

20 Years of Analyzing Multilingual Propaganda Content on the Web

Mark Last

Department of Software and Information Systems Engineering, Ben-Gurion University of the Negev, Beer Sheva, Israel

Abstract

From the early years of the Internet as a global information infrastructure, multilingual propaganda content has been circulating on the web. For example, the WWW played a critical role in the planning of the 9/11 attacks, both as a source of inspirational information and as a safe means of covert communication between the plotters. Shortly after the tragic events of 2001, our binational team of US and Israeli researchers started to explore the online activities of various hate groups. Initially, we developed a prototype of a monitoring system aimed at detecting the frequent visitors of terrorist websites, which could be influenced by terrorist propaganda and eventually develop into what we call today “the lone wolf attackers”. Shortly after, we focused on another, closely related question: what makes terrorist-generated propaganda content in various languages different from unbiased news reports discussing similar topics? Over the years, we developed prototypes of several additional text analysis tools such as text-summarization algorithms, which can automatically summarize large amounts of untranslated content in any language, as well as AI tools for automated detection of metaphoric language. After presenting the motivational and ethical foundations of our research, I plan to describe some of the methods developed during the last two decades and finally, discuss past and future challenges in this important and fascinating domain.

Short Bio

Prof. Mark Last is a Full Professor at the Department of Software and Information Systems Engineering, Ben-Gurion University of the Negev, Israel and the Head of the Data Engineering Program. Prof. Last has published over 210 peer-reviewed papers, two monographs, and 11 edited volumes on data mining, text mining, and cyber security. According to Google Scholar, his works were cited more than 6,000 times. He is a Senior Member of the IEEE Computer Society and a Professional Member of the Association for Computing Machinery (ACM). Prof. Last currently serves as an Action Editor of Data Mining and Knowledge Discovery and an Editorial Board Member of Machine Learning Journal and ACM Transactions on Intelligent

In: M. Litvak, I.Rabaev, R. Campos, A. Jorge, A. Jatowt (eds.): Proceedings of the IACT'23 Workshop, Taipei, Taiwan, 27-July-2023

✉ mlast@bgu.ac.il (M. Last)

ORCID [0000-0003-0748-7918](https://orcid.org/0000-0003-0748-7918) (M. Last)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

Systems and Technology. Previously, he has served as an Associate Editor of IEEE Transactions on Systems, Man, and Cybernetics - Part C (2004–2012), Pattern Analysis and Applications (2007- 2016), and IEEE Transactions on Cybernetics (2013-2019). His main research interests are focused on data mining, cross-lingual text mining, soft computing, cyber intelligence, and medical informatics.