

Understanding Microtargeting Pattern on Social Media

Tunazzina Islam

Department of Computer Science, Purdue University
West Lafayette, Indiana 47907 USA
islam32@purdue.edu

Abstract

We now live in a world where we can reach people directly through social media, without relying on traditional media such as television and radio. On the other hand, social media platforms collect vast amounts of data and create very specific profiles of different users through targeted advertising. Various interest groups, including politicians, advertisers, and stakeholders, utilize these platforms to target potential users to advance their interests by adapting their messaging. This process, known as *microtargeting*, relies on data-driven techniques that exploit the rich information collected by social networks about their users. Microtargeting is a double-edged sword. It enhances the relevance and efficiency of targeted content, can influence people to take action based on personal beliefs. This could be great, increasing the relevance based on users to help guide people in making better health decisions and offering them opportunities for career growth. On the other hand, it can influence people to make decisions against their own interests, foster echo chambers, and increase polarization. My research is motivated by the fact that some of these risks can be mitigated by providing transparency, identifying conflicting or harmful messaging choices, and indicating bias introduced in messaging in a nuanced way. I provide computational frameworks to analyze microtargeting patterns, which will help policymakers make better decisions. This is crucial for promoting healthy public discourse in the digital age and maintaining a cohesive society.

Research Contributions

The landscape of social media is highly dynamic, with users generating and consuming a diverse range of content. This dynamic environment enables microtargeting, allowing advertisers to target specific user groups based on their demographics, interests and behaviors. Fig. 1 shows an example of microtargeting where the same ad source tailors its messaging based on different demographics. When targeting the **older population**, it emphasizes the message “*vaccine passport is oppression*”. Conversely, while targeting **women of reproductive age**, it claims “*vaccine is dangerous for pregnant women*”. My research vision is to understand **microtargeting and activity patterns on social media** by developing computational approaches and frameworks blend-

Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

ing **computational social science (CSS)**, **natural language processing (NLP)**, and **artificial intelligence (AI)**.

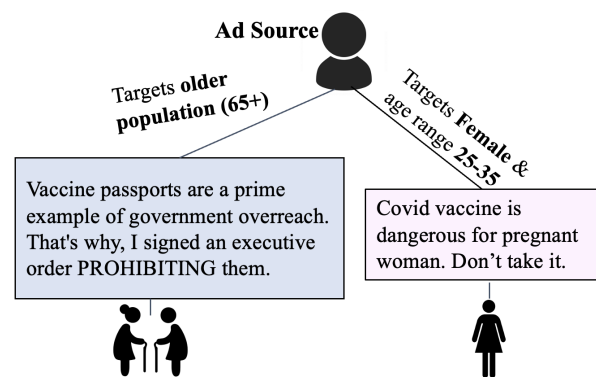


Figure 1: Example of Microtargeting.

To analyze the impacts of microtargeting, understanding messaging from both the sender’s and recipient’s perspectives is essential. For the sender, we need to know what are their motivations. For the recipient, we need to know something about their demographic properties and interests, according to which we hypothesize that messaging would change. A significant challenge lies in understanding the messaging and how it changes depending on the targeted user groups. Another challenge arises when we do not know who the users are and what their motivations are for engaging with content. My research is driven by characterizing users and messaging on social media. I address the challenges by developing computational approaches for (1) characterizing user types and their motivations for engaging with content (Islam and Goldwasser 2022, 2021a,b, 2020), (2) analyzing the messaging based on topics relevant to the users and their responses to it (Islam et al. 2023b,a, 2022), and (3) delving into the deeper understanding of the themes and arguments involved in the content (Islam and Goldwasser 2025, 2024c; Pacheco, Islam et al. 2023, 2022a,b).

Characterizing User Types and Their Motivations

Social media data, comprising text, user interactions, and social connections, forms an information graph in social networks, presenting integration challenges. I present a joint

embedding model that incorporates social and textual information to understand users, especially in lifestyle choices like **Yoga** and **Keto diet** (Islam and Goldwasser 2021a). A follow-up study tackles label scarcity with a weakly supervised graph embedding framework, leveraging social homophily and an EM-style iterative process to infer user types (e.g., `practitioners` vs `promoters`) (Islam and Goldwasser 2022).

Characterizing Messaging

My research examines social media’s influence on public campaigns and advocacy. I used a weakly supervised graph embedding model to analyze political ads focusing on **2020 U.S. presidential election**, identifying **stances** (e.g., `pro-biden`, `anti-trump`, `pro-trump`, `anti-biden`) as well as **issues** (Islam et al. 2023b). I also studied **COVID-19 vaccine** campaigns, revealing narrative shifts targeting different demographics, such as when targeting the *older population*, same ad source emphasizes on *liberty/oppression* **moral foundation** and the **vaccine mandate theme**. In contrast, when *targeting women of reproductive age*, it focuses on *care/harm* **moral foundation** and the **theme** of *vaccine efficacy* (Fig. 1) (Islam et al. 2022). Additionally, my research on **climate advertising** leverages an unsupervised approach to identify campaign **themes** and a minimally supervised model soup approach to determine **stances** (e.g., `pro-energy`, `clean-energy`, `neutral`) (Islam et al. 2023a). One key finding is that *pro-energy* ads receive the most views in Texas, while *clean-energy* ads are most viewed in California, reflecting their respective roles in energy and climate activism.

