Modeling Quarantine Intervention for Varied Toxic Intensities

Prof. Dr. Nitin Agarwal[1,2]
[1]COMOS Research Center, University of Arkansas at Little Rock
[2]International Computer Science Institute, University of California, Berkeley
nxagarwal@ualr.edu

The Fifteenth International Conference on Social Media Technologies, Communication, and Informatics (SOTICS 2025)
September 28, 2025 to October 02, 2025 - Lisbon, Portugal

Author Biographies

Prof. Dr. Nitin Agarwal, Ph.D. | Director, COSMOS Research Center

- Jerry L. Maulden-Entergy Chair & Donaghey Distinguished Professor, UA-Little Rock.
- Faculty Fellow, ICSI, University of California, Berkeley.
- IARIA Fellow, ARA Fellow, AAOC Fellow, IEEE Senior Member.

Introduction

- Wildfire-like spread of toxicity on social media makes it difficult to manage
- Users exposed such toxic contents amplifies its reach, thus affecting more individuals
- Understanding how toxicity spreads on social media is essential to mitigate its negative impact.
- Epidemiological models, originally used for disease spread, are now being applied to analyze toxicity propagation online.
- This study extends the $SEIDR$ model to examine the impact of toxicity intensity on the spread of toxic content on social media **by incorporating quarantine (Q) intervention**.

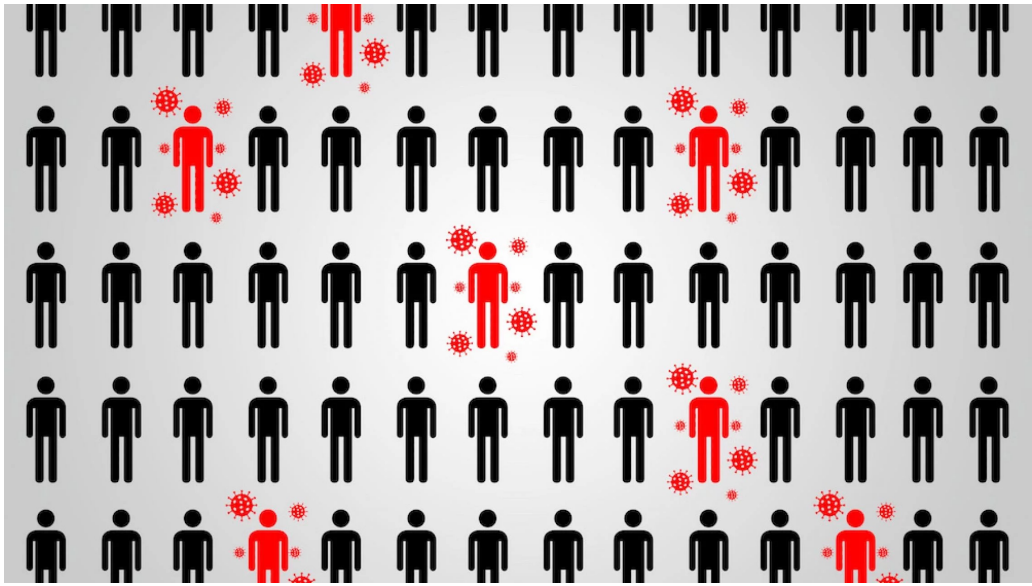
Research Questions

RQ1:

- Do variations in toxicity levels influence its dissemination?
- Does the degree of toxicity intensity impact the model's performance?

RQ2:

- What are the key parameters in the SEIQR model that have the most significant impact on the effective reproduction number R0?
- How can these parameters be examined to control the spread of toxicity on social media?

