

Michael Baldassaro
Digital Threats Project Lead
The Carter Center

Prepared for The Workshop on
Comparative Approaches to Disinformation
Harvard Law School
October 4, 2019

Identifying and Tracking Disinformation during the May 2019 South Africa Elections

Analytical and Methodological Lessons from a Carter Center Pilot Mission

THE
CARTER CENTER



Context

- On May 8, 2019, South Africans voted in the sixth general elections since the end of Apartheid.
- Past elections have been largely credible, but a few recent high-profile coordinated disinformation campaigns to distort opinion heightened concerns ahead of 2019 elections.
- Through desktop research and interviews, TCC learned of a binary domain classifier created and maintained by Media Monitoring Africa (MMA) to evaluate news sources.
- TCC also identified Twitter as the platform on which the bulk of political discourse takes place and which to identify and track possible mis / disinformation.
 - Note: although more South Africans use Facebook (16M±) than Twitter (8M±), interviewees noted that Twitter was the predominant platform for political discourse. This is reinforced by the larger numbers of followers of parties and candidates on Twitter than Facebook.



Methodology

- The Carter Center mission focused on three main goals:
 - Identify possible sources of mis / disinformation (“dodgy” domains) that have entered the “political mainstream”
 - Assess the lifespan and potential reach of “dodgy” news links
 - Identify possible computational propaganda to amplify mis / disinformation.
- Between April 8 and May 8, 2019, TCC scraped 379,877 tweets from the “political mainstream” using the Twitter API.
- Researchers extracted domains shared using custom R code and used the KnowNews browser extension to identify “dodgy” domains.



Findings

- 32,202 tweets contained URLs with 884 distinct domains
- 10 “dodgy” news sources entered the “political mainstream”
- Extracted additional 13,373 tweets searching by “dodgy” domains
- 608 unique links were shared from “dodgy” news sources
- 161,206 total tweets + retweets of “dodgy” news links
 - Note: 151K+ tweets + retweets of one dodgy source: southafricatoday.net
- Potential reach of “dodgy” news links: 10,756,201 followers
 - Note: Proxy count - includes possible duplication of follower accounts



Findings

- Although unable to verify the accuracy of each news article, researchers identified 21 suspicious links that may contain mis / disinformation that could distort public opinion.
 - One “dodgy” news link alleging white contractors were sabotaging state-run electrical stations to obtain repair contracts gained the most traction after it was retweeted by a senior ruling party official (African National Congress Director of Elections and former Deputy Minister of Police Fikile Mbalula) to his more than 1.6 million followers.
 - The “dodgy” news link, which cited unnamed sources, was tweeted + retweeted 4,000+ times between April 9 and May 6 (two days before Election Day), with an estimated possible reach of nearly 1.8 million Twitter followers



Findings

- Links from the “dodgy” news source southafricatoday.net - identified by researchers as a hyper-partisan news source - had a suspiciously high number of retweets disproportionate to its follower base.
 - The source’s Twitter account had just 8,543 followers, but its 144 links were collectively retweeted 151,378 times.
 - TCC sampled 557 Twitter user profiles that shared southafricatoday.net links and used a bot detection algorithm, that identified 364 users as “likely bots” with a 60% or greater probability.
 - Due to time and resource limitations, TCC was unable to investigate and verify how many accounts were “actual bots” however it is notable that the frequency distribution of “likely bots” is negatively skewed, indicating a strong likelihood that many “likely bots” are indeed bots.



Conclusions

- The approach utilized by the TCC mission highlighted the obvious benefits of using a domain classifier to facilitate the identification of possible political and election-related mis / disinformation.
- Most links from “dodgy” sources were considered to be generally credible or ignorable (i.e. apolitical) content however possibly “dodgy” links that were identified would have been far more difficult to detect absent a domain classifier.
- The use of a domain classifier allowed for a sharper focus on content from possible mis / disinformation sources and, in turn, the efficiency gain increased opportunity to devote to investigating spread.
- This enabled researchers to assess how mis / disinformation can be amplified and legitimized by political figures (presumably unintentionally) as well as how it can be amplified nefariously vis-à-vis computational propaganda (intentionally).



Conclusions

- The approach did not provide a contextualize lens through which to understand possible mis / disinformation within the broader political media ecosystem – or even within sub-ecosystems in which specific narrative frames may pervade
- The use of binary domain classification typology was of limited utility given that it did not distinguish “dodgy” sources in any meaningful way that may indicate motivation (i.e. satire, clickbait, hyper-partisan journalism, etc.).
- The development of a multi-class typology domain classifier and identification of sub-ecosystems would be beneficial and could be mutually reinforcing exercises.
 - Typological designations of domains as hyper-partisan journalism could be sub-categorized in accordance with editorial biases and, in turn, be used to identify ideological sub-ecosystems for monitoring purposes.
 - Conversely, if ideological sub-ecosystems are first identified, typological designations of domains could be denoted in accordance with the sub-ecosystems in which they are propagated.