Deeper Understanding of Themes and Arguments

Previously, my research used fixed labels and themes to analyze messaging, which limited capturing nuanced insights. To address this, I developed a holistic framework for social media posts on the **COVID-19 vaccine** (Pacheco, Islam et al. 2022a), integrating stance, reason, morality frames. This approach models dependencies across levels of analysis and incorporates **human insights**. I further advanced this with an interactive framework to extract nuanced arguments (Pacheco, Islam et al. 2022b) and a concept learning framework for theme discovery (Pacheco, Islam et al. 2023), applied to **COVID-19 vaccine** and **immigration** debates.

In my early work on deeper understanding, I used a *human-in-the-loop* approach for theme and argument discovery. I then shifted to a **machine-in-the-loop** approach, leveraging LLMs to identify finer-grained themes and arguments (Islam and Goldwasser 2025, 2024c), applied to **climate** and **COVID-19 vaccine** campaigns. This method also maps *text* → *theme* and *text* → *argument*, showing how messaging is tailored to demographics and how themes and talking points shift in response to real-world events.

Future Research Directions

A major challenge is understanding the harmful effects of messaging choices when it comes to **reinforcing bias and**

stereotypes. Doing that requires us to scale up this analysis and adapt to ongoing continuous changing messaging, LLMs provide us an opportunity to reason about it and deal with how this analysis can scale up. Currently, I am working on leveraging LLMs to analyze societal opinions, biases in microtargeting (Islam and Goldwasser 2024b) to ensure equitable digital practices, and fostering human-AI collaboration in complex psycho-linguistic tasks (Islam and Goldwasser 2024a), create AI-driven insights that inform policymaking and promote positive societal change. This integrated approach ensures that AI serves as a catalyst for understanding and improving human experiences within diverse social contexts. My future research will utilize advanced AI technologies to bridge the gap between societal needs and technological solutions.

References

- Islam, T.; and Goldwasser, D. 2020. Does yoga make you happy? analyzing twitter user happiness using textual and temporal information. In *IEEE Big Data*.
- Islam, T.; and Goldwasser, D. 2021a. Analysis of Twitter Users’ Lifestyle Choices using Joint Embedding Model. In *ICWSM*.
- Islam, T.; and Goldwasser, D. 2021b. Do You Do Yoga? Understanding Twitter Users’ Types and Motivations using Social and Textual Information. In *IEEE ICSC*.
- Islam, T.; and Goldwasser, D. 2022. Twitter user representation using weakly supervised graph embedding. In *ICWSM*.
- Islam, T.; and Goldwasser, D. 2024a. Can LLMs Assist Annotators in Identifying Morality Frames? - Case Study on Vaccination Debate on Social Media. In *preprint*.
- Islam, T.; and Goldwasser, D. 2024b. Post-hoc Study of Climate Microtargeting on Social Media Ads with LLMs: Thematic Insights and Fairness Evaluation. *arXiv preprint arXiv:2410.05401*.
- Islam, T.; and Goldwasser, D. 2024c. Uncovering Latent Arguments in Social Media Messaging by Employing LLMs-in-the-Loop Strategy. In *arXiv preprint arXiv:2404.10259*.
- Islam, T.; and Goldwasser, D. 2025. Discovering Latent Themes in Social Media Messaging: A Machine-in-the-Loop Approach Integrating LLMs. In *ICWSM*.
- Islam, T.; et al. 2022. Understanding COVID-19 Vaccine Campaign on Facebook using Minimal Supervision. In *IEEE Big Data*.
- Islam, T.; et al. 2023a. Analysis of Climate Campaigns on Social Media Using Bayesian Model Averaging. In *AIES*.
- Islam, T.; et al. 2023b. Weakly Supervised Learning for Analyzing Political Campaigns on Facebook. In *ICWSM*.
- Pacheco, M. L.; Islam, T.; et al. 2022a. A Holistic Framework for Analyzing the COVID-19 Vaccine Debate. In *NAACL*.
- Pacheco, M. L.; Islam, T.; et al. 2022b. Interactively uncovering latent arguments in social media platforms: A case study on the covid-19 vaccine debate. In *DaSH*.
- Pacheco, M. L.; Islam, T.; et al. 2023. Interactive Concept Learning for Uncovering Latent Themes in Large Text Collections. In *ACL Findings*.