Joseph Helbing

Senior AI Engineer

☎ +1 630 240 4190
www.joehelbing.net
✉ Joehelbing@gmail.com
★ github.com/JoeHelbing

EDUCATION

University of Chicago, *M.A. Computational Social Science* **[2023]**
Ohio State University, *M.A. East Asian Languages and Literatures (Mandarin)*
University of Illinois at Urbana-Champaign, *B.A. Political Science and China Studies* **[2013]** **[2011]**

LANGUAGES

English	*Native*
Mandarin	*Professionally Fluent*

SKILLS

- Python, Mojo, R, SQL, HTMX
- High Performance Computing (HPC)
- Natural Language Processing (NLP)
- Reinforcement Learning (RL)
- Machine Learning (ML)
- Web Scraping
- AWS EKS, Kubernetes, Docker, KServe, KEDA, Karpenter, vLLM
- Agent Based Modeling
- Agentic Systems, LLM Finetuning
- Data Collection, Analysis and Visualization
- Communicating Analysis to Non-Technical Audiences

EXPERIENCE

Imprivata **[2026 - Present]**

Senior AI Engineer (Full-Time)

- Migrating model inference from Amazon Bedrock to self-hosted open-weight models on Amazon EKS using KServe and vLLM
- Building a gateway-routed inference stack that uses KEDA to scale services from zero in response to queued requests and Karpenter to provision GPU-backed EC2 Spot Instances on demand
- Continuing development of agentic alert-summarization and security-data analysis systems using Pydantic AI and SQL over ephemeral DuckDB datasets

Imprivata **[2025 - 2026]**

Data Scientist (Full-Time)

- Rewrote legacy EMR data processing pipeline using Polars lazy frames, reducing processing time from 5 hours to under 2 minutes per ~30GB file across 1000+ daily runs
- Developing agentic analysis system using Pydantic AI, automating security dataset analysis through SQL querying of ephemeral DuckDB instances to triage authentication failures and accelerate incident response
- Built the core analysis platform as a FastAPI service, deployed on Kubernetes via Helm charts and ArgoCD
- Built TensorZero-based LLM gateway for API load balancing and per-customer usage and cost tracking
- Containerized MCP server wrapping AWS Bedrock Knowledge Bases, providing internal company documentation as context for agentic tabular dataset analysis

US Digital Corps at Equal Employment Opportunity Commission (EEOC) **[2024 - 2025]**

Data Scientist (Full-Time)

- Assisted EEOC investigators on AI and ML related investigations with technical consulting, forensic source code review, and case related testing
- Modernized analytics infrastructure with cloud computing and VM systems for data science tasks
- Assisted EEOC data analysts and statisticians on case investigations by training Language Model based classification systems, network analysis, and other modern data science techniques
- Took ownership of existing codebases of internal analytics tools—upgrading and refactoring with modern best practices to achieve large speedups, limit concurrency issues, and minimize resource footprints, while adapting them to lower cost platforms
- Co-led US Digital Corps NLP working group, managing group administration, facilitating inter-agency knowledge sharing, and training on NLP methods

Highlighted Project—Language Model Based Case Analysis Support

- Fine-tuned bi-directional encoder decoder vectorization models using BERT and modernBERT architectures for free form textual resume classification in support of case investigations for applicant flow analysis

- Used unstructured application text to predict job titles for unhired applicants, using hired applicants' job titles as training data with test-train splits and post training accuracy exploration to assist statisticians in identifying group differences in discrimination investigations

Library of Congress Federal Research Division **[2023 - 2024]**

Data Analyst (Full-Time)

- Long form research reporting for military and federal clients, including literature review, statistical data analysis, and visualization
- Technical lead on large-scale Natural Language Processing (NLP) project utilizing a combination of AWS GovCloud and local compute based open weights foundation models in a data extraction pipeline from legal document images
- Contributor on dual-use technologies research reports via Chinese source material for US government client

Highlighted Project—Large-Scale Document Processing System

- Architected and implemented an end-to-end pipeline for extracting structured data from 500,000+ military court martial documents across all U.S. military branches
- Architected an advanced document processing pipeline integrating AWS Textract, custom bounding box algorithms, and LLM refinement to extract 60+ variables from heterogeneous military forms, with built-in quality assurance through a purpose-built GUI for sampling based human verification
- Designed and implemented SQL database architecture for efficient storage and retrieval of extracted information
- Led technical training sessions for US Digital Corps NLP working group on LoC standardize form extraction methodology

University of Chicago Data Science Institute **[2022]**

Research Assistant (Full-Time)

- Developed a web scraper for the Security and Exchange Commission (SEC) EDGAR API to access corporate reports
- Utilized statistical textual matching techniques, XBRL scraping, and open-source pretrained machine learning models to create an information extraction pipeline

Paratech Inc **[2015 - 2021]**

Marketing Coordinator (Full-Time) (2019 - 2021)

- Oversaw the update of Paratech's corporate website, managed marketing materials, and fostered dealer partnerships, focusing on tech integration and staff training
- Launched a new WordPress website, handling design, content, and CMS
- Developed webinar series utilizing YouTube, trained the sales team, and ran live training events in the field broadcast to customers

Industrial Sales Manager (Full-Time) (2017 - 2019)

- Led the development of a new industrial sales market, establishing partnerships with distributors and manufacturing representatives
- Produced industrial, military, and maritime marketing materials, including brochures and videos, using tools like InDesign, Photoshop, and DaVinci Resolve
- Contributed on the production floor in product assembly working from engineering drawings during peak times to meet manufacturing targets

China Regional Sales Manager (Full-Time) (2015 - 2017)

- Overhauled Chinese operations strategically and organizationally, improving the company's position and navigating away from problematic relationships without harming client networks
- Transitioned local contacts to direct company relationships, retaining all dealers and clients
- Handled contract design and translation between Chinese and English, and directly negotiated partnerships through multiple visits to the country

SiMple International Inc **[2014 - 2015]**

Founder (Full-Time)

- Founded company selling international telecommunications to exchange students through partnership distribution channels

INTO University Partnerships **[2013 - 2014]**

US Recruitment Manager Beijing and North China (Full-Time)

China - Liaoning Province

- Worked with education consultancy network to recruit students to attend degree programs in the United States in representation of 6 US universities

- Northern China recruitment territory based in Dalian, Liaoning China. Self-planned and executed travel, events, and position responsibilities executed exclusively in Mandarin Chinese