

# Lead the Artificial Intelligence Revolution of the Federal Government!

VIRTUAL STUDENT FEDERAL SERVICE



Project Title	Lead the Artificial Intelligence Revolution of the Federal Government!
Project Summary	Are you a data analyst looking for the ultimate challenge? If you can explain “classification via logistic regression” than this project is for you!
Country	Singapore

## Project Description

Embassy Singapore Information Resource Management section seeks to research AI Machine Learning and produce a prototype environment specific to Keyphrase extraction, Sentiment analysis, Text analysis, Entity recognition, Topic modeling, language identification and translation.

With full creative control and a deadline that demands a ready to go tangible AI and Machine Learning product for the production environment, you will be tasked to have AI ingest large datasets that will ultimately make useful prediction models that will be used at the bleeding edge of application deployment in the State Department.

The focus will be to run labs to train, validate, test and deliver to end-users a realistic keyphrase identification model that will be used in the field to predict classification of large sets of texts and classifying them into subjects (i.e. science/tech documents vs resumes from the hospitality industry) with the ability to take into account specific keyword and phrases of interest that needs to be flagged.

## Required Skills or Interests

Skill(s)

Coding

Data analysis

Research

Software development

## **Additional Information**

You would fill the role of a research associate with minimal supervision but mesh well with a high collaboration tempo. Experience with discrete event simulation models and the infrastructure that comes along with customer facing products is a must.

- Experience using Google AutoML NLP, Microsoft Azure, Amazon Comprehend or IBM Watson (Lab experience from coursework is considered experience)
- Academic Coursework in probability and statistics
- The ability to analyze and interpret data from stochastic simulation output
- Programming experience in a language like Python
- Basic data analytics skills to get the data needed to quantify parameters in a model
- Experience with TensorFlow, BigQuery, Pandas and other “Ingestion” programs

## **Language Requirements**

*None*