YUKUN YANG

Boston, MA • contact@yukunyang.info • yukunyang.info

RESEARCH SUMMARY

I study the collective and communicative practices on social media and how these acts uphold or challenge the dominant narratives, power hierarchies, and hegemonic ideologies, with a specific concentration on race and ethnicity.

Keywords: Online Communication, Race and Technology, Computational Social Science, Social Media, Marginalization

EDUCATION

Northeastern University	*Boston, MA, U.S.A.*
Ph.D. in Interdisciplinary Design and Media	*May 2028 (Expected)*

University of North Carolina at Chapel Hill	*Chapel Hill, NC, U.S.A.*
Master of Science in Information Science	*May 2020*

Master's Paper: Friends or Foes: Understanding Communication and Interaction Patterns of Homogeneous and Cross-Cutting Spaces in Online Activism. https://doi.org/10.17615/efhv-bf11.

Northwest University	*Xi'an, China*
Bachelor of Management in Library Science	*July 2018*

RESEARCH EXPERIENCE (* for Independent/Self-guided Research)

Northwestern University	*Boston, MA, U.S.A.*
Research Assistant	

Dramaturgical Performance of TikTok Refugees* – Work in Progress   Oct 2024-
- Analyzed the discursive construction of refugeeness by U.S. TikTok users on their migration to RedNote with content analysis

Mapping Migration Flow with Social Media Data* – Work in Progress   Sept 2024-
- Designed a data analysis workflow to infer migration intention disclosures on r/IWantOut computationally
- Performed inferential analysis to examine the effects of sociopolitical shocks on migration

DNA Test Discourses and Race on Social Media* – [J1]   *Jan 2024-May 2024*
- Proposed the conceptualization of "Platformed Racial Spectacle" contextualized in online discourses on DNA test results sharing with selfies on r/23andme
- Analyzed the discourses on race on r/23andme qualitatively, conceptualized the phenomenon called digital race-(re)making

Subversive Humor, Platformed Asianness, and Neologism* – [J3]   *Sept 2023-Dec 2023*
- Designed and performed a mixed-method analysis on neologisms on the subreddit r/2asia4u
- Created a data pipeline to detect neologisms online and constructed two coding schemas to capture the morphological processes and pragmatical usages of the neologisms

Self-Disclosure of Experiences of Racism on Reddit* – [C1]   *Feb 2022-Nov 2023*
- Designed a framework to categorize voluntarily disclosed narratives of experiences of racism on r/racism
- Performed content analysis to reveal the themes and intentions behind self-disclosure

Online Hate Speech and Social Reward Theory – [J2]   *Oct 2023-March 2025*
- Led data collection, cleaning, and data analysis on hate speech data on Gab

Local and National Media Dynamics on Transgender Discourses – [J4]   *Oct 2023-Feb 2025*
- Assisted with the data cleaning and data analysis on news articles discussing transgender issues

Scientific Discovery and Patent Teamwork Network – Work in Progress   *Sept 2024-*
- Conducted a literature review on the effects of diversity on team collaboration, team structure, and performance.
- Conducted primary data analysis with USPTO patents data.

Center for Antiracist Research, Boston University   *Boston, MA, U.S.A.*
Data Scientist

Built Environment, Redlining, and Racial Inequities in Health –[J1] [J6]   *May 2022-Oct 2023*
- Conducted quantitative data analysis to investigate the relationship between current-day built environment conditions and historical redlining with visual analytics, factor analysis, clustering analysis, multi-level modeling, mediation analysis, and non-parametric tests in R
- Led paper writing and data visualization

Structural Racism in COVID-19 Vaccine Distribution – [J7]   *Nov 2021-Dec 2023*
- Developed and standardized the data wrangling pipeline that scraped data from public-facing websites, formatted and geocoded the addresses, and merged it with census data
- Performed secondary regression diagnostics, participated in paper writing and revision, and developed visualizations for the publication

National Police Foundation (National Policing Institute)   *Arlington, VA, U.S.A.*
Research Data Scientist