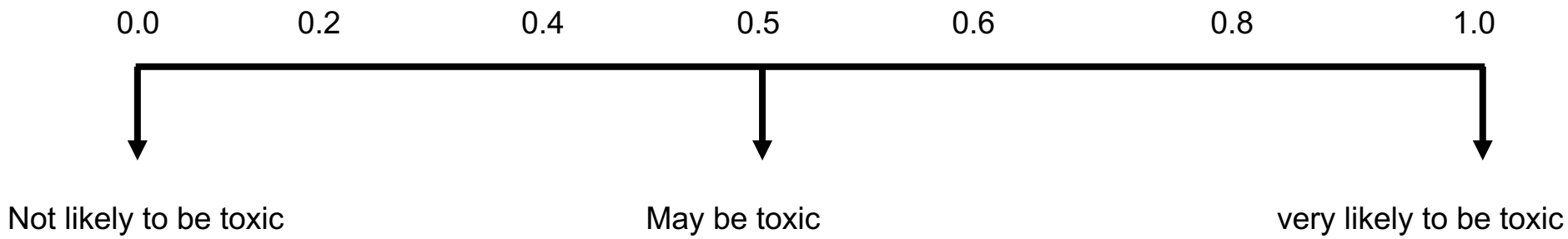


Data Collection

- **Data Collection:**
  - Used Twitter Academic APIs to gather data from two areas: COVID-19 and social movements.
  - COVID-19 data covers discussions from February 2020 to June 2021. The datasets are:
    - **Anti Lockdown, 5G, Anti covid, Anti mask, 5G**
  - Social movement data focuses on protests in Brazil from November 1, 2022, to February 25, 2023. The datasets are:
    - **Brazil anti government, Brazil Pro government**
  - Social movement data focuses on **Peru**, anti-government protests erupted after President Pedro Castillo was removed from power on December 7,2022, resulting in violent clashes.

Toxicity Score Assessment

- Detoxify model is applied to calculate the toxicity score for each dataset.



- Infected individual categories:
  1. the first group comprises individuals with a toxicity score greater than the average toxicity score of the infected population.
  2. The second group consists of individuals with scores between 0.5 and the average toxicity score.

Data Overview

- The table shows the statistical information regarding seven datasets.

Dataset	Num. of Posts	Avg. Toxicity	Num. of High Toxic	Num. of Moderate Toxic
F*Covid	28,131	0.91	4,684	2,082
F*Mask	2,423	0.91	538	217
F*Lockdown	1,995	0.82	598	493
5G	33,403	0.84	1,096	703
Brazil Anti	405,160	0.70	1,221	2,309
Brazil Pro	44,415	0.75	105	131
Peru	195,290	0.71	511	546

Epidemiological Modeling

- **Epidemiological Model**

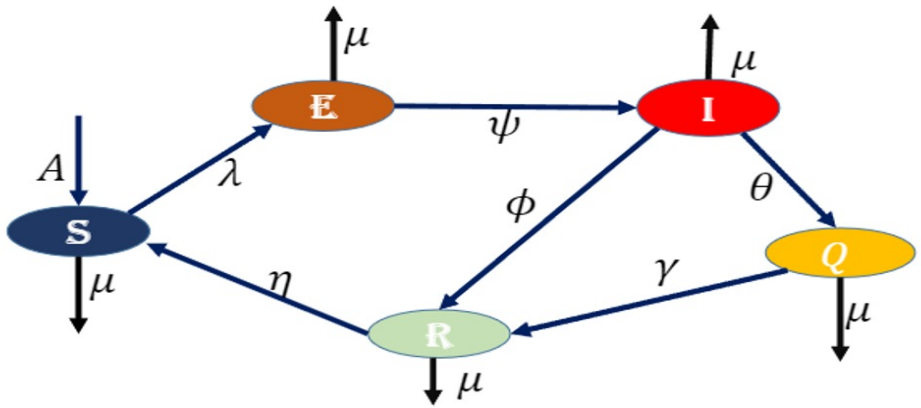
SEIQR (Susceptible-Exposed-Infected-Quarantine-Recovered)

