

MICHAEL W. KEARNEY, PH.D.

NLP/AI SCIENTIST

mike.wayne.kearney@gmail.com

(816) 447-8650

mikewk.com [in/mikewk](https://www.linkedin.com/in/mikewk)
[mkearney](https://github.com/mkearney) [kearneymw](https://twitter.com/kearneymw)

PROFESSIONAL SUMMARY

Internationally recognized expert in Data Science, Natural Language, and Artificial Intelligence with over seven years of professional experience, strong theoretical and quantitative backgrounds, and a knack for finding and implementing innovative solutions to complex problems.

PROFESSIONAL EXPERIENCE

- | | | | |
|--|--------------------------------------|--|-----------|
| AI SCIENTIST | <i>Aware (client)</i> | <i>Michael Kearney AI Consulting LLC</i> | 2023 |
| <ul style="list-style-type: none">Delivered recommendations on use and development of generative AI and large language modelsDescribed strengths and limitations in data sources and models and estimated project timelinesDrafted plans to implement/monitor human-in-the-loop AI labeling and offer multi-tier API access | | | |
| SENIOR AI SCIENTIST | <i>BI Labs</i> | <i>Aware</i> | 2022-2023 |
| <ul style="list-style-type: none">Proposed and built information quality model to supercharge summarization and generative AICurated external dataset of half a billion recent, diverse, and heuristically labeled social messagesCreated internal packaging for data visualization and sampling, aggregating, and uploading to S3Refactored large language models in PyTorch, implemented best practices, and enabled GPUsConverted and parameterized model training workflows from Airflow to Sagemaker Pipelines | | | |
| SENIOR DATA SCIENTIST | <i>BI Language</i> | <i>Aware</i> | 2021-2022 |
| <ul style="list-style-type: none">Owned, improved, and expanded (to non-English) featured NLP models—sentiment & toxic speechAutomated model reports on inputs, parameters, and metrics (precision, recall, f1, accuracy)Mentored junior team members and often contributed tutorials and other documentationDelivered public- and client-facing thematic analysis of workplace collaboration messagesUsed SQL and Athena to efficiently query customer data in DynamoDB, Athena, and RedshiftFrequently worked with datalakes, Apache Arrow columnar format, and unstructured JSON dataCreated labeling rule books, documented edge cases, and trained third party annotators | | | |
| ASSISTANT PROFESSOR | <i>Data Science & Journalism</i> | <i>Univ of Missouri</i> | 2017-2021 |
| <ul style="list-style-type: none">Developed and taught courses on social media data, networks, machine learning, and statisticsMentored dozens of data science and journalism capstones and chaired two doctoral dissertationsDeveloped and reviewed open-source data science tools, including popular bot detection softwareServed on various school and university committees related to data, education, and diversity | | | |
| DATA CONSULTANT | <i>Research Methods Center</i> | <i>Univ of Kansas</i> | 2016-2017 |
| <ul style="list-style-type: none">Delivered feedback on interdisciplinary research projects across university systemLed data science working groups on python, machine learning, and natural language processingContributed to project automating personalized analytics for students in a large school district | | | |
| INSTRUCTOR | <i>Communication</i> | <i>Univ of Kansas</i> | 2012-2016 |
| <ul style="list-style-type: none">Created rtweet, a massively popular R package for collecting and analyzing Twitter dataInstructor of record for quantitative methods, public speaking, leadership, and persuasionUpdated and refined course materials for core quantitative research methods course | | | |

TOOLS

LANGUAGES

Python | R | SQL | Git | Docker | Bash | JavaScript

AWS RESOURCES

S3 | Sagemaker | ECR | Athena | Redshift | DynamoDB

WEB DATA COLLECTION

REST/Stream APIs | requests | BS4 | HTTP | Rvest

DEV-OPS

Azure | Jira | GitHub | Gitlab | Confluence | Trello

DEEP LEARNING

PyTorch | TensorFlow | HuggingFace | Keras

MACHINE LEARNING

Scikit-learn | XGBoost | FastText | Tidymodels

NATURAL LANGUAGE

Spacy | HuggingFace | NLTK | RegEx | Scikit-learn

DATA SCIENCE

Polars | Pandas | Numpy | Tidymodels | Data.table

DATA VIZUALIZATION

Ggplot2 | Ggvis | Plotly | Matplotlib | D3

REPORTING

JupyterNB | Markdown | Knitr | Quarto | LaTeX

SKILLS

THEORY

Communication | Media | Interpersonal | Opinion/Topic Mining

RESEARCH METHODS

Experiments | Survey | Annotation | Algorithmic Labeling | Sampling

ARTIFICIAL INTELLIGENCE

LLMs | Transformers | Human in the Loop | Reinforcement

SUPERVISED LEARNING

Linear Models | Multilevel | Time Series | ANOVA | Structural Equation

UNSUPERVISED LEARNING

Topic Modeling | Network Analysis | PCA | Factor Analysis | Similarity

LEADERSHIP

Project Management | Team Supervisor | Committee Member & Chair

EDUCATION

DOCTOR OF PHILOSOPHY (Ph.D. with Honors)

University of Kansas

Lawrence, KS

2017

[Communication & Behavioral Sciences Methodology](#)

MASTER OF ARTS (M.A. with Honors)

University of Kansas

Lawrence, KS

2014

[Communication Studies](#)

BACHELOR OF SCIENCE (B.S.)

Missouri State University

Springfield, MO

2012

Socio-Political Communication

PUBLICATIONS (SELECTED)

PEER-REVIEWED ARTICLES

- [Comparative approaches to misinformation](#), *International Journal of Communication*, 2021
- [Exploring gender differences on Twitter with NLP](#), *Journal of Language and Social Psychology*, 2021
- [Collecting and analyzing Twitter data](#), *Journal of Open-Source Software*, 2019
- [Analyzing change in network polarization](#), *New Media Society*, 2019
- [Two tests of social displacement through social media use](#), *Information, Communication, & Society*, 2018
- [Interpersonal goals and political uses of Facebook](#), *Communication Research Reports*, 2017

POPULAR PRESS ARTICLES

- [March Madness office talk on par with Vegas odds](#), *PR Newswire*, 2023
- [Doubts cast over Elon Musk's Twitter bot claims](#), *BBC News*, 2022
- [Using data science to uncover mystery author of Times' Op-ed](#), *Reynolds Journalism Institute*, 2018
- [Trusting News project report](#), *Trusting News*, 2017

OPEN-SOURCE SOFTWARE

- [tweetbotornot2: Detecting Twitter bots](#), 100,000 downloads / 27 citations, 2019
- [textfeatures: Extracts features from text](#), 43,000 downloads / 4 citations, 2017
- [rtweet: Collecting Twitter data](#), 1,100,000 downloads / 491 citations, 2016