Active Bystander Training and Ethical Policing – [J8]   *May 2021-August 2021*
- Conducted exploratory data analysis on the survey data responses with Exploratory Factor Analysis and Canonical Correlation in R
- Generated static and interactive data visualizations and participated in paper writing

Police Union Contract and Police Accountability   *Dec 2020-August 2021*
- Supervised the qualitative data collection, including MTurk worker management and double-coding review, and performed clustering on the coded documents
- Analyzed the LexisNexis news data with NLP techniques in Python

Assessment of LAPD's Response to George Floyd Protests – [R1]   *Sept 2020-March 2021*
- Identified policing hotspots and community opinion leaders with spatial-temporal clustering and tweets analysis and developed interactive dashboards
- Developed an interactive E-report, which was later presented to the LA City Council

University of North Carolina-Chapel Hill   *Chapel Hill, NC, U.S.A.*
Master's Student

Computational Music Trend Analysis* – [C2]   *Sept 2019-Jan 2020*
- Collected and wrangled textual and acoustic data from lyrics datasets and Spotify's API
- Proposed novel metrics for quantifying lyrical repetitiveness using closed itemset mining and performed time-series analyses with trend detection and Time-lagged Cross-Correlation

Power Abuse and Extreme Online Moderation on Reddit* – [C3]   *Jan 19-May 2019*
- Scraped and cleaned 110k submissions and comments from r/xkcd and r/xkcdcomic
- Conducted a quantitative case study on the effects of excessive moderation on Reddit with time-series analysis (ARIMA), breakpoint detection, and social network analysis in R

Guilty Information Disclosure on Social Media* – [C4]   *Nov 2018-Jan 2019*
- Retrieved, cleansed 225k submissions and comments from r/confession, and visualized the topic distribution of submissions with LDA in Python
- Constructed a thematic model capturing the intents, contents, and interactions of the guilty information-sharing behavior on social media with sampled data

Communication and Interaction Patterns of Homogenous and Cross-Cutting Spaces in Online Activism* – [Master's Paper]
- Scraped and classified 60k+ historical tweets about #SilentSam with word embeddings
- Conducted non-parametric significant tests using LIWC and Perspective API and analyzed the subgraph pattern of the network

Digital Innovation Lab, University of North Carolina-Chapel Hill   *Chapel Hill, NC, U.S.A.*
Social Media Data Analyst   *Feb 2019-May 2020*
- **#SilentSam Tweets Analysis**: Analyzed the tweets of a campus activism event with visual and spatial analytics, topic modeling, and network analysis and delivered multiple interactive dashboards in Python and R
- **Pinterest for Digital Humanities**: Designed, developed, and deployed an online application for a graduate-level digital humanities class that allows scraping images from Pinterest and

provides a descriptive analysis of the search results

OTHER WORK EXPERIENCE

ArcLight Analytics   *Chapel Hill, NC, U.S.A.*
Data Scientist   *July 2020 -June 2023*
- Scraped and wrangled multiple large-scale datasets from public/private sectors, including Crunchbase, NSF, PPP loans FIOA, Federal Workforce, NCSL, BLS, etc.
- Performed data analysis with survey data, administrative data, and text data in Python
- Developed and deployed a static project website and two interactive data dashboards

PUBLICATIONS

Peer-Reviewed Conference Publications (Full papers/Short papers/Posters)

[C1] **Yang, Y.**, Tagore, P., & Welles, B. F. "I'm so sick and tired of this": Understanding the Roles and Scenarios in Self-Disclosure of Everyday Experiences of Racism on Social Media. In *Proceedings of ACM SIGCHI Conference on Computer-Supported Cooperative Work & Social Computing (CSCW)* 2025. Accepted.

[C2] **Yang, Y.** (2020). "Musicalization of the Culture": Is Music Becoming Louder, More Repetitive, Monotonous, and Simpler? In *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, 14(1), 750-761. Retrieved from https://aaai.org/ojs/index.php/ICWSM/article/view/7340.

[C3] **Yang, Y.** (2019). When Power Goes Wild Online: How Did a Voluntary Moderator's Abuse of Power Affect an Online Community? In *Proceedings of the Association for Information Science and Technology (ASIS&T)*, 56: 504-508. Retrieved from doi:10.1002/pra2.55.