$$\begin{cases} \frac{dS(t)}{dt} = \mathcal{A} + \eta R - \frac{\beta IS}{N(t)} - \mu S, \\ \frac{dE(t)}{dt} = \frac{\beta IS}{N(t)} - (\mu + \psi)E, \\ \frac{dI(t)}{dt} = \psi E - (\mu + \phi + \theta)I, \\ \frac{dQ(t)}{dt} = \theta I - (\gamma + \mu)Q, \\ \frac{dR(t)}{dt} = \gamma Q + \phi I - (\mu + \eta)R, \end{cases}$$

where $\lambda = \frac{\beta IS}{N(t)}$, the entire population, we define the quantity $N(t)$ by

$$N(t) = S(t) + E(t) + I(t) + Q(t) + R(t).$$

Parameter	Explanation
$\mathcal{A}$	recruitment rate of human
β	effective contact rate
ψ	the rate at which exposed become infected
θ	the rate at which I transfer to quarantine class
η	recovery rate
γ	the rate at which Q transfer to recovery
ϕ	the rate at which I transfer to recovery
μ	the rate at which people exit autonomously



Error Rate

- Used nonlinear least-squares curve-fitting tool for model fitting.
- Calculated error rates to evaluate model performance.

$$Error\ Rate = \frac{\left\| Observed - Predicted \right\|_2}{\left\| Predicted \right\|_2}$$

Results (RQ1)

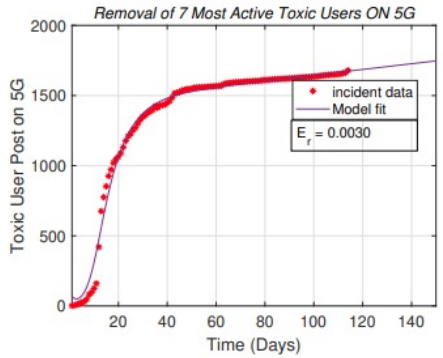
- Comparison of error rates for various epidemiological models applied to seven different toxicity datasets.
- The models evaluated include the SIS, SIR, SIRS, SEIRS, and SEIQR models.

Error Rates for Model Variations					
Dataset	**SIS-Model**	**SIR**	**SIRS**	**SEIRS**	**SEIQR**
F*Covid	0.2001	0.21	0.174	0.18	**0.081**
F*Mask	0.263	0.09	0.025	0.11	**0.073**
Lockdown	0.248	0.188	0.106	0.15	**0.061**
5G	0.526	0.101	0.097	0.041	**0.03**
Brazil Anti	0.338	0.126	0.066	0.09	**0.094**
Brazil Pro	0.533	0.089	0.076	0.08	**0.088**
Peru	0.404	0.445	0.385	0.432	**0.41**

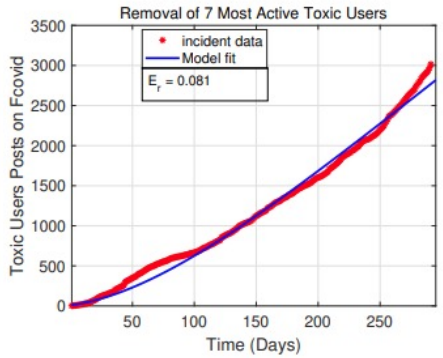
Results (RQ1)

- High error rates across the various datasets without split.
- The model was not able to capture the nuances and complexities in the data.

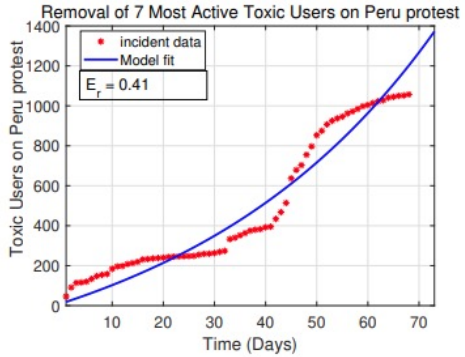
Error rates for SEIQR without datasets split	
Dataset	**Error rate**
F*Covid	0.081
F*Mask	0.073
Lockdown	0.061
5G	0.03
Brazil Anti	0.094
Brazil Pro	0.088
Peru	0.41



