Cyprus University of Technology

Department of Electrical Engineering and Computer Engineering and Informatics

Bachelor thesis

Retrieving and Analyzing Toxic and Polarized Discourse on BlueSky: A Study of Decentralized Social Network

Christodoulos Stavrou

Limassol, May 2025

CYPRUS UNIVERSITY OF TECHNOLOGY

Faculty of Engineering and Technology

Department of Electrical Engineering, Computer Engineering, and Informatics

Bachelor thesis

Retrieving and Analyzing Toxic and Polarized Discourse on BlueSky: A Study of Decentralized Social Network

Christodoulos Stavrou

Advisor:

Dr Michael Sirivianos Dean of the School of Engineering and Technology

Limassol, May 2025

Copyrights

Copyright © 2025 Christodoulos Stavrou

All rights reserved.

The approval of the dissertation by the Department of Electrical Engineering, Computer Engineering, and Informaticsdoes not necessarily imply the approval by the Department of the views of the writer.

Acknowledgements

I would like to express my deepest gratitude to Dr. Nikos Salamanos for his valuable guidance, expertise, and mentorship throughout the development of this thesis.

I am also sincerely thankful to Pantelitsa Leonidou for her continued support, encouragement, and input during this project.

Their contributions and feedback were instrumental in shaping the research and analysis presented in this thesis.

ABSTRACT

The proliferation of hate speech on decentralized social media platforms, such as BlueSky, presents a growing challenge for ensuring healthy online discourse. Unlike traditional centralized networks, decentralized platforms lack a single governing authority, making moderation and content regulation more complex. This thesis focuses on the retrieval and analysis of data from BlueSky to better understand the nature and spread of hate speech and polarization within such decentralized ecosystems. By collecting and processing real-world data from the platform(13699 posts, 58788 users), we explore patterns of information dissemination and evaluate the potential of applying machine learning approaches—such as Federated Learning (FL) and Graph Neural Networks (GNN), for future disinformation detection in decentralized contexts.

Keywords: Federated learning, decentralized social media, data retrieval, data analysis.