[C4] **Yang, Y.**, & Huang, Y. (2019). Dumping the Closet Skeletons Online: Exploring the Guilty Information Disclosure Behavior on Social Media. In *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, 13(01), 663-666. Retrieved from https://wvvw.aaai.org/ojs/index.php/ICWSM/article/view/3267.

Peer-Reviewed Journal Papers (* for co-first authorship)

[J1] **Yang, Y.** r/23andMe and My Pictures: Platformed Racial Spectacle as Collective Racial Identity Formation. Revise & Resubmit.

[J2] **Yang, Y.**, Lange, R. A. & Walther, J. B. How Social Responses to Online Hate Messages Affect Hatefulness. Submitted.

[J3] **Yang, Y.**, & Welles, B. F. Subversive Humor and Platformed Asianness: Resisting West Homogenization and Asian Ultranationalism through Satirical Neologism. Submitted.

[J4] Smith A., Kumar S., **Yang, Y.**, Goel P. Locating the Asymmetry in Information Flow between Local and National Media on Transgender Discourses. *The Bulletin of Applied Transgender Studies (BATS)*. Accepted.

[J5] Cho, A., **Yang, Y.***, Nguyen, Q., & Nsoesie, E. O. Census Tract Built Environment Resources

in Relation to Contemporary Social Characteristics and Historical Redlining. Submitted.

[J6] Kang, M., Cho, A., **Yang, Y.** (2025). My "Kind" of Food: How Ethnic Food Categories Were Changed in Local Mainstream Media From the Case Study of Boston Magazine. *Food, Culture & Society*, 1-26.

[J7] Holtzman, G. S., **Yang, Y.**, Louis, P., West, S. G., & Kandaswamy, P. (2024). Racial disparities in Phase 1 COVID-19 vaccine shipments to Neighborhood sites in Pennsylvania by the Federal Retail Pharmacy Program. *Scientific Reports*, 14(1), 23591.

[J8] **Yang, Y.*,** Cho, A., Nguyen, Q., & Nsoesie, E. O. (2023). Association of Neighborhood Racial and Ethnic Composition and Historical Redlining With Built Environment Indicators Derived From Street View Images in the US. *JAMA Network Open*, 6(1), e2251201-e2251201. doi:10.1001/jamanetworkopen.2022.51201

[J9] Taniguchi, T., Vovak, H., Cordner, G., Amendola, K., **Yang, Y.**, Hoogesteyn, K., & Bartness, M. (2022). The Impact of Active Bystander Training on Officer Confidence and Ability to Address Ethical Challenges. *Policing: A Journal of Policy and Practice*, 16(3), 508-522.

Reports

[R1] National Police Foundation. (2021). A crisis of trust: A National Police Foundation report to the Los Angeles Board of Police Commissioners on the Los Angeles Police Department's response to First Amendment assemblies and protests occurring May 27 – June 7, 2020.

[R2] Vovak, H., Riddle, T., Taniguchi, T., Hoogesteyn, K., & **Yang, Y.** (2021). Strategies for Policing Innovation (SPI) in Wilmington, Delaware: Targeting Violent Crime. Retrieved from https://www.smart-policing.com/sites/default/files/inline-files/Wilmington%20SPI_Final%20report.pdf.

Public Writing

[P1] **Yang Y.**, & McKnight, J. (2023, Apr 28). Students of Color are Disproportionately Bullied and Harassed at School. Public Health Post. Retrieved Apr 28, 2023, from https://www.publichealthpost.org/databyte/students-of-color-are-disproportionately-bullied-and-harassed-at-school.

[P2] **Yang Y.**, & Vu, C. (2023, Jan 25). The Persisting Racial Inequities in Food Insecurity. Public Health Post. Retrieved Jan 25, 2023, from https://www.publichealthpost.org/databyte/the-persisting-racial-inequities-in-food-insecurity.

