Volha Katebi (Murashka)

San Diego, California

(706) 206-1493 | volhamurashka.health@gmail.com | **GitHub**: https://github.com/volhakatebi |
https://www.linkedin.com/in/volha-katebi/

QUALIFICATIONS

University of Georgia, Athens GA – May 2022
PhD in Communication Studies GPA: 3.93
Specialized in **Health Communication** using
Quantitative-Method Approaches

Ohio University, Athens OH – May 2017
M.A. in Communication and Development;
Graduate Certificate in Global Health GPA: 4.0

Minsk State Linguistic University, Belarus – June 2013
**B.A. in Linguistic Support of Intercultural
Communication** GPA: 3.60

Languages: English (fluent), Russian (native),
Belarusian (native), German (intermediate)

PROFESSIONAL EXPERIENCE

Division of Global Migration and Quarantine, CDC Atlanta, GA
Post-Doctoral Fellow in Health Communication & Marketing August 2022 – Present

- Lead a team of four data scientists in a text classification project using various machine learning and transformer-based models (Logistic Regression; Random Forrest; BERT; DistilBERT; RoBERTa; BERTscore in Python) and explainable AI approaches (SHAP in Python) to identify COVID-19 anti-vaccine content in cruise-related social media messsages and inform health communication efforts for clients in CDC's Maritime Unit. The project has been selected along with the other 3 NLP projects across the whole organization for agency prioritization though the *Data Science Upskilling Program*
- Lead a project on topic modeling using BERTopic and SimCSE/Lonformer with hierarchical clustering methods in Python on Databricks to facilitate faster manual reviewing of qualitative health data
- Collaborate with health communication and evaluation teams to evaluate effectiveness of natural language processing (NLP) approaches for public health projects
- Create reproducible data cleaning, analysis, and report processes using RStudio/RMarkdown and GitLab to enable collaborative time-sensitive reporting of national and international travel patterns using public and private travel and mobility data
- Collaborate with research fellows in optimizing and automating R code to enable rapid-turn response to leadership requests
- Create RShiny and Power BI dashboards to vizualize textual, travel and public health data to inform public health decisions for senior leadership
- Conduct evaluation of user acceptance testing of the national-level public health software
- Develop and lead training workshops for research fellows on topics such as managing conda environments and using cluster computing
- Organize and lead study group on mathematical modeling in public health for research fellows

The Communication, Health and Emerging Media Lab, University of Georgia Athens, GA
Research Affiliate July 2022-Present

- Lead an NLP project on identifying misinformation in discussions about electronic cigarettes using unsupervised (BERTopic) and supervised (BERT; BERTweet) machine learning approaches (data collection using Twitter API; codebook development; training text annotators; testing algorithms; data visualization)

Division of Global Migration and Quarantine, CDC Atlanta, GA
ORISE Fellow in Data Science July 2021 – July 2022

- Led a a team of data scientists in a topic modeling project using Latent Dirichlet Allocation in R focusing on online users' messages about traveling during COVID-19 to inform public health interventions targeted at travelers; conducted trend analysis and sentiment analysis (transformer-based; LIWC; dictionary-based syuzhet) of topics

- Provided NLP support (sentiment analysis; text data visualization in R) to the evaluation team to understand general insights on public health initatives from qualitative interview data
- Provided data analysis and visualization support to project leads on modeling the spread of infectious diseases and developing public health metrics using R, RShiny and Power BI; reviewed public health reports to identify relevant model parameters; extracted relevant social media data using R and Python
- Designed communication materials in R Markdown for dissemination across the division
- Organized and led weekly sessions on data science in public health for research fellows

Department of Communication Studies, University of Georgia   Athens, GA
Graduate Reseach Assistant   March – April 2021
- Conducted topic modeling in Python to understand online users' reactions to fitness-promoting content on social media (proposal development; data collection and analysis; summary reports)
- Managed, instructed, and collaborated with faculty and research assistants in eye-tracking study design, data collection, data analysis, and report

Department of Communication Studies, University of Georgia   Athens, GA
Teaching Assistant, Health Communication   Spring 2018, Spring 2020
- Developed and co-instructed course 2 days per week (1 class section, 75 students) designed to help students understand how media and society exert an influence on health-related behaviors

Pan American Health Organization|World Health Organization   Washington, D.C.
Intern   May 2016 – August 2016
- Conducted the systematic review of opinions of policymakers on public health issues (screened 200+ articles); created health messages for prenatal care intervention campaigns

National Academy of Sciences of Belarus, Research Center for Food Products   Minsk, Belarus
Communication Specialist   July 2013 – July 2015
- Coordinated implementation of international projects supported by the European Commission; systematized and presented research findings on food industry technologies

RELEVANT CERTIFICATION
- Certified SPARK NLP Data Scientist from John Snow Labs
- Alumna of CDC's Data Science Upskilling 10-month program (September 2022 – July 2023)

SELECTED PUBLICATIONS & PRESENTATIONS

Dissertation: *Understanding How Message Features of Fitspiration Posts Affect Body Image Perceptions and Exercise Intentions*

Katebi, V., O'Sullivan, T., Panasci, A., Lan, K. (2023, July). Detection of Misinformation about COVID-19 Vaccines in Cruise-Related Tweets Using Transformer-Based Models.Presented at Data Science Upskilling Symposium at Centers for Disease Control and Prevention.

Katebi, V. & Lan, K. (2023, July). Topic Modeling Best Practices on EDAV Databricks. Presented at EDAV Workgroup at Centers for Disease Control and Prevention.

Murashka, V., Liu, J., & Peng, Y. (2021). Fitspiration on Instagram: Identifying topic clusters in user comments to posts with objectification features. *Health Communication*, 1-12
Media coverage: ABC WBAY https://tinyurl.com/43zvej7p

Murashka, V. & Liu, J. (2018, September). Identifying topic clusters in comments to gender-specific fitness images on Instagram. Presented at the National Conference on Health Communication and Media, Atlanta, GA.