

# 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

<b>Carnegie Mellon University</b> Master of Information Technology (Data Analytics Concentration)	Expected 08/2026
<b>University of Illinois Urbana-Champaign</b> Bachelor of Science in Computer Science and Linguistics (Statistics Minor)	05/2025 <i>Summa Cum Laude (GPA: 4.0)</i>

## Work Experience

<b>Amazon Web Services (AWS)</b> <i>Software Development Engineer Intern (AI/ML) – AWS SageMaker Unified Studio</i>	Bellevue, WA 05/2025 – 08/2025
<ul style="list-style-type: none"><li>Developed infrastructure to enable persistent monitoring, diagnostics, and resource optimization for thousands of customers' <b>ML</b> workloads by reporting enhanced app-level metrics (CPU, memory, etc.) from <b>SageMaker Studio</b>.</li><li>Engineered and deployed new production services in <b>Go</b> and <b>Java</b>, leveraging <b>CloudFormation</b> for automation and internal pipelines for continuous integration.</li></ul>	
<b>Common Crawl Corpus Research</b> <i>Undergraduate Researcher</i>	Champaign, IL 09/2024 – 05/2025
<ul style="list-style-type: none"><li>Partnered with Professor Jonathan Dunn to build a Python pipeline that generated a high-quality, diverse global language corpus from the <b>Common Crawl</b>, filling a critical gap in natural language datasets.</li><li>Applied <b>NLP</b> techniques and linguistic heuristics to filter, clean, and classify hundreds of billions of multilingual text samples for <b>large-scale corpus construction</b>.</li></ul>	
<b>CyberGIS Research Center</b> <i>Software Developer and Undergraduate Researcher</i>	Champaign, IL 01/2023 – 01/2025
<ul style="list-style-type: none"><li>Empowered less technical <b>GIS</b> researchers at multiple universities to perform complex spatial analysis on large datasets by putting powerful <b>HPC</b> resources at their fingertips.</li><li>Maintained and extended a <b>JavaScript</b> backend for <b>CyberGIS-Compute</b>, interfacing with the <b>SLURM</b> job scheduler to enable remote job execution via a <b>Jupyter Notebook</b> interface.</li></ul>	
<b>Phillips 66</b> <i>Cloud Innovation Intern</i>	Houston, TX 05/2024 – 08/2024
<ul style="list-style-type: none"><li>Refactored and standardized 23 <b>Terraform</b> modules across <b>AWS</b> and <b>Azure</b>, ensuring adherence to company standards and security policies, which eliminated the need for teams to recreate configurations and improved code reuse.</li><li>Developed a dedicated pipeline in <b>Azure DevOps</b> to periodically validate <b>Terraform</b> configurations, proactively catching infrastructure errors before they impacted our customer teams.</li><li>Spearheaded the company's first <b>quantum computing</b> initiative; built a scheduling optimization proof-of-concept using <b>AWS Braket</b> to evaluate its applicability for optimizing the feed and output of refinery operations.</li></ul>	

## Independent Projects

<b>Comparing and Evaluating Synonymy Representations Across World Englishes</b>	05/2025
<ul style="list-style-type: none"><li>Engineered a research pipeline using <b>Python</b> to evaluate synonym representations from traditional resources and contemporary <b>LLMs</b> like <b>GPT-4o</b>, <b>LLaMA3.3</b>, and <b>Deepseek-V3</b>, demonstrating a systematic bias in both toward Inner Circle English norms.</li></ul>	
<b>Scrutinizing LLM Bias Towards Groups Protected Under U.S. Employment Law</b>	12/2024
<ul style="list-style-type: none"><li>Investigated biases in the <b>GPT-4o-mini</b> model using the <b>VADER sentiment analyzer</b> on stories generated from biographical prompts in <b>Python</b>, exposing a statistically significant disparity across different protected classes.</li><li>Identified a secondary skew in how the <b>GPT-4o-mini</b> model would assign biographical labels to stories generated with different sentiments reinforcing the same disparity across protected classes.</li></ul>	

## Leadership

<b>TEEL Lab: AI Technicians Bootcamp</b> <i>Teaching Assistant</i>	Pittsburgh, PA 08/2025 – Present
<ul style="list-style-type: none"><li>Designed and lead hands-on coding exercises to teach students <b>Python</b> programming, <b>data analytics</b>, <b>CloudOps</b>, <b>DevOps</b>, and <b>AI</b> development from scratch within 12 months.</li></ul>	
<b>League of Linguists</b> <i>President and Undergraduate Representative</i>	Champaign, IL 08/2023 – 05/2025
<ul style="list-style-type: none"><li>Lead and organized club functions like professional workshops, research talks, and meetings with career professionals, growing attendance and engagement by 70% in my time as President.</li></ul>	