

Lucas Cecchi

New York, NY | lucascecchi@gmail.com | lucasc-99.github.io

WORK EXPERIENCE

J.P. Morgan Chase & CO

New York, NY

Software Engineer III, AI Research

March 2022 - present

- Collaborated with researchers on projects building scalable AI-powered internal applications built in Python and C++
- First and co-authored papers on agentic systems and NLP, notably a first author paper at EMNLP 2024
- Promoted from Software Engineer II to III in February 2023
- Fine-tuned open source models from HuggingFace for text generation and classification tasks
- Scaled and optimized proof-of-concept applications using AWS Services (ECS, S3, Lambda, DynamoDB, Opensearch, Glue, and Sagemaker) provisioned with Terraform
- Lead an application from proof-of-concept stage to production: a platform for scraping and analyzing SEC data with embedding-based retrieval and LLMs, currently serving 1,000+ users and generating \$36 million in revenue per year

The Natural Language Processing Laboratory at Penn State

University Park, PA

Research Intern

Aug 2021 – Dec 2021

- Designed and implemented NLP models for sentiment analysis, part-of-speech tagging and data augmentation using BERT, GloVe, GPT-2, and GANs
- Utilized a machine translation pipeline to generate a synthetic dataset of question/answers, used to train a BERT model for automatic short answer grading

MILA Quebec Artificial Intelligence Institute

Montreal, Canada

Research Intern

May 2020 – Aug 2020, May 2021 - Aug 2021

- Contributed to open source Python projects [continuum](#) and [Sequoia](#), co-authored Sequoia publication
- Ran large scale Continual Learning and Reinforcement Learning experiments on MILA compute cluster using Slurm
- Audited graduate level courses on Reinforcement Learning and Self-Supervised Learning

TECHNICAL SKILLS

Programming Languages: Python, Terraform, C, C++, CUDA, SQL, Java, Rust

Libraries and Frameworks: Pandas, Pytorch, Numpy, Huggingface, Ray, Xgboost, Sklearn, FairSeq, Faiss, MPI, OpenMP

Data Management: Redis, Spark, Opensearch, MySQL

Containerization and Deployment: Docker, Kubernetes

Cloud: AWS, S3, ECS, Glue, DynamoDB, Sagemaker, Opensearch Service

EDUCATION

The Pennsylvania State University

University Park, PA

Bachelor of Science

December 2021

Major in Computer Science; Minor in Mathematics

Cumulative GPA: 3.4/4.0; Major GPA: 3.7/4.0

Coursework: Operating Systems; Data Structures and Algorithms; Artificial Intelligence

RECENT RESEARCH

- **Lucas Cecchi**, Petr Babkin. [ReportGPT: Human-in-the-Loop Verifiable Table-to-Text Generation](#), Empirical Methods in Natural Language Processing Industry Track 2024
- William Watson, Nicole Cho, Nishan Srishankar, Zhen Zeng, **Lucas Cecchi**, Daniel Scott, Suchetha Siddagangappa, Rachneet Kaur, Tucker Balch, Manuela Veloso. [LAW: Legal Agentic Workflows for Custody and Fund Services Contracts](#), International Conference on Computational Linguistics Industry Track 2025
- Nicole Cho, Nishan Srishankar, **Lucas Cecchi**, William Watson. [FISHNET: Financial Intelligence from Sub-querying, Harmonizing, Neural-Conditioning, Expert Swarms, and Task Planning](#), International Conference on AI in Finance 2024