**Lawrence Gray, Ph.D.**

**Location:** Boise, Idaho **Email:** [lwgray@gmail.com](mailto:lwgray@gmail.com)

**Phone:** (503)764-8451

**SUMMARY**

Open-source community leader and AI expert with a proven track record of growing open-source contributor communities and delivering enterprise ML solutions. As Chair of NumFocus, stewarded major projects like Jupyter, Scikit-learn, and NumPy while expanding the Yellowbrick contributor base from 20 to 100+. Delivered $6M+ in enterprise ML solutions while providing technical direction to 15+ engineers. As Adjunct Professor at Georgetown University, demonstrated ability to communicate complex AI concepts to diverse audiences, complementing experience creating Yellowbrick documentation, tutorials, and delivering community talks at conferences including StrangeLoop and PyData NYC.

**SKILLS**

* **Open Source Community:** Contributor Engagement, Open Source Governance, Technical Documentation, Community Growth
* **AI & Engineering:** Machine Learning, Deep Learning, Natural Language Processing, Data Analysis & Visualization, Business Intelligence
* **Programming:** Python (10+ years), Git, CI/CD (GitHub Actions), API Development, Documentation Tools
* **Software:** PyTorch, Python Data Science Stack (Pandas, NumPy, Scikit-learn), Deep Learning Frameworks (Keras), NLP Libraries (NLTK, gensim), Web Development (Flask), SQL
* **Technical Direction:** Provided oversight and technical guidance for projects using Transformers(BERT), RAG Architectures (MMR), LangChain, and Vector Databases (ChromaDB)
* **Leadership:** Technical Direction, Cross-functional Team Leadership, Developer Mentorship, Project Management

**OPEN-SOURCE AND COMMUNITY LEADERSHIP**

**Chair - Board of Directors, NumFocus** | 2022 - 2025

* Led strategic initiatives impacting 50+ major open source projects in scientific computing

**Chair - Affiliated Projects Committee, NumFocus** | 2021-2022

* Reviewed applications from 50+ open-source projects seeking affiliate status

**Member - Small Development Grants, Numfocus** | 2019-2021

* Reviewed and awarded $150,000+ in grants to open-source projects

**Organizer - Salt Lake City PyData Meetup** | 2021

* Organized monthly virtual presentations for a community of 400+ data science and machine learning practitioners during the pandemic
* Facilitated knowledge sharing among the local Python data science community through online events

**Core Contributor/Maintainer, Yellowbrick** | 2018 - Present | [github.com/districtdatalabs/yellowbrick](https://github.com/districtdatalabs/yellowbrick)

* Led development of Python ML Visualization Library supporting 386,000 monthly PyPi downloads
* Authored comprehensive tutorials and documentation pages [scikit-yb.org](https://www.scikit-yb.org/)
* Created and delivered technical workshops and sprint sessions that grew the developer community from 20 to 100+
* Reviewed 200+ technical questions on StackOverflow
* Authored, reviewed, and merged 125+ PRs and 30+ Issues while maintaining code quality standards
* Developed governance framework and community guidelines for sustainable project growth

**CONFERENCE SPEAKING**

* PyData New York City, 2022 - [20 Ideas to Build Social Capital in the Data Science Ecosystem](https://youtu.be/CQlQZRWHVcY)
* Strange Loop, St. Louis, MO September 2023 - [Building Strong and Sustainable Open Source Projects](https://youtu.be/TE84-XFUT2w?si=MkcaC0kiQThD6h5n)
* PyData New York City, 2024 - Keynote Speaker - [Open Source, Open Roads: Mapping the Human Element in Tech](https://youtu.be/DKNT_AEmjEw?si=2EQ7fgPuXf-LJQtI)
* OSCON March 2025 - Keynote Speaker - [Higher-Order Thinking in an AI World](https://ospo.gwu.edu/open-source-conference-gw-oscon)

**PROFESSIONAL LEADERSHIP**

**Director of Machine Learning Engineering, KUNGFU.AI** | 2022 - 2025

* Successfully delivered $6M+ in enterprise ML solutions across 12+ projects while leading team of 15 engineers
* Architected and implemented PyTorch-based solutions for multiple enterprise clients
* Led development of key AI initiatives including:
  + Generative AI system using RAG architecture reducing response time from hours to minutes
  + Time series forecasting using PyTorch achieving 20% accuracy improvement
  + Transformer-based NLP solution with 95% accuracy in data harmonization
  + XGBoost model tripling click attribution accuracy
* Provided technical oversight and direction for projects implementing modern AI architectures including Transformers, RAG, LangChain, and vector databases
* Strengthened client partnerships with Department of State, Deloitte, DataRobot, and Wendy’s
* Promoted inclusive engineering culture through collaborative development and mentorship

**Head of Data Science, KPMG Spark** | 2020 - 2022

* Defined organization-wide data strategy and AI initiatives, increasing data literacy by 10%
* Led development of ML/AI capabilities improving productivity 4-fold across enterprise
* Implemented customer analytics reducing churn by 10% within first 100 days
* Developed automation roadmap and re-engineered processes with cross-functional teams

**Data Scientist, Maxar Technologies** | 2019 - 2020

* Developed ML systems for government clients to detect anomalous patterns in flight and maritime data
* Built end-to-end ML pipelines for classification models using Python
* Created automated processes for large-scale data processing and analysis

**ACADEMIC & EDUCATIONAL LEADERSHIP**

**Program Director - Data Analytics, Georgetown University** | July 2022 - Present

* Drive program innovation and growth, serving 300+ students across Data Analytics, Data Science, and GenAI certificates
* Source and recruit faculty while ensuring curriculum stays aligned with industry demands
* Spearhead development of new certificate programs based on emerging technology trends
* Successfully manage and oversee faculty team, maintaining program quality and educational standards

**Adjunct Professor, Georgetown University** | 2019 - Present

* Design and teach advanced courses in Data Science, Data Analytics, and GenAI
* Develop practical, industry-relevant curriculum incorporating real-world case studies and projects
* Consistently receive high student satisfaction ratings for making complex technical concepts accessible
* Pioneer new teaching methodologies for remote learning environments

**EDUCATION**

* Ph.D. in Cellular and Molecular Physiology, Johns Hopkins University, School of Medicine
* Bachelor of Science in Biology and Chemistry, California State University - Fullerton