**REFERENCES**

[1] J. Ronson, So You’ve Been Publicly Shamed. Picador, 2015.

[2] E. Spertus, “Smokey: Automatic recognition of hostile messages,” in

AAAI/IAAI, 1997, pp. 1058–1065.

[3] S. Sood, J. Antin, and E. Churchill, “Profanity use in online communities,”

in Proceedings of the SIGCHI Conference on Human Factors in

Computing Systems. ACM, 2012, pp. 1481–1490.

[4] S. Rojas-Galeano, “On obstructing obscenity obfuscation,” ACM Transactions

on the Web (TWEB), vol. 11, no. 2, p. 12, 2017.

[5] E. Wulczyn, N. Thain, and L. Dixon, “Ex machina: Personal attacks

seen at scale,” in Proceedings of the 26th International Conference on

World Wide Web. International World Wide Web Conferences Steering

Committee, 2017, pp. 1391–1399.

[6] A. Schmidt and M. Wiegand, “A survey on hate speech detection using

natural language processing,” in Proceedings of the Fifth International

Workshop on Natural Language Processing for Social Media. Association

for Computational Linguistics, Valencia, Spain, 2017, pp. 1–10.

[7] Hate-Speech, “Oxford dictionaries,” retrieved August 30, 2017 from

https://en.oxforddictionaries.com/definition/hate speech.

[8] W. Warner and J. Hirschberg, “Detecting hate speech on the world wide

web,” in Proceedings of the Second Workshop on Language in Social

Media. Association for Computational Linguistics, 2012, pp. 19–26.

[9] I. Kwok and Y. Wang, “Locate the hate: Detecting tweets against blacks.”

in AAAI, 2013.

[10] P. Burnap and M. L. Williams, “Cyber hate speech on twitter: An

application of machine classification and statistical modeling for policy

and decision making,” Policy & Internet, vol. 7, no. 2, pp. 223–242,

2015.

[11] Lee-Rigby, “Lee rigby murder: Map and timeline,” retrieved December

07, 2017 from https://http://www.bbc.com/news/uk-25298580.

[12] Z. Waseem and D. Hovy, “Hateful symbols or hateful people? predictive

features for hate speech detection on twitter.” in SRW@ HLT-NAACL,

2016, pp. 88–93.

[13] P. Badjatiya, S. Gupta, M. Gupta, and V. Varma, “Deep learning for hate

speech detection in tweets,” in Proceedings of the 26th International

Conference on World Wide Web Companion. International World Wide

Web Conferences Steering Committee, 2017, pp. 759–760.

[14] D. Olweus, S. Limber, and S. Mihalic, “Blueprints for violence

prevention, book nine: Bullying prevention program,” Boulder, CO:

Center for the Study and Prevention of Violence, 1999.

[15] P. K. Smith, H. Cowie, R. F. Olafsson, and A. P. Liefooghe, “Definitions

of bullying: A comparison of terms used, and age and gender differences,

in a fourteen–country international comparison,” Child development,

vol. 73, no. 4, pp. 1119–1133, 2002.

[16] R. S. Griffin and A. M. Gross, “Childhood bullying: Current empirical

findings and future directions for research,” Aggression and violent

behavior, vol. 9, no. 4, pp. 379–400, 2004.

[17] H. Vandebosch and K. Van Cleemput, “Defining cyberbullying: A

qualitative research into the perceptions of youngsters,” CyberPsychology

& Behavior, vol. 11, no. 4, pp. 499–503, 2008.

[18] H. Vandebosch and K. Van Cleemput, “Cyberbullying among youngsters:

Profiles of bullies and victims,” New media & society, vol. 11, no. 8,

pp. 1349–1371, 2009.

[19] K. Dinakar, B. Jones, C. Havasi, H. Lieberman, and R. Picard,

“Common sense reasoning for detection, prevention, and mitigation

of cyberbullying,” ACM Transactions on Interactive Intelligent Systems

(TiiS), vol. 2, no. 3, p. 18, 2012.

[20] P. Singh, T. Lin, E. T. Mueller, G. Lim, T. Perkins, and W. L. Zhu,

“Open mind common sense: Knowledge acquisition from the general

public,” in OTM Confederated International Conferences” On the Move

to Meaningful Internet Systems”. Springer, 2002, pp. 1223–1237.

[21] H. Hosseinmardi, S. A. Mattson, R. I. Rafiq, R. Han, Q. Lv, and

S. Mishra, “Detection of cyberbullying incidents on the instagram social

network,” arXiv preprint arXiv:1503.03909, 2015.

[22] J. Cheng, C. Danescu-Niculescu-Mizil, and J. Leskovec, “Antisocial

behavior in online discussion communities.” in ICWSM, 2015, pp. 61–70.

[23] J. Cheng, C. Danescu-Niculescu-Mizil, J. Leskovec, and M. Bernstein,

“Anyone can become a troll,” American Scientist, vol. 105, no. 3, p. 152,

2017.

[24] P. Tsantarliotis, E. Pitoura, and P. Tsaparas, “Defining and predicting

troll vulnerability in online social media,” Social Network Analysis and

Mining, vol. 7, no. 1, p. 26, 2017.

[25] S. O. Sood, E. F. Churchill, and J. Antin, “Automatic identification of

personal insults on social news sites,” Journal of the Association for

Information Science and Technology, vol. 63, no. 2, pp. 270–285, 2012.