(a) 5G dataset

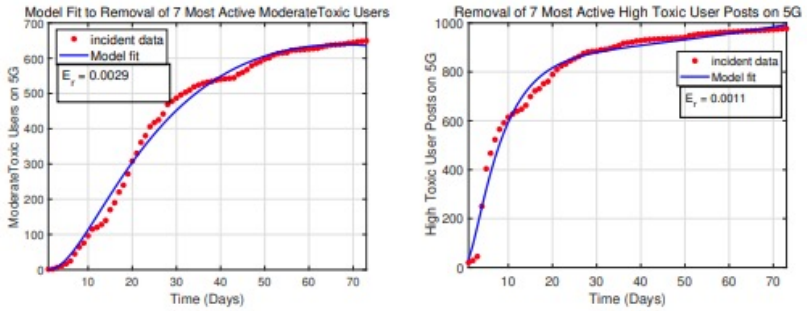


(b) F*covid-19 dataset

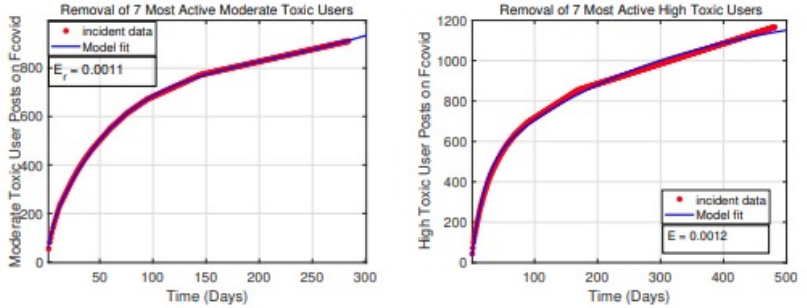


(c) Peru protest dataset

Results (RQ1)

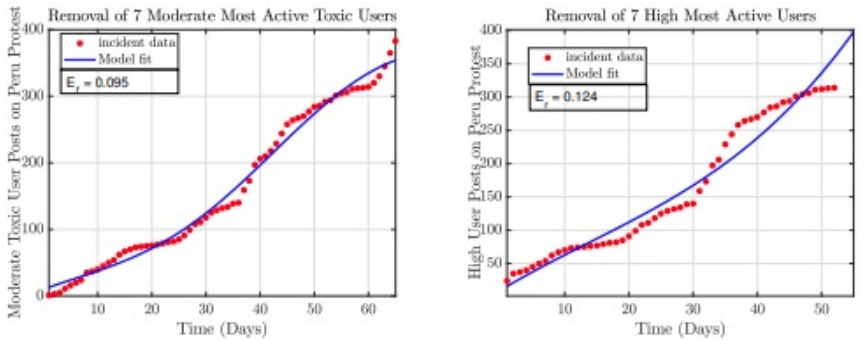


(a) 5G data (left: moderate toxic, right: high toxic).



(b) F*Covid data (left: moderate toxic, right: high toxic).

- Low error rates across the various datasets with split.
- The model was able to capture the nuances and complexities in the data.

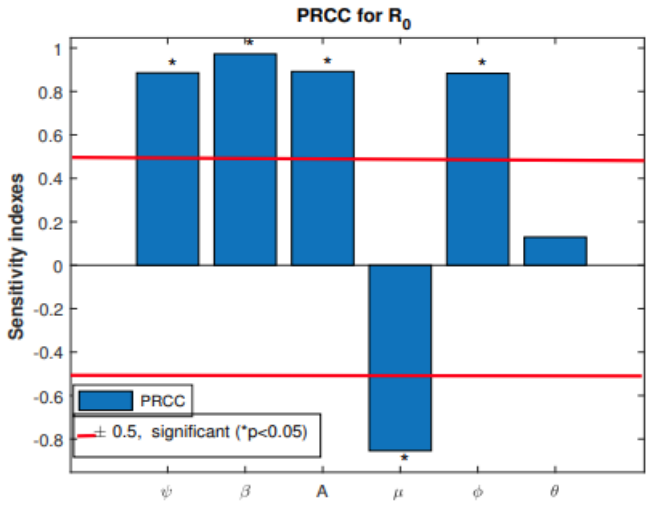


(c) Peru protest data (left: moderate toxic, right: high toxic).

