

# SHIRA EISENBERG

---

[shira.dev](https://shira.dev) | shiraeisenberg114@gmail.com

## Artificial Intelligence Engineer

### AREAS OF SPECIALIZATION

- Natural Language Processing (NLP)
- Human Computer Interaction (HCI)

### TECHNICAL PROFICIENCIES

Languages: Python, C, Unity, some JavaScript (React, NextJs)

NLP: Large Language Models and transformer architectures, prompt engineering, fine tuning, embeddings, summarization, question answering, semantic search, other NLP tasks, sentence embeddings, deep learning, code generation, Langchain, vector databases

Other Writing, editing, SQL, MongoDB, Flask, word games

## Education

### BACHELOR OF SCIENCE | UNIVERSITY OF CHICAGO | JUNE 2020

- Computer Science, Specialization: Human Computer Interaction (HCI), focus Neuroscience
- Odyssey Scholar, Dean's List
- *Advanced Courses:* Networks & Distributed Systems, Machine Learning and Cancer, Computational Linguistics, Computer Architecture, Data Visualization, Disability & Design, Human Computer Interaction Engineering, Inventing, Engineering, and Understanding Interactive Devices, Theory Sequence (Discrete Mathematics, Theory of Algorithms, Formal Languages), Neuroscience of Communication, Neurobiology of Stress, Psycholinguistics, Social Neuroscience, Psychoanalysis

## Experience

**Raqia AI** | Virtual for now

**Founder / CEO** | January 2024-Present;

- \* Founder of an AI startup in stealth mode
- \* Stay tuned

**EVOLVE24 INNOVATION LABS** | Hybrid Herndon, VA

**AI Developer** | September 2023 - January 2024;

- \* Worked on a team responsible for data science and capabilities work for OSINT products for government and commercial clients using open source data
- \* Solo developed a trustworthiness metric pipeline to evaluate the reliability of publications, authors/journalists, and Twitter users. Pipeline included extracting tweets from a given author, evaluating where a document lay on an intent scale, assessing a given individual's area of expertise if any and level of expertise, extracting tweets or articles on the same topic of conversation, and maintaining a database of user trustworthiness
- \* Contributed to a core library with functionalities for clustering, classification, embedding, sentiment analysis, emotion analysis, relevance scoring, and information sourcing

**SPHINX AI STUDIOS** | Virtual

**Founder** | April 2023 - September 2023;

- \* Founded a bootstrapped AI games studio using generative AI to bring more life to gaming
- \* Managed a team of 2 game developers
- \* Developed AI Non Playable Characters (NPCs) with their own backstory, motivations, character traits, and voices
- \* Implemented speech to text and text to speech for AI NPCs
- \* Led team in developing a vertical slice of game play

**INDEPENDENT CONSULTING** | Virtual

**Consultant** | January 2023 - April 2023;

- \* Spearheaded generative AI integration for various startups
- \* Helped companies build products with guidance on fine-tuning, engineering, Retrieval-Augmented-Generation, and prompting

**INDEPENDENT AI RESEARCH GROUP** | Virtual

**AI Researcher** | March 2023 - April 2023;

- \* Worked on in silico gain of function research on the H5N1 Avian flu
- \* Used large language models trained on protein sequences to generate mutations of virus
- \* Used diffusion based docking simulation diffdock to simulate docking at both human and avian docking sites to identify potential highly transmissible candidates

**CARPER AI / ELEUTHER AI | Virtual****Open Source Contributor | January 2023 - March 2023;**

- \* Reimplemented bots called walkers in a 2D environment as in OpenAI's Evolution Through Large Models paper
- \* Worked on a crossover operator to evolve one function for a walker bot from two or three separate walker programs

**NEW ATLANTIS DAO | Virtual****AI Researcher | June 2022 - March 2023;**

- \* Trained large language model on oceanic metagenome data
- \* Used metadata such as geocoordinates, temperature, and depth to predict plankton resilience to climate change
- \* Partnered with Orcas to apply framework to data collection in Mexican marine reserves
- \* Worked with data team to engineer framework for production

**MASSACHUSETTS INSTITUTE OF TECHNOLOGY (MIT) LINCOLN LABORATORY | Lexington, MA****AI Language Systems Researcher | June 2020 – May 2021;**

- \* Worked with large language models to predict gender, age range, political alignment, and other characteristics from tweets using zero and few shot classification
- \* Built a conflict of interest detection and resolution tool for MIT campus
- \* Left to pursue own startup

**MASSACHUSETTS INSTITUTE OF TECHNOLOGY (MIT) LINCOLN LABORATORY | Lexington, MA****AI Research Intern | June 2019 – September 2019;**

- Created a novel framework to enable flexible and robust Automated Machine Learning on raw text data
- Achieved state-of-the-art performance on prolific downstream NLP tasks
- Outperformed current AutoML benchmarks while overcoming restrictive memory and structural limits
- Personally presented research to organizational group and broader laboratory community
- Expanding project to create meta learner for text embedding and universal language models

**TOYOTA TECHNOLOGICAL INSTITUTE AT CHICAGO (TTIC) | Chicago, IL****Undergraduate Researcher, Liew Family Research Fellow | September 2018 – June 2019**

