# Jesse Cambon

jesse.cambon@gmail.com jessecambon.github.io jessecambon Boston, MA

## SUMMARY

Experienced machine learning engineer with expertise in NLP and large language model applications. Broad experience with cloud and machine learning infrastructure. Creator of the R package tidygeocoder.

- **Skills**: Machine Learning • NLP • LLMs • Software Development • Statistics • Data Visualization
- **Tools**: Python • R • SQL • Git • Linux • Docker • AWS • Azure

## EXPERIENCE

**Lead Machine Learning Engineer • Booz Allen Hamilton** Aug 2023 – Present

- Key machine learning developer on a RAG chatbot for searching FDA guidance documents. Led deployment of the app to AWS and implemented key features such as chat history. Utilized AWS Bedrock and Opensearch.
- Built API to provide analysis on medical device documentation using LLMs via AWS Bedrock.
- Led development of a new AI solution for extractive question answering and token classification on GPU cloud compute. Utilized Kubernetes, pytorch, CUDA, and transformers.

**Senior Data Scientist • Lighthouse** Mar 2022 – Aug 2023

- Built a flexible framework for training and deployment of large language models. Utilized Azure, pytorch, deepspeed, and Hugging Face libraries. The framework enabled faster experimentation and larger models.

**Senior Data Scientist • Embedded Healthcare (acquired by Clarify Health)** Mar 2021 – Feb 2022

- Led analytics efforts to measure the impact of a key company intervention on healthcare utilization.
- Built data pipelines, data infrastructure, and quality assurance algorithms for critical project performance analysis.

**Lead Data Scientist • Booz Allen Hamilton** Dec 2017 – Mar 2021

- Served as a technical lead and interim project lead. Managed 1-3 members of a data science team.
- Led the design and implementation of an NLP algorithm that used word embeddings to match employees to internal job opportunities. The algorithm was deployed as a web application for use by all employees and hiring managers.
- Overhauled and redesigned a python AWS Lambda application for enterprise-wide time charge accounting to provide efficient testing with AWS SAM, improved error handling, and smaller dependencies.
- Constructed regression models with Census microdata to analyze characteristics of metropolitan areas that attract immigrants. Presented at the 2018 Joint Statistical Meetings conference in Vancouver on behalf of DHS.

**Analytics Consultant • Humana** Aug 2014 – Nov 2017

- Built a predictive model to identify social isolation in a population of 3 million Medicare members.
- Performed observational and randomized studies such as measuring the effects of vaccinations and blood sugar test compliance on medical costs.
- Published research in the American Health & Drug Benefits Journal on the impact of outreaches on out-of-network and emergency room utilization.
- Key analytic contributor to the survey design and analysis for a core company initiative featured in a Harvard School of Public Health case study and published in Population Health Management (2017).

## PROJECTS

- Created the R package tidygeocoder which provides a uniform interface for using a variety of geocoding services. Tidygeocoder has over 250,000 downloads and was published in the Journal of Open Source Software (2021).

## EDUCATION

- **University of Louisville • M.Eng. & B.S**, Mechanical Engineering, highest honors (2008 – 2013)
- University of Illinois / Coursera • Accelerated Computer Science Fundamentals Specialization
- DeepLearning.AI / Coursera • 2 courses: Neural Networks and Deep Learning • Improving Neural Networks