

# Kyle Arbide

## EDUCATION

### George Washington University, Washington D.C, MD— *Master of Science (GPA: 3.92/4.00)*

August 2021 - May 2023

Data Analytics, School of Engineering and Applied Science

### University of Miami, Miami, FL — *Bachelor of Science (GPA: 3.22/4.00)*

August 2017 - May 2021

Industrial Engineering, College of Engineering

## WORK EXPERIENCE

### Noblis

Data Scientist, June 2022- Current

- Creating Knowledge Graphs by combining information extracted from both unstructured and structured sources.
- Researching and Developing LLM-based tools, including conversational agents and text summarization. This includes working with LangChain and vector stores such as ChromaDB, following industry leading research in the fields of Retrieval Augmented Generation and MapReduce, and leveraging open source models through HuggingFace.
- Building cloud solutions in the AWS environment, including end user applications, Amazon Machine Images, and security architectures. Following best practices for deploying in the cloud environment by using Terraform and Cloudformation.
- Working directly with clients to understand user requirements, develop user stories, and discuss additional features for existing applications.

### Base Operations

Statistics Intern, September 2021- May 2022

- Performed descriptive analysis and created customer reports using user interactions through Mixpanel
- Created labeled training and testing data for an NER labeling campaign using Prodigy
- Worked on reports, maps, and other data visualizations that describe and interpret crime data in different locations
- Identified and addressing data quality issues and tracing any error back to their source
- Day-to-day included using EC2 and S3 from AWS, as well as git and GitHub

### NPC-GPT

January 2023-May 2023

- Developed interactive, fully conversational NPC agents using Large Language Models (LLMs).
- Performed fine-tuning on LLMs using a synthetic dataset of NPC conversations and PyTorch.
- Leveraged Named Entity Recognition (NER) with spaCy to extract game items from NPC outputs and convert them into player quests.
- Developed a game mod for Stardew Valley in C#, integrating this project into a custom in-game NPC.

## CERTIFICATIONS

### AWS Certified Cloud Practitioner

May 2024

### Six Sigma Green Belt

1320 4th St NE Apt 1001, Washington, DC, 20002

(305)4965737

[kyle.arbide@gmail.com](mailto:kyle.arbide@gmail.com)

## SKILLS

- Experienced Analytical Programming Languages: 3 years with Python and R, 1 year with Spark
- SQL (PostgreSQL, SQLite) and NoSQL (MongoDB, Neo4j, AllegroGraph, ArangoDB, Amazon Neptune) Databases and Query Languages (SQL, cypher, SparQL)
- AWS services (Lambda, Step Functions, EC2, S3, API Gateway, IAM, ECR) and Infrastructure as Code (Terraform, CloudFormation)
- Machine Learning using Scikit- Learn and TensorFlow
- Team Skills
- Proficient in Microsoft Office programs
- Knowledge of social media platforms: Facebook, Instagram, Tumblr, LinkedIn, and Twitter

## PORTFOLIO

<https://kylearbide.github.io/>

### Special Skills

- Experienced with Descriptive, Predictive, Text, and Graph Analytics
- Proficient in Python and R and experienced with other analytical languages such as PySpark
- Bilingual: fluent in English and Spanish