

# Christopher H. Greer, Ph.D

## Basic Information

(773) 680-6951  
chgreer@gmail.com

[linkedin.com/in/chgreer/](https://www.linkedin.com/in/chgreer/)

## Professional Experience

### Children's Hospital Colorado

Aurora, Colorado

#### *Data Scientist Advanced*

**November, 2019 – Present**

I lead implementation of real-time machine-learning models into the production electronic health record (EHR) providing clinical decision support. I provide expertise around software engineering and devops best practices, and pilot new technologies for clinical applications.

- Designed and built a pipeline for model implementation in the production EHR to predict: risk of complications from influenza, risk of septic shock, and risk of serious bacterial infection, (Epic, Python, Podman)
- Designed and built a process using a RAG pattern with ChatGPT to cohort patient using information contained in clinical notes, doubling the precision of existing methods. Eliminated 16 hours of manual chart-review per 1000 patients. (Python, Azure OpenAI Service & AI Search)
- Replaced existing process to analyze employee survey comments using ChatGPT in a Retrieval Augmented Generation (RAG) pattern, saving of 400 developer hours per year. (Python, Azure OpenAI Service & AI Search)
- Developed a risk stratification model for a serious hospital-acquired injury attaining 80% recall with 75% precision; results to prioritize valuable nursing resources. (Python, R)
- Supervised junior team members modeling the risk of hospital acquired injury, likelihood of employee turnover, and a time series of respiratory-season hospital volumes using epidemiology data. (Data Robot)
- Designed and built a planning process to allocate data science resources and provide visibility on commitments to stakeholders.
- Co-author on clinical trial results for the sepsis models. Spoke nationally to audiences on real-time implementation and MLOps pipelines for real-time EHR models.

### Oracle

Broomfield, Colorado

#### *Principle Data Scientist*

**February, 2017 – November, 2019**

- Incorporated geolocation data into the Oracle Data Cloud Identity Graph. (Python, Spark, AWS EMR, Docker)
- Designed and built a privacy-preserving record linkage algorithm between incoming and fulfilled datasets, improving the quality of the match by 45%, scale by 30%, and standardizing the approach across 1000s of datasets. (Scala, Spark, EMR, Docker)
- Designed and built a graph-quality measurement algorithm using a Monte-Carlo approach, demonstrating a factor of  $\sim 6$  improvement over deterministic graph approaches. (Scala, Spark, EMR, Docker)

### KPMG

Denver, Colorado

#### *Sr. Associate Data Scientist*

**October, 2015 – February, 2017**

- Designed and build a document classification tool for end-users. Wrote a domain-specific language for ease-of-use. (Apache OpenNLP, Spark, Python, Elasticsearch)
- Used the document classifier for information security and control for KPMG as well as data separation for large, multinational clients across millions of documents hundreds of TB in size.

## Skills

**Data:** Bayesian statistics, machine learning, natural language processing, Fourier signal analysis, MCMC, record linkage, visualization, large-language model prompt engineering, retrieval augmented generation (RAG)

**Technology:** Apache Spark, Python, MATLAB, C, SQL, git, BASH, Docker/Podman, Luigi, Azure DevOps, Epic electronic health record, Azure/AWS, Data Robot, Elasticsearch, R, Scala

**Leadership:** Experience organizing and leading workshops and collaboration meetings, Teaching and mentoring junior team members, Public speaking, Agile development, writing/publishing, 2020 Breakthrough Prize in Fundamental Physics Laureate for contributions to the Event Horizon Telescope.

## Education

### University of Chicago, Chicago, IL

- Ph.D., Astronomy and Astrophysics, 2012
- M.S., Astronomy and Astrophysics, 2004

### Northwestern University, Evanston, IL

- B.A., Physics and Mathematics, 2002