

Charlie Guthrie

Data scientist and visualization engineer

charles.d.guthrie@gmail.com

charlieguthrie.net

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

WORK

Facebook

Data Scientist

Jul 2017 - Present

- On Community Integrity team, which enforced content policies, removing porn, hate speech, and other bad behavior from 40M jobs per month.
- Empowered team to reduce errors by designing and prototyping system for tracking accuracy of our automated and human reviewers. Using expert reviewers to establish ground truth, I determined appropriate metrics, sampling distribution, and headcount requirements to achieve reliable metrics. Worked with engineer partners to publish dashboard.
- Developed model to identify ambiguous cases, which realigned team to focus on improving accuracy in more actionable, straightforward cases.

Amazon (Audible)

Data Science Intern

Spring 2017

- Natural language processing research at leading audiobook provider.
- Clustered user comments and identified representative samples.
- Built model to extract keywords from Audible's entire library of books to tag for search. Experimented with automatic summarization of books.

Cadent

Data Science Intern

Summer 2016

- TV & digital advertising network. Used variational auto-encoding to simulate & preserve sensitive data, critical for our forecasting models.
- Overhauled design of ad forecasting system.

The New York Times

Data Vis Engineer

May 2013 - Jun 2015

- Broadly expanded data literacy, access, and decision-making by developing interactive visual dashboard used for business intelligence analytics across 10+ departments, including finance & advertising.

Dstillery

Data Science Analyst

Mar 2011 - Apr 2013

- Data analysis and testing on ad targeting models for a digital advertiser.
- Queries in Hive, modeling in R, visualization in D3, shell scripting.
- Managed intern and educated employees on our bidding process.

Carfax

Marketing Analyst

Sep 2009 - Feb 2011

- Researcher in company that produces used vehicle history reports.
- Focused sales efforts by building model to identify best client prospects.
- Developed experimental product modeling annual miles driven for insurers.

EDUCATION

See projects at charlieguthrie.net

New York University

Sep 2014 - May 2017

MSc Data Science

- Coursework included Statistics, Deep Learning, Machine Learning, Big Data, Inference, Natural Language Processing, and Algorithms

Dartmouth College

Sep 2003 - Jun 2007

B.A. Mathematics

- Coursework included Mathematical Statistics, Modeling, CS

SKILLS

Data Science

- Analysis in Python (Pandas, NumPy, SKLearn), R, Excel; queries in SQL, Hive/Hadoop; deep learning with Tensorflow/Torch.

Data Vis / Front-end

- HTML, CSS, Javascript, specializing in D3 and JQuery.

Other

- Shell scripting, Microsoft Office Suite; Conversant in French.