

# Victor Vassallo

---

2501 Porter St NW, Apt 601, Washington, DC 20008  
(703) 975-4851 / victorevassallo@gmail.com  
[victorevassallo.com](http://victorevassallo.com)

---

## Professional Objective

*Data Scientist and AI Systems Builder who designs production-ready analytics, machine learning, and decision-support systems that turn complex information into actionable outcomes. Experienced across data pipelines, model development, experimentation, semantic modeling, and applied AI, with a strong record of translating structured and unstructured data into reliable insights for technical and non-technical stakeholders.*

## CORE SKILLS

---

- **Machine Learning & Statistical Modeling**
- **Applied AI, NLP, & LLM Workflows**
- **Exploratory Data Analysis & Validation**
- **Data Engineering & ETL Pipelines**
- **Semantic Modeling & Data Architecture**
- **Data Visualization & Stakeholder Communication**

## PROFESSIONAL EXPERIENCE

---

**Kaptive LLC**  
Data Scientist

**Alexandria, VA**  
*May 2024–Present*  
*40 hours/week*

- Design and maintain production data pipelines integrating structured records, documents, and narrative inputs across multiple federal and operational data sources to enable decision-ready analytics.
- Develop and deploy NLP and semantic modeling systems that classify qualitative research inputs, extract meaning at scale, and transform unstructured text into structured, queryable knowledge.
- Build data validation, anomaly detection, and monitoring frameworks with schema-constrained extraction pipelines and threshold alerting to ensure integrity across multi-source environments.
- Create dashboards and analytical deliverables used by 30+ program staff, translating complex operational and policy datasets into clear, actionable insights for technical and non-technical stakeholders.
- Architect and implement several retrieval-augmented generation systems including a contract opportunity intelligence platform that cut review time 70% and supported \$4M+ in successful federal funding applications.

**Data Society LLC**  
Data Science/Machine Learning  
Intern

**Washington, DC**  
*May 2023–August 2023*  
*40 hours/week*

- Developed supervised and unsupervised machine learning workflows for client-facing research projects, improving predictive performance and supporting reproducible, well-documented analytical outputs.
- Engineered Python- and SQL-based ETL processes to clean, transform, and organize large datasets for modeling, reporting, experimentation, and downstream analytics applications.
- Applied natural language processing and automated text analysis techniques to extract insights from unstructured data, improving research efficiency and analytical consistency.
- Evaluated model behavior, validated datasets, and produced technical documentation that improved transparency, usability, and communication of analytical results across stakeholders.

- Built predictive and descriptive models analyzing member behavior, supporting retention strategy, campaign optimization, and more targeted decision-making across business units.
- Automated recurring data preparation and reporting workflows, reducing manual effort while improving consistency, timeliness, and accessibility of operational insights for stakeholders.
- Integrated CRM, behavioral, and reporting data into unified analytical views that improved visibility into lifecycle trends, segmentation opportunities, and performance drivers.
- Developed dashboards and presentation-ready analyses that translated complex membership data into clear recommendations for planning, strategy, and cross-functional collaboration.

---

## KEY PROJECTS

### **Intelligent Opportunity Analysis and Decision Support System**

- Built a scalable pipeline ingesting and standardizing multi-source business development opportunity data, enabling structured evaluation workflows and consistent prioritization across large volumes of incoming information.
- Developed a retrieval-augmented analytical system using embeddings and LLM workflows to surface relevant matches, generate summaries, and support faster, higher-confidence decisions.

### **Document Intelligence and Semantic Analytics System**

- Designed document-processing and semantic analysis workflows that converted unstructured text into searchable, structured outputs for research, reporting, and downstream analytics applications.
- Applied clustering, classification, and similarity techniques to large text corpora, revealing patterns, shifts, and high-value signals across complex policy and business datasets.

---

## TECHNICAL PROFICIENCIES

- **Programming:** Python, SQL, R, Java, Bash, PowerShell
- **Applied AI / NLP:** LLM workflows, RAG, text classification, clustering, topic modeling, prompt engineering
- **Data Engineering:** ETL development, pipeline automation, API integration, data validation, orchestration
- **Dashboarding & BI:** Power BI, Tableau, Qlik Sense, ArcGIS, Google Data Studio
- **Databases:** Databricks, Snowflake, PostgreSQL, MySQL, MongoDB
- **Cloud Computing:** AWS, Azure, GCP, Linux

---

## CERTIFICATIONS

- **Google Cloud Data Engineering, Big Data, and Machine Learning**
- **Tableau Desktop Specialist / Tableau Data Visualization**
- **Databricks Certified Associate Developer**
- **Python for Data Science**
- **SQL for Data Science**

---

## EDUCATION

### **UNIVERSITY OF VIRGINIA**

**Charlottesville, VA**  
Graduated: May 2024

- Bachelor's Degrees in Economics and Statistics (Concentration in Data Science), with distinction
- Echols Scholar
- University Achievement Award – Awarded on basis of outstanding leadership
- Cumulative GPA of 3.72