

# MaryClare T. Roche, PhD

---

LinkedIn: [maryclare-roche](#)

Website: [mcroche.github.io](#)

## Areas of Focus

---

R Package Development | Time Series Models | AI/ML | Mentoring | Data Visualization | Training & Development | Data Communication

## Work Experience

---

### **Arab Barometer, Princeton University; Washington, D.C.**

*Director of Technology & Innovation*

*August 2022 to present*

- Develops sophisticated analytical models, disseminates groundbreaking insights, and delivers high-impact presentations at prestigious venues including the United Nations, the State Department, various US Embassies, renowned think tanks and non-profit organizations.
- Represents organization at both professional and academic national and international conferences discussing primarily AI/ML uses in surveys and political research.
- Evaluates existing and developing technologies related to survey work, data organization and management, making recommendations on which technologies will increase efficiency.
- Leads weekly debrief and status meetings of the Data Team (5 colleagues, 2 research assistants) and developed the Data Elites, which are 2 members of the Data Team who I mentor and instruct on the more in-depth, backend development challenges
- Designs and implements dynamic, data-driven dashboards utilizing Smartsheet, effectively monitoring and analyzing Key Performance Indicators (KPIs) to drive informed decision-making and optimize organizational performance.
- Championed the progressive evolution and scaling of SQL database infrastructure.
  - Prioritized the meticulous harmonization of data records over time, accommodating nuanced changes in response coding and variable nomenclature.
  - Spearheaded the expansion of the database to encompass comprehensive question and response option wordings, meticulously tracking alterations in both Arabic and English languages to ensure data integrity and accessibility.
- Directed the Data Team in the programming of our flagship survey.
  - Taught team members how to program in ODK/XLSForms with original training modules and materials; spent time one-on-one developing specific skills as needed.
  - Crafted and executed comprehensive project management strategies, outlining timelines, milestones, and resource allocation to ensure timely survey development.
  - Implemented A/B testing methodologies to rigorously stress test the server hosting survey forms, enhancing system robustness and reliability.
  - Reduced form size by over fifty percent year-over-year while implementing new efficiencies and solving previous challenges.
- Oversaw and aided in the development of proprietary quality control and fraud detection software, which allows for daily monitoring of incoming data.
  - Software performs 150+ fraud detection and quality control tests utilizing timers, location, and response patterns at the response, enumerator, and supervisor levels.
  - Provided comprehensive training to colleagues on software utilization and result interpretation.
  - Prioritized user experience by actively soliciting and incorporating feedback to optimize software functionality.
- Collaborates with the Media Team to meticulously craft and convey our research findings and data, ensuring precise and compelling representation across diverse audiences in both Arabic and English, fostering impactful engagement and understanding.
- Directed a comprehensive talent acquisition initiative resulting in the successful recruitment and onboarding of a skilled Data Analyst.

# MaryClare T. Roche, PhD

---

*Senior Research Specialist*  
*August 2020 to August 2022*

- Spearheaded initiative to bring programming in-house, increasing programming design and data collection transparency as well as allowing for increased quality monitoring.
- Initiated construction of SQL database of over ten years of historic data across 16 countries.
- Built proprietary R package for data visualization based on established patterns of need and updated branding for both Arabic and English data.
- Established initial data processing pipeline incorporating SPSS, STATA, and R which collected, cleaned, and processed data from multiple countries simultaneously to prepare data for publishing.
  - The new processes reduced time-to-publication by over six months.
- Developed SMS messaging program in Python to routinely deliver surveys across the Middle East and North Africa and evaluated settings to ensure optimal response rate.
- Transitioned organization code storage to GitHub.
  - Trained colleagues on the use of git and GitHub with original guides and reference materials, and led several training sessions.

## **Thesis Work; University of Rochester, NY**

*PhD Candidate*

*June 2015 to June 2017*

- Scraped original text data with python BeautifulSoup and selenium via multithreading over a generated cluster of temporary, localized IP addresses with Microsoft Azure and Docker.
- Created topic models using both LDA and GloVe on Arabic text in both R and python.
- Utilized Optical Character Recognition (OCR) API on official government documents in Arabic to supplement missing information.
- Built an Arabic-to-English dictionary using Google and Bing translator APIs.
- Analyzed Irregular Time Series data to uncover temporal speaking patterns.
- Uncovered unexpected gender dynamics in Middle Eastern semi-democratic parliaments.

## **Selected Presentations**

---

### **European Political Science Association 2023; Glasgow, Scotland**

*Panel Chair & Discussant: New Developments in Text Analysis*

*Presenter: Natural Language Processing for Short Survey Responses*

### **World Association of Public Opinion Research 2022; Dubai, U.A.E.**

*Panel Chair & Discussant: Big Data, Sentiment Analysis, & Machine Learning*

*Presenter: Advances in Natural Language Processing for Surveys*

## **Education**

---

### **University of Rochester, Rochester, NY**

Doctorate of Philosophy in Political Science; Fields: Political Methodology, Formal Theory, Comparative Politics

### **University of Maryland, College Park, MD**

Bachelor of Science in Physics and Government & Politics; Double Major