[P3] Nsoesie, E. O., & **Yang Y.** (2022, November 22). The high suicide rates among indigenous populations. Public Health Post. Retrieved December 22, 2022, from https://www.publichealthpost.org/databyte/suicide-rates-indigenous.

[P4] **Yang Y.**, & Nsoesie, E. O. (2022, October 25). Are racial differences in flu vaccination rates widening? Public Health Post. Retrieved Oct 25, 2022, from https://www.publichealthpost.org/databyte/race-flu-vaccination-rates.

INVITED TALKS/CONFERENCE PRESENTATIONS (NON-ARCHIVAL)

[T1] **r/23andMe and my pictures: Platformed Racial Spectacle as Collective Racial Identity Formation**, 75th Annual Conference of the International Communication Association (ICA), Denver, CO, 2025.

[T2] **Mapping Real-World Migration Flow Using Context-Rich Social Media Data**, 75th Annual Conference of the International Communication Association (ICA), Denver, CO, 2025.

[T3] **How Social Responses to Online Hate Messages Affect Hatefulness**, 75th Annual Conference of the International Communication Association (ICA), Denver, CO, 2025.

[T4] **Subversive Humor and Platformed Asianness: Resisting West Homogenization and Asian Ultranationalism on Social Media through Satirical Neologism**, 110th National Communication Association's Annual Convention (NCA), New Orleans, LA, 2024.

[T5] **Digitized Genome, Detailed Maps, and Dazzling Selfies: Platformed Racial Spectacle and DNA Test Results Image Sharing on Social Media**, Association of Internet Researchers (AOIR) Doctoral Colloquium, 2024, Remote.

[T6] **Mapping Real-World Migration Flow Network Using Context-Rich Social Media Data** (Poster). 10th International Conference on Computational Social Science (IC^{2}S^2), Philadelphia, PA, 2024.

[T7] **Locating the Asymmetry in Information Flow between Local and National Media on Transgender Discourses**, Local Journalism Researchers Workshop, Durham, NC, 2024.

[T8] **Using Google Street View Images to Examine Systematic Racism in Built Environment**, Strategic Directions Workshop: Cities and Health, Boston University, Boston, MA, 2023.

GRANTS

[G1] "**Mapping Real-World Migration Flow Network Using Context-Rich Social Media Data**", NULab for Texts, Maps, and Networks Seeding Grant ($1,725). Role: Principal Investigator. 2023-2024.

[G2] "**City Equity: Are we making progress on racial equity**", Boston University School of Public Health Investigator Innovation Award ($20k). Role: Contributor. 2023-2024.

HONORS & AWARDS

HASTAC Scholar	2024-2025
Best Paper Reviewer for ASIS&T 2023	*2023*
Burmeister Fellowship, UNC School of Information and Library Science (SILS)	*August 2019*

Best Project Award, 5th Annual UNC SILS Project Fair *April 2019*
China National Scholarship (3 times), Chinese Ministries of Education *2014-2017*

SKILLS

Quantitative Research Methods: Exploratory Data Analysis, Supervised & Unsupervised Machine Learning, Data Mining (Pattern, Time-series, Network), Statistical Modelling (Regressions, CFA/EFA, SEM), Data Visualization

Qualitative Research Methods: Content Analysis, Thematic Analysis, Survey Testing & Administration

Tools/Languages: Python, R, Tableau, MS Office, Git, Arc/QGIS, AWS, Markdown, SQL, JavaScript, HTML/CSS, Gephi, Java, Qualtrics, Figma

SERVICE & LEADERSHIP

President, Network Science Institute Graduate Student Association *2024-2025*
Founder/Leader, Network Science Institute Social Science Club *2024-*
Group Leader, Northeastern University LGBTQIA2S+ Intl. Student Affinity Group *2023-*
Center for Information, Technology, and Public Life (CITAP) Affiliate *2023-*
Peer Reviewer (Full papers/Short papers/Posters/Datasets): ICWSM, CSCW & ASIS&T *2020-2023*
Master Student Representative, Diversity Committee, UNC SILS *2019-2020*
Master Student Mentor, UNC SILS MENTOR Program *2019-2020*