[SORSE] Towards an NLP Pipeline for Conflict Narrative Detection as well as Inclusive Definition and Use of Terms for the US RSE Community

Report of Contributions

Contribution ID: **1**   Type: **not specified**

Towards an NLP Pipeline for Conflict Narrative Detection

Thursday, October 22, 2020 2:00 PM (30 minutes)

This talk is about PhD research into developing an NLP pipeline for Conflict Narrative Detection. In response to increased incidences of online abuse, a new industry of hate speech detection using NLP has emerged. Accordingly, we tested NLP technologies used by this industry to discover how quantitatively analysing language distorts meaning. We compiled a dataset comprising "Mein Kampf" from Hitler, "War on Terror" texts from George Bush and Osama bin Laden, and in how he advocated for non-violence, speeches from Martin Luther King provide control data. We tested both general-purpose and state-of-the-art sentiment analysis technologies from TextBlob, Google and IBM. Where distinctive results would be expected from a dataset of extremes, our tests show that regardless of technical sophistication, these technologies are unable to distinguish abusive from non-abusive texts. We address this problem with quantitatively analysing language by offering Conflict Narrative Detection as a new approach.

Using a series of experiments published on GitHub, participants will learn about developing a sociotechnical pipeline to detect conflict narratives using the spaCy NLP python library. "Conflict narrative" means a narrative produced by an orator who intends to legitimise violence against their outgroup. Accordingly, guiding technical design is the theory of "cultural violence" from Peace Research, which explores processes of violence legitimisation. Detecting a conflict narrative means inferring what cultural violence calls the "Self-Other Gradient". What follows is a hypothesis whereby the steeper the Self-Other Gradient in favour of an orator's ingroup, the more legitimate acts of violence against their outgroup become. To infer this gradient, we move beyond quantitatively analysing language by employing qualitative methods, such as hypernymy. Accordingly, qualitative data produced by the pipeline represent the language patterns used to legitimise violence. Participants will learn how the Self-Other Gradient and these language patterns provide new data for tackling online abuse.

Presenter: ANNING, Stephen (University of Southampton)

Session Classification: Talks

Contribution ID: **2**   Type: **not specified**

The Continual Road to an Inclusive Definition and Use of Terms for the US Research Software Engineer Community

Thursday, October 22, 2020 2:30 PM (15 minutes)

An important part of building a strong and resilient community is fostering collaboration between individuals with varying backgrounds, expertise, and viewpoints and building a diverse and inclusive community. While the term Research Software Engineering was initially suggested in 2010, there are many different definitions of Research Software Engineers (RSEs) and their community. The US-RSE Association uses "We like an inclusive definition of Research Software Engineers to encompass those who regularly use expertise in programming to advance research." One of the first community events of the US-RSE Association was a Birds-of-Feather session at PEARC19, where the discussion revealed that participants wanted to learn if the RSE community could include them, and questions arose who would be interested in such an association and community beyond RSEs themselves. These discussions have continued since then and we realized that any definition of RSE will probably change over time.

The mission of US-RSE focuses on three areas: 1) Creating a professional community to share knowledge, connections, and resources; 2) Promoting RSEs' impact on research, highlighting the critical and valuable role RSEs serve; and 3) Providing access to information and material to support individuals and RSE groups. The main focal point is the community and we used it in the name of the association: RSE in US-RSE is "Research Software Engineer".

In addition to naming the association, creating a code of conduct, and announcing and organizing events, reaching out to the full, diverse community needs a careful selection of terms in order to be inclusive and make everyone feel welcome as part of the community. The talk will go into detail about how we have been approaching inclusivity. Our goal is to actively discuss and raise more and more awareness of the topic to improve inclusivity in our outreach and in the community in general.

Presenters: GESING, Sandra (University of Notre Dame); Dr KATZ, Daniel S. (University of Illinois at Urbana-Champaign, US); Dr PARSONS, Lance; Dr FERENBAUGH, Charles; Dr COSDEN, Ian

Session Classification: Talks