- Demonstrated that, when applied to the summarization of long-form scientific articles, LexRank, an unsupervised, non-neural, extractive baseline, with original TF-IDF-based similarity scores outperforms state-of-the-art neural methods
- Introduced a novel method text summarization method augmenting LexRank with BERT
- Co-authored a conference paper and poster
- Received fellowship grant for project developing a novel adaptive system for summarizing and simplifying complex texts according to literacy level and user-specific needs

**INTEL CORPORATION, SAFFRON ARTIFICIAL INTELLIGENCE | Santa Clara, CA****Software Engineering Intern | June 2018 – August 2018**

- Engineered a novel tool for building word vectors in domain-specific vocabularies (e.g. medicine, Information Technology, etc.), using state-of-the-art embedding techniques (Gensim + Word2vec, fasttext)
- Built an interactive visualization aid to explore spatial maps of concept connections learned from data using t-distributed Stochastic Neighbor Embedding
- Re-designed and implemented a an event discovery pipeline on Issue-Resolution Data, substantially boosting entity-to-entity match detection accuracy by 20% from prior implementations
- Delivered Skype presentation to corporate audience of clientele across multiple continents

**CENTERS FOR DISEASE CONTROL AND PREVENTION (CDC) | Atlanta, GA****Scientific Editorial Assistant | June 2017 – August 2017**

- Co-authored publication: *National Amyotrophic Lateral Sclerosis (ALS) Registry Impact, Challenges, and Future Directions*,
- Publication is currently in the top 5% of all cited Research outputs on Altmetric
- Reviewed epidemiological articles for publication in Morbidity and Mortality Weekly Report (MMWR)
- Assisted top medical officials with peer reviews of high profile medical journals

**Machine Learning Engineer & Data Science Specialist | June 2017 – August 2017**

- Developed a machine learning pipeline to automate selection of "Top Articles" for CDC weekly Science Clips, accounting for population impact and translation potential, media interest, relevance to CDC Winnable Battles and Strategic Priorities, and balance of topics
- Presented finished project to senior medical officials, librarians, and technical staff
- Collaborated with Microsoft Azure federal accounts to host service on cloud
-

## Writing Experience

### WRITER | Virtual

**AlphaSignal AI** | February 2023 - April 2023;

- \* Wrote paper summaries and op-eds for a newsletter going out to 60,000 ML researchers weekly
- \* Helped edit emails

### WRITER | Virtual

**Sino Global Capital** | January 2023 - April 2023 ;

- \* Wrote weekly news and research articles on AI for VC substack and newsletter
- \* Started an AI column for a blockchain focused VC firm

### CONTRIBUTING FELLOW | Virtual

**AI & Faith** | Present;

- \* Wrote monthly reflective pieces on AI and Genomics, Consciousness, and the connections between AI and Faith

## Leadership & Community Service

### COMPILEHER COMMUNICATIONS DIRECTOR | October 2016 -- 2020

- Introduced middle school girls to computer science concepts in fun, interactive workshops, field trips, hackathons, and capstone events
- Acknowledged by the Obama Foundation
- Lead outreach operations to promote accessibility to girls in lower income neighborhoods in Chicago

## Awards & Publications

### RESEARCH AWARDS

- **Intel International Science and Engineering Fair (2016):** 1<sup>st</sup> place regional winner; 2<sup>nd</sup> place international winner
- Synaptics Special Award for the Innovative use of Human-Computer Interaction
- American Psychological Association (APA) Award for Achievement in Research in Behavior and the Psychological Sciences
- Intel Science Talent Search Research Report Award (2016)
- Liew Family Fellowship Grant for Undergraduate Research in Computer Science (2018)

### ACADEMIC AWARDS

- President's Award for Educational Excellence and Outstanding Academic Achievement – President Barack Obama (2016)
- University of Chicago Odyssey Scholarship (2016-2020)
- Dean's List University of Chicago (2016—Ongoing)
- National Merit Commended Scholar (2015)
- Merit-based grant to attend Grace Hopper Celebration of Women in Computing (2018)
- Kimpton Fellowship (2020)

### PUBLICATIONS

- CDC Grand Rounds: National Amyotrophic Lateral Sclerosis (ALS) Registry Impact, Challenge, and Future Direction. MMWR Vol. 66 No. 50. (2017)
- Going for The Goal: The Effects of Removing Preparatory Information on The Fast and Unconscious Reading of Action Goals in a Computer-Simulated Competitive Interaction. Scientia Winter 2017: 33-42 (2017)
- Outperforming Long-form Neural Abstractive Summarizers with Simple Unsupervised Baselines. Midwest Speech and Language Day Conference (2019)
- AI & Genomics, *AI & Faith*, (2022)

### TEACHING EXPERIENCE

- *Automated Machine Learning (AutoML): What it is, How to Use it, and Where it's Going*; Workshop; October 24, 2019
- Keynote Speech at the Ethics and Responsible Business Forum hosted by the Monterey Bay College of Business, in conjunction with the School of Computing and Design