

# Leo Torres

Lead Data Scientist & Full-Stack Developer

leo@leotrs.com ♦ leotrs.com ♦ github.com/leotrs ♦ Leipzig/Berlin

## Technical Skills

### Languages & Frameworks

- **Python (Expert):** pandas, NumPy, SciPy, FastAPI, NetworkX, scikit-learn
- **JavaScript (Advanced):** Vue.js, HTMX, Node.js, D3.js
- **AI & LLMs:** GPT-4, Claude, RAG systems, vector databases, prompt engineering
- **Other:** SQL (advanced), Mathematica, C/C++, R, Julia

### Infrastructure & Tools

- **Cloud & Data:** GCP (BigQuery, Dataflow, Cloud Run), Azure ML Studio
- **Databases:** BigQuery, PostgreSQL, experience with billion-row datasets
- **DevOps:** Docker, Airflow, GitHub Actions, Netlify, Fly.io
- **Development:** Git, pytest, CI/CD, OOP design, performance optimization

## Experience

### FGS Global - Lead Data Scientist & Tech Lead Berlin, Germany

May 2023 - Present

- Built LLM-powered pipelines processing 10M+ news articles daily for Fortune 500 clients
- Architected data infrastructure handling 1B+ log entries with sub-second query performance
- Tech lead for flagship internal product, managing team of 6 engineers
- Implemented RAG architecture and vector databases for domain-specific information retrieval
- Reduced processing time by 10x through optimized data pipelines and caching strategies
- Built robust API layer with FastAPI serving multiple internal products

### Max Planck Institute - Postdoctoral Fellow Leipzig, Germany

Aug 2021 - May 2023

- Developed spectral graph theory algorithms for complex networks
- Co-led open-source software development for research community

### Yahoo! Research - Research Intern New York, NY

May 2019 - Jul 2019

- Built graph learning models on terabyte-scale social network data
- Developed distributed computing pipelines using PySpark

## Open Source Projects

---

### The Aris Program - Founder & Lead Developer

2025 - Present

[aris.pub](https://aris.pub)

- Open-source academic publishing platform with Python, FastAPI, Vue.js
- Real-time collaboration architecture, launching 2025

### XGI (Complex Group Interactions) - Co-Lead Developer

2021 - Present

[xgi.readthedocs.io](https://xgi.readthedocs.io)

- Python library for hypergraph analysis, NumFOCUS affiliated
- NumPy pandas pytest GitHub Actions

### Manim Community Edition - Core Developer

May 2020 - May 2021

[manim.community](https://manim.community)

- Mathematical animation library popularized by 3Blue1Brown
- Graphics programming, code review, community management

### netrd - Co-Lead Developer

Jan 2019 - Jul 2019

[github.com/netsiphd/netrd](https://github.com/netsiphd/netrd)

- Implemented 40+ algorithms for network reconstruction and comparison
- Led team of 12+ developers, established coding standards
- Published in Journal of Open Source Software (JOSS)

## Professional Development

---

- **Talk:** "LLMs: Computation or Measurement?" - Berlin Data Meetup, 2024 (YouTube)
- **Workshop:** "Communicating with Data" - Designed and delivered technical workshop for enterprise clients
- **Recurse Center:** 12-week programming retreat focused on algorithm design and software craftsmanship (2016)
- **Publications:** Complete list available on [Google Scholar](#)

## Education

---

### Ph.D. Network Science - Northeastern University, Boston, MA

2016-2021

Dissertation: Spectral Aspects of Mining Complex Networks

Advisor: Tina Eliassi-Rad

### B.Sc. Mathematics - Pontificia Universidad Católica del Perú, Lima

2009-2015

Top 3% grades in 75-year history of Sciences Department

Honors Thesis: Small-World Random Graphs Beyond Watts-Strogatz