

Terence Zhi Liu

[linkedin.com/in/terence-z-liu](https://www.linkedin.com/in/terence-z-liu)

terence.z.liu@gmail.com

ML/DL Researcher | Software/Data Engineer | Computational Physicist (PhD)

(419) 936-4932

I love building and deploying **deep learning models**, **parallel compute systems** and **neural search engines**. I enjoy the cutting-edge while seeking to understand established working patterns before change. For detailed projects, write-ups and media highlights please visit terencezl.github.io.

EXPERIENCE

- **Vice President, ML Research | Clearview AI, Inc., NY** 2021 - Present
 - Created a facial recognition [algorithm](#) ranked **US #1** and **World #2**.
 - Researched & implemented a **deca-B** face search engine **500+ times** the SOTA throughput. Near-term 1/10 server cost & much higher capacity. Long-term \$10M annual savings.
 - Perfected efficient model training using tens of nodes/GPUs, enabling Research function.
 - Drove model efficiency using distillation training, quantization, pruning & trimming and hardware acceleration. Streamlined deployment with model encryption and serving engines.
 - Led cross-functional collaboration to design & review multiple compact ML data pipelines & inference infra for images/videos, resulting in \$100k cost reduction per recurring batch job.
 - Established foundational ML practices & tooling and built a team of ML engineers.
 - Planned & executed on product development & GTM of [PAD](#) feature with team.
 - Co-wrote marketing materials, public letters and engaged with policy officials.
 - Liu, Ton-That, *Scalable Training Data Pipeline...* (2022) U.S. Patent No. [11,443,553](#)
 - *MXNet, PyTorch, OpenCV, ONNX Runtime, OpenVINO, ncnn, CuPy, TensorRT, GPU batch serving, Faiss, hnswlib, TensorBoard, MLFlow, DVC, protobuf & gRPC, Redis Stream, RocksDB, AWS, Datadog, GitLab CI/CD*
- **Senior Software Engineer | Bloomberg L.P., NY** 2017 – 2021
 - Maintained & monitored widely-used bare-metal SFTP infrastructure with 7M daily logins.
 - Designed & implemented reliable account management, auth, routing, caching, messaging for cloud-based next-gen SFTP. Wrote OS-level modules and web servers for auth.
 - Leveraged S3 Storage as OpenSSH SFTP subsystem with file system emulation and cross-data-center replication & failover.
 - Spearheaded successful multi-year high-stake account migration as technical lead.
 - *C++, Golang, OpenSSH, Flask, SQLAlchemy, Postgres, Redis, MongoDB, RabbitMQ, Apache Kafka, Grafana, Splunk, Humio, Jenkins, Chef, Docker, Kubernetes, OpenStack*
- **Researcher & Developer | University of Toledo, OH** 2013 – 2017
 - Specialized in materials simulations with parallel computing clusters. 14 [papers](#) published.
 - Calculated electronic ground states with gradient descent & residual minimization schemes, (non-)linear regression - similar routines as in ML/DL frameworks. Had various ML [projects](#).
 - Authored open-source projects: [ScriptsForVASP](#), [pydass_vasp](#), [pyvasp-workflow](#).
 - Built a materials database [website](#) with a modern stack, supporting tabulation/graphing, user auth/contribution. Project helped get a \$100k research grant.
 - *Linux, Python, R, Java, NumPy, SciPy, scikit-learn, XGBoost, LightGBM, pandas, matplotlib*

EDUCATION

- **Ph.D. in Computational Physics | University of Toledo, OH** 2012 – 2017
[Dissertation](#): *A First-Principles Simulation Approach with Computational Predictive Models*
- **B.S. in Materials Physics | Nanjing University, China** 2008 – 2012
Trained in computer science, statistics, operations research, math, physics, etc.