Error rates for SEIQR with datasets split		
Dataset	**Moderate Toxicity**	**High Toxicity**
F*Covid	0.0011	0.0021
F*Mask	0.021	0.049
Lockdown	0.032	0.045
5G	0.0029	0.0011
Brazil Anti	0.062	0.055
Brazil Pro	0.060	0.061
Peru	0.095	0.124

Results (RQ2)

- The sensitivity analysis using the Latin Hypercube Sampling-Partial Rank Correlation Coefficient (LHS-PRCC) method identified parameters φ, β, ψ, A, and θ as having a positive influence on the reproduction number R_0, while parameter μ has a negative effect.
- This analysis shows that specific parameters significantly impact the spread of online toxicity, providing insights for effective management strategies.



Sensitivity of $\mathcal{R}_0$ of the online toxicity contagious

Key Findings

- **Improved Model Accuracy**: Categorizing toxicity based on intensity, we obtained a lower error rate, meaning we were able to capture the propagation of toxicity more accurately.
- **Parameters Impact on the Model**: We carried out a sensitivity analysis to determine the important model parameters that significantly affect the dynamics of toxic spread on social media platforms.
- **Future research** will further refine and extend the $SEIQR$ model to cross-platform validation since each of the platforms has their own algorithm to detect toxicity.

Opportunities @ COSMOS

- **5-year $15 million grant from the U.S. Department of Defense, PI: Nitin Agarwal, Ph.D.**
- 20+ Graduate Research Assistants (MS/PhD)
  ($27k, Full Tuition coverage, Health Insurance covered)
- 10+ Postdoctoral fellowships
  ($60k - $110k, Full benefits, Health insurance)
- 10+ Full stack developer, Dev Ops, Data engineer
  ($60k - $100k, Full benefits, Health insurance)
- 2 System Engineer
  ($65k, Full benefits, Health insurance)

Apply at https://cosmos.ualr.edu/careers/
Contact Prof. Dr. Nitin Agarwal (nxagarwal@ualr.edu)

Thank You!

COSMOS Tools Developed:

- **Blogtrackers** - https://btracker.host.ualr.edu
- **YouTubeTracker** - https://vtracker.host.ualr.edu
- **Focal Structure Analysis** - http://fsa.host.ualr.edu/
- **COVID-19 Misinfo Tracker** - https://cosmos.ualr.edu/covid-19

Blogtrackers tracker FOCAL STRUCTURE ANALYSIS

Follow **cosmographers** on https://cosmos.ualr.edu/

Nitin Agarwal, Ph.D.
Director, COSMOS Research Center
nxagarwal@ualr.edu

ACKNOWLEDGEMENTS

This research is funded in part by the U.S. National Science Foundation (OIA-1946391, OIA-1920920), U.S. Office of the Under Secretary of Defense for Research and Engineering (FA9550-22-1-0332), U.S. Army Research Office (W911NF-23-1-0011, W911NF-24-1-0078, W911NF-25-1-0147), U.S. Office of Naval Research (N00014-21-1-2121, N00014-21-1-2765, N00014-22-1-2318), U.S. Air Force Research Laboratory, U.S. Defense Advanced Research Projects Agency, the Australian Department of Defense Strategic Policy Grants Program, Arkansas Research Alliance, the Jerry L. Maulden/Entergy Endowment, and the Donaghey Foundation at the University of Arkansas at Little Rock. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations. The researchers gratefully acknowledge the support.