COLING 2018

The 27th International Conference on Computational Linguistics

Proceedings of the First Workshop on Natural Language Processing for Internet Freedom (NLP4IF-2018)

August 20, 2018
Santa Fe, New Mexico, USA

Copyright of each paper stays with the respective authors (or their employers).

ISBN 978-1-948087-58-2

Introduction

Welcome to the First Workshop on NLP for Internet Freedom. Our workshop has been inspired by the recent report produced by Freedom House (freedomhouse.org), an “independent watchdog organization dedicated to the expansion of freedom and democracy around the world", which states that Internet freedom declined in 2016 for the sixth consecutive year. 67% of all Internet users live in countries where criticism of the government, military, or ruling family are subject to censorship. Social media users face unprecedented penalties, as authorities in 38 countries made arrests based on social media posts over the past year. Globally, 27% of all internet users live in countries where people have been arrested for publishing, sharing, or merely “liking" content on Facebook. Governments are increasingly going after messaging apps like WhatsApp and Telegram, which can spread information quickly and securely. Various barriers exist to prevent citizens of a large number of countries to access information. Some involve infrastructural and economic barriers, others violations of user rights such as surveillance, privacy and repercussions for online speech and activities such as imprisonment, extralegal harassment or cyberattacks. Yet another area is limits on content, which involves legal regulations on content, technical filtering and blocking websites, (self-)censorship. Large internet providers are effective monopolies, and themselves have the power to use NLP techniques to control information flow. Users are suspended or banned, sometimes without human intervention, and with little opportunity for redress. Users react to this by using coded, oblique or metaphorical language, by taking steps to conceal their identity such as the use of multiple accounts, raising questions about who the real originating author of a post actually is. This workshop brings together NLP researchers whose work contributes to the free flow of information on the Internet.

The papers in this volume all focus on censorship in China. Heng Ji and Kevin Knight discuss obfuscated language that people often create to avoid Internet censorship. The authors also give an overview of automated techniques needed to simulate human encoding. Knockell et al. conduct an in-depth study of blacklists and their variability across and within domains. Ng et al. discuss the linguistic properties of censored and uncensored social media posts.

Thanks to the US National Science Foundation support (award No. 1828199), we were able to bring in three speakers from the NLP community and beyond. Dr. Jedidiah Crandall (University of New Mexico) will give a talk entitled “How to Talk Dirty and Influence Machines". Dr. Jennifer Pan (Stanford University) will give a talk on how the Chinese government fabricates social media posts for strategic distraction, not engaged argument. Nancy Watzman (Dot Connector Studio) will discuss what journalists really want from NLP researchers and how to help build trust in media and democracy by helping journalists make sense of big data. Finally, we will also hold a panel session on NLP and Disinformation.

Last but not least, we would like to thank the program committee for their help with reviewing the papers, organizing and advertising the workshop.

See you in Santa Fee on August 20!

Chris Brew, Anna Feldman, and Chris Leberknight

Organizers:

Chris Brew
Anna Feldman, Montclair State University
Chris Leberknight, Montclair State University

Program Committee:

Joan Bachenko, Deception Discovery Technologies, NJ, USA
Jedidiah Crandall, University of New Mexico, NM, USA
Chaya Hiruncharoenvate, Mahasarakham University, Thailand
Lifu Huang, Rensselaer Polytechnic Institute (RPI), NY, USA
Zubin Jelveh, The University of Chicago, IL, USA
Judith Klavans, Columbia University, NY, USA
Jeffrey Knockel, University of New Mexico, NM, USA
Will Lowe, Princeton University, NJ, USA
Rada Mihalcea, University of Michigan, Ann Arbor, MI, USA
Prateek Mittal, Princeton University, NJ, USA
Rishab Nithyanand, Data & Society, NY, USA
Noah Smith, University of Washington, WA, USA
Thamar Solorio, University of Houston, TX, USA
Mahmood Sharif, Carnegie Mellon University, PA, USA
Evan Sultanik, Trail of Bits, NY, USA
Svitlana Volkova, Pacific Northwest National Laboratory, WA, USA
Brook Wu, New Jersey Institute of Technology, NJ, USA

Invited Speakers:

Dr. Jedidiah Crandall, University of New Mexico
Dr. Jennifer Pan, Stanford University, CA
Nancy Watzman, Dot Connector Studio

Table of Contents

The effect of information controls on developers in China: An analysis of censorship in Chinese open source projects
Jeffrey Knockel, Masashi Crete-Nishihata and Lotus Ruan 1

Linguistic Characteristics of Censorable Language on SinaWeibo
Kei Yin Ng, Anna Feldman, Jing Peng and Chris Leberknight 12

Creative Language Encoding under Censorship
Heng Ji and Kevin Knight 23

Conference Program

09:00–10:00 **Invited talk: Jennifer Pan (Stanford University):** ***How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument***

10:00–10:30 *The Effect of Information Controls on Developers in China: An Analysis of Censorship in Chinese Open Source Projects*
Jeffrey Knockel, Masashi Crete-Nishihata and Lotus Ruan

10:30–11:00 **Coffee Break**

11:00–12:00 **Invited talk: Jed Crandall (University of New Mexico):** ***How to Talk Dirty and Influence Machines***

12:00–12:30 *Linguistic Characteristics of Censorable Language on SinaWeibo*
Kei Yin Ng, Anna Feldman, Jing Peng and Chris Leberknight

12:30–02:00 **Lunch**

02:00–03:00 **Invited Talk: Nancy Watzman (Dot Connector Studio):** ***What do Journalists Really Want from NLP Researchers? How to Help Build Trust in Media and Democracy by Helping Journalists Make Sense of Big Data***

03:00–03:30 *Creative Language Encoding under Censorship*
Heng Ji and Kevin Knight

03:30–04:00 **Coffee Break**

04:00–05:00 **Panel: NLP and Disinformation (Moderator: Chris Brew)**