

Jingfei Cai, Ph.D.

I am a data scientist who can quickly adapt to various challenges and is motivated to acquire new skills. I am eager to combine data science techniques with my experience in life sciences research and communication to catalyze data-driven decision-making and deliver high-quality work.

SKILLS

Data Science: Python (numpy, pandas, sklearn, statsmodels, matplotlib, seaborn, requests, nltk, etc.), SQL, Tableau

Development: Git, Linux, Unix

Cloud Computing: Amazon Web Services, Google Cloud Platform

Computer (other): Word processing, spreadsheet, slideshow, desktop publishing, web publishing, video editing

Languages: Fluent in English, Cantonese, and Mandarin (both spoken and written)

EXPERIENCE

General Assembly Data Science Immersive | *Student* | Washington DC | June 2017 – September 2017

Participated in a 12-week, full-time, project-based training program that focuses on statistics, data analysis and visualization, and machine learning. Projects include:

- **Liquor Sales + Linear Regression** (linear models to predict sales; cross-validation to evaluate model performance; regularization tactics to compare model performance; executive summary)
- **Web Scraping & Classification** (web scraping; BeautifulSoup; classification models (random forest, logistic regression, etc.); executive summary)
- **Predicting Terrorism/Bayesian Inference** (Markov Chain Monte Carlo; Bayesian Inference; filling in missing data)
- **Capstone: Topic Evolution in Life Sciences Research** (Extracting data from XML files with lxml; Natural Language Processing; dynamic topic modeling)

LanguageLine Solutions | *Over-the-phone Interpreter for Cantonese/Mandarin (contractor)* | August 2015 – present
Interpret conversations between Cantonese/Mandarin and English speakers (~500 hours), assisting clients in diverse fields including health care, 911, government, law enforcement, court, and general business

The Journal of Visualized Experiments | *Scientific Video Editor* | Cambridge, MA | June 2013 – April 2017

- Edited manuscripts and accompanying videos (over 1000) from various scientific disciplines (*e.g.* medicine, biology, neuroscience, immunology, behavior, bioengineering, engineering, chemistry, and environment)
- Addressed comments from authors and enhanced presentation of research methods and results
- Performed quality control on manuscripts/videos as a final step before publication

Boston College, Department of Chemistry | *Research Assistant* | Chestnut Hill, MA | July 2006 – April 2013

- Overexpressed and purified EF-hand domain of mammalian phosphatidylinositol-specific phospholipase; developed binding assays to identify potential lipid ligands of EF-hand protein; designed and produced site-directed protein mutants to identify lipid binding site(s); developed NMR assays using model membrane systems to study lipid-specific interactions of lung surfactant proteins A and D; synthesized site-specific ^{13}C or ^2H labeled phospholipids using both traditional organic methods and chemoenzymatic methods (for use in high resolution field cycling NMR spectroscopy to study membrane dynamics)
- Summarized research progress in forms of slideshow/poster presentations, written reports, abstracts to three scientific conferences and manuscripts to journals

EDUCATION

Boston College, Graduate School of Arts & Sciences | Chestnut Hill, MA

Ph.D. in Chemistry | May 2013

Thesis Title: “Investigations of Lipid-specific Interactions of Lung Surfactant Proteins and EF-hand Domain of Mammalian Phosphatidylinositol-specific Phospholipase C”

Peking University, College of Chemistry and Molecular Engineering | Beijing, China

Bachelor of Science in Chemistry | July 2005