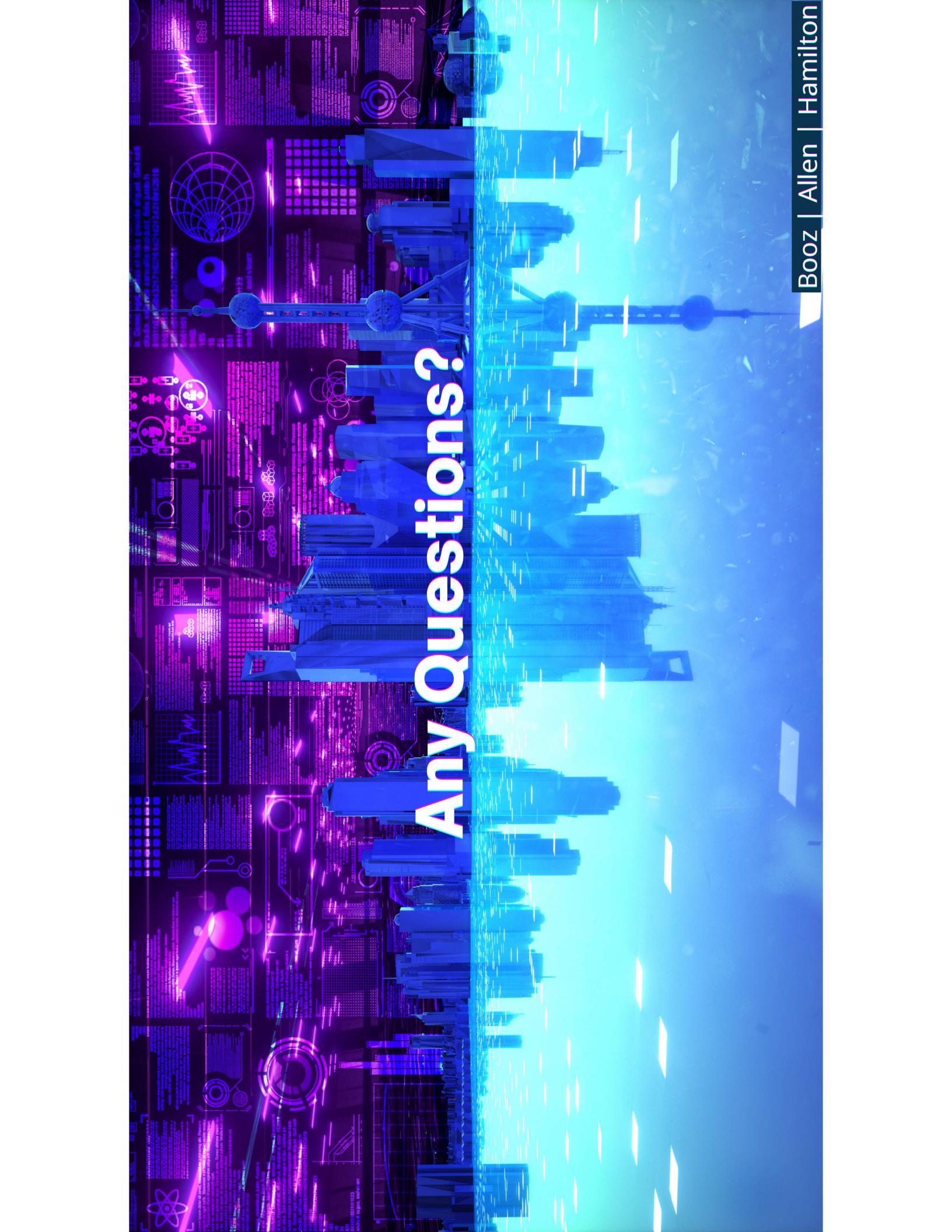


Booz | Allen | Hamilton

Booz | Allen | Hamilton

Conclusion

"What we should more concerned about is not necessarily the exponential change in artificial intelligence or robotics, but about the stagnant response in human intelligence." - ANDERS SORMAN-NILSSON



Any Questions?

Thank you!

SHALL WE PLAY A GAME?

HAL 9000

