

# Michelle A. McSweeney

8612 Ridge Blvd Apt 2H Brooklyn, NY 11209

☎ 347-314-0072 | ✉ michelleamcsweeney@gmail.com | 🏠 www.michelleamcsweeney.com | 📷 michellejm | 📺 michelleamcsweeney

## Summary

---

Computational linguist seeking to improve dating through better messaging. Over six years experience in text analysis, corpus development, and natural language processing. Award-winning presenter, podcaster, and published author on the language of text messaging.

## Experience

---

### Director of Data Quality and Annotation

*Converseon.AI*

NLP, MACHINE LEARNING, DATA MANAGEMENT, RESEARCH & DEVELOPMENT

*Sep 2018 - Present*

Oversee data quality for building machine learning models in 10 languages. Evaluate and maintain performance of over 150 classifiers. Standardized data selection, cleaning, annotation, arbitration, and input processes. Wrote annotation guidelines for 16 tasks; hired, trained, and manage 54 freelance coders. Decreased time to product by 30% with better data handling. Developed bootstrapping technique to maximize sparse data. Saved over \$20,000 in coding costs in 3 months with statistical techniques to measure reliability.

### Assistant Professor and Lecturer

*Pratt, Columbia, CUNY, LCC*

DATA ANALYSIS, DATA VISUALIZATION, LINGUISTICS, MATH

*Aug 2005 - Jan 2007 & Jul 2011 - Present*

Teach Data Analysis with R, Data Visualization with Tableau, Mapping with QGIS, Linguistics, and Developmental Mathematics Courses.

## Projects

---

### Digital Communication of Bilingual Youth

*CUNY Graduate Center*

NLP/TEXT ANALYSIS/ANNOTATION/CLASSIFICATION

*May 2014 - Sep 2017*

Analyzed texting behavior of 50 students by ranking key word usage in 45,000 messages using Python (NLTK, SpaCy, Pandas); found that abbreviations and emojis spread in networks, but punctuation and laughter do not. Used machine learning (linear regression, sklearn) to predict if a message is between a romantic or platonic pair. Discovered patterns between literacy and digital fluency.

### Polarizing Narratives

*Columbia University*

NLP/WEB SCRAPING/TOPIC MODELLING/SENTIMENT ANALYSIS

*Jan 2018 - Aug 2018*

Scraped articles on polarizing topics and used topic modeling to find differences in argument structure by political affiliation.

### Conflict Urbanism: Urban Language Ecologies

*Columbia University*

QGIS/DATA COLLECTION/PROJECT MANAGEMENT

*Sep 2016 - Jan 2018*

Visualized spatial organization of languages in NYC to illustrate that transportation and rent affect language maintenance.

### Languages Above the Subway

*CUNY Graduate Center*

QGIS/HTML/CSS/DATA VISUALIZATION

*Aug 2012 - Aug 2014*

Analyzed the languages spoken around each subway station with QGIS, visualized the results as line charts on an interactive website.

## Education

---

2016 **PhD, Linguistics**, The Graduate Center at CUNY, New York, NY

2014 **Certificate, Interactive Technology & Pedagogy**, The Graduate Center at CUNY, New York, NY

2004 **B.A., Chemistry & Mathematics**, Michigan State University, East Lansing, MI

## Publications & Media Coverage

---

2018 **The Pragmatics of Text Messaging**, Routledge. | **Network Analysis of Feature Adoption**, LSA 91, Salt Lake City, UT. || BuzzFeed, Jul | Gioia! Aug | La Repubblica, Oct

2017 **Chatting with the Natives**, Harvard Design Magazine (44). | **Texting and academic proficiency**, Language 2(3). || New York Times, Jul | New York Magazine, Jun

## Skills

---

**Programming** Python: SciKit-Learn, Pandas, NLTK, TextBlob, SpaCy | R | **Javascript**: Leaflet | **Web**: Django, HTML, CSS  
**Software** QGIS, Tableau, Brandwatch, Excel