

# John Speaks

Personal Website/Portfolio: [johnspeaks.com](https://johnspeaks.com)

[johnspeaksiv@gmail.com](mailto:johnspeaksiv@gmail.com) | Pittsburgh, PA | [linkedin.com/in/jtsiv](https://linkedin.com/in/jtsiv) | [github.com/JTSIV1](https://github.com/JTSIV1)

## EDUCATION

### Carnegie Mellon University

*Master of Information Technology Strategy*

08/2026

GPA: In Progress

### University of Illinois Urbana-Champaign

*Bachelor of Science in Computer Science and Linguistics (Summa Cum Laude)*

05/2025

GPA: 4.0

## WORK EXPERIENCE

### Amazon Web Services (AWS)

*Software Development Engineer Intern (AI/ML) – AWS SageMaker Unified Studio*

Bellevue, WA

05/2025 – Present

- Created a system to report enhanced app-level metrics (CPU, memory, etc.) from SageMaker Studio to customers' CloudWatch accounts, enabling persistent monitoring, diagnostics, and resource optimization for ML workloads.
- Implemented production services in Go and Java using CloudWatch Agent, CloudFormation, and internal pipelines; followed unit testing and design review standards.
- Collaborated with engineers and PMs to scope and deliver infrastructure features that improved visibility and control for ML-driven applications.

### Common Crawl Corpus Research

*Undergraduate Researcher*

Champaign, IL

09/2024 – 05/2025

- Collaborated with Professor Jonathan Dunn to build a Python pipeline for generating a high-quality global language corpus from Common Crawl web data.
- Applied NLP techniques and linguistic heuristics to filter, clean, and classify multilingual text for large-scale corpus construction.

### CyberGIS Research Center

*Software Developer and Undergraduate Researcher*

Champaign, IL

01/2023 – 01/2025

- Maintained and extended the JavaScript backend for CyberGIS-Compute to interface with HPC job schedulers, enabling remote job execution via a web interface.
- Wrote and published open-source documentation for CyberGISX and its API to support reproducible geographic analysis using HPC resources.
- Co-authored a peer-reviewed paper on accessible HPC software ([dl.acm.org/doi/10.1145/3626203.3670614](https://dl.acm.org/doi/10.1145/3626203.3670614)).

### Phillips 66

*Cloud Innovation Intern*

Houston, TX

05/2024 – 08/2024

- Refactored and standardized 23 Terraform modules to streamline cloud deployments on AWS and Azure, reducing development time and improving code reuse across teams.
- Developed a pipeline in Azure DevOps to automate infrastructure testing and validate Terraform configurations.
- Spearheaded the company's first quantum computing initiative; built a scheduling optimization proof-of-concept using AWS Braket to evaluate applicability of quantum methods.

## LEADERSHIP

### League of Linguists

*President and Undergraduate Representative*

Champaign, IL

08/2023 – 05/2025

- Lead and organize club functions with the club executive team and coordinate with school organizations.
- Elected representative to the Linguistics Department Diversity and Inclusion Committee and All-Faculty Council.

### 'Jila' - An Educational App for Q'anjob'al Speakers

*Backend Tech Lead*

Champaign, IL

09/2024 – 12/2024

- Manage the backend team: Vercel, Clerk, and REST API for data management and user auth in the ReactJS app.
- Coordinate with project managers and other teams to create an accessibility tool for local Q'anjob'al speakers.

## PROJECTS AND INDIVIDUAL RESEARCH

### Natural Language Processing and Machine Learning Projects

[johnspeaks.com/projects](https://johnspeaks.com/projects)

- Scrutinized bias in GPT-4o-mini toward groups protected under U.S. employment law using story generation and labeling prompts.
- Compared bias of static and LLM-generated synonym sets across World Englishes using Jaccard similarity and contextualized embedding evaluation.
- Measured gender bias in social media embeddings by developing novel metrics and training custom vector spaces using scraped online text.
- Evaluated GPT-4 modeled Arabic-English code-switching using Word2Vec-based semantic drift analysis.