CSE 291/DSC 291
Information Manipulation: Trustworthiness of Information in Cyberspace

Molly Roberts & Stefan Savage, Winter 2023

International Sources of Misinformation
Election meddling, Bush and Prather (2022)

- Foreign interference in elections by international actors not new – Thomas Jefferson, George Washington worried about this in the founding of the U.S.
- Dov Levin, “Between 1946 and 2000, the United States and the Soviet Union/Russia have intervened in about one of every 9 competitive national-level executive elections.”
- Lobbing elected officials, providing foreign aid, contributing to campaigns, election monitoring, disinformation campaigns
- Some democracy-promoting, some democracy-undermining
- To what extent do they undermine or promote trust in elections?
Foreign propaganda campaigns also common

- US intervention in Chile’s 1964 presidential election
- US favored candidate Eduardo Frei Montalva, worried that the other candidate Allende was too close to Cuba
- Engaged in disinformation campaigns, propaganda campaigns, vote buying
- CIA propaganda often attributed to domestic source
Has the Internet made this easier?

• Political campaigns/news increasingly online
• Easier for foreign (and domestic) governments to manipulate information transmission
• Easier to pretend to be someone else?
• Online foreign influence effort: coordinated effort by a foreign government to influence politics of another country, masquerading as online users from that country
• 84 Online Foreign Influence Efforts, 62% from Russia

Online Political Influence Efforts Dataset*

Diego A. Martin¹ Jacob N. Shapiro¹ Julia G. Ilhardt³

Version 3.0
February 3, 2022

Abstract

This report version 3.0 of ESOC's data on the covert use of social media to influence politics by promoting propaganda, advocating controversial viewpoints, and spreading disinformation. Influence efforts are defined as: (i) coordinated campaigns by a state, or the ruling party in an autocracy, to impact one or more specific aspects of politics at home or in another state, (ii) through media channels, including social media, by (iii) producing content designed to appear indigenous to the target state. Our data draw on more than 1000 media reports and 500 research articles/reports to identify IEs, track their progress, and classify their features. The data cover 84 foreign influence efforts (FIEs) and 30 domestic influence efforts (DIEs)—in which governments targeted their own citizens—against 56 different countries from 2011 through 2021, five more countries than in the previous version of this report. The Influence Effort dataset measures covert information campaigns by state actors, facilitating research on contemporary statecraft.
Strategies of online foreign interference

- Promoting/advertising
- Defaming
- Dividing/Polarizing
- Distraction
- Signal-jamming
- Gaslighting
Strategies of foreign interference

Panel B: Share of attacks involving strategies

- downstream
- persuade
- Poalrize
- Shift Agenda
- Undermine institutions

- DIE
- FIE
What is the context for the IRA report?

- **Who are the authors?**
  - New Knowledge (now Yonder), company specializing in information integrity
  - Renee DiResta (now at the Stanford Internet Observatory)

- **Who commissioned the report?**
  - United States Senate Select Committee on Intelligence
  - Congressional oversight committee for U.S. Intelligence agencies (provide budget recommendations as well)
  - Previous oversight – intelligence in Iraq, Snowden revelations, etc

- **How did they get the data?**
  - Twitter, Facebook, and Alphabet
  - What were their complaints about the data?
What is the IRA? Who runs it?

- Established in 2013 in St. Petersburg, Russia
- Ties to Russian intelligence, financed by people close to Putin
- “Troll farm” focused on international influence
  - Interviews with workers there: managed Facebook/Instagram/Twitter accounts, commented on news stories, developed personas
  - Attracted followers with non-political content, bought followers, etc
  - Active in Ukraine, U.S., also Europe
- In 2018, federal grand jury indicted the IRA along with 13 Russian nationals for sowing information warfare against the United States
  - Foreigners barred from spending money to influence U.S. election
  - Identify theft, wire fraud, bank fraud, conspiracy charges
What platforms did they use from 2015-2018?
What were the themes/strategies of the campaign?

<table>
<thead>
<tr>
<th>Themes/Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black culture, community, Black Lives Matter</td>
</tr>
<tr>
<td>Blue Lives Matter, pro-police</td>
</tr>
<tr>
<td>Anti-refugee, pro-immigration reform</td>
</tr>
<tr>
<td>Texas culture, community, and pride</td>
</tr>
<tr>
<td>Southern culture (Confederate history)</td>
</tr>
<tr>
<td>Separatist movements and secession</td>
</tr>
<tr>
<td>Muslim culture, community, and pride</td>
</tr>
<tr>
<td>Christian culture, community, and pride</td>
</tr>
<tr>
<td>LGBT culture, community, and pride</td>
</tr>
<tr>
<td>Native American culture, community, and pride</td>
</tr>
<tr>
<td>Meme and “red pill” culture</td>
</tr>
<tr>
<td>Patriotism and Tea Party culture</td>
</tr>
<tr>
<td>Liberal and feminist culture</td>
</tr>
<tr>
<td>Veteran’s Issues</td>
</tr>
<tr>
<td>Gun rights, pro-2nd Amendment</td>
</tr>
<tr>
<td>Political Pro-Trump, anti-Clinton content</td>
</tr>
<tr>
<td>Pro-Bernie Sanders and Jill Stein content</td>
</tr>
<tr>
<td>Syria and ISIS, pro-Assad, anti-U.S. involvement</td>
</tr>
<tr>
<td>Trust in media</td>
</tr>
</tbody>
</table>
Lots of engagement
Inspired real-world offline activities

- Encouraged:
  - Self-defense classes
  - Prayer groups
  - Volunteers
  - Help phone lines
  - Protest organization
Extensive ad operations

- Targeted particular categories of Americans (e.g. coal-miners who support Trump)

- Selling of merchandise
  - Collect personal information about people
  - Identify people to target more advertising

<table>
<thead>
<tr>
<th>Ad Copy</th>
<th>Impressions</th>
<th>Clicks</th>
<th>CTR</th>
<th>Spend</th>
<th>Currency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community of people who support our brave Police Officers. Back The Badge</td>
<td>1,334,544</td>
<td>73,063</td>
<td>5.5%</td>
<td>110,588.00</td>
<td>RUB</td>
</tr>
<tr>
<td>Together we are a shining beacon of faithfulness! Join us and light your way to Jesus! Army Of Jesus</td>
<td>368,065</td>
<td>28,137</td>
<td>7.6%</td>
<td>89,081.10</td>
<td>RUB</td>
</tr>
<tr>
<td>Heritage, not hate. The South will rise again! South United</td>
<td>511,224</td>
<td>40,134</td>
<td>7.9%</td>
<td>78,357.30</td>
<td>RUB</td>
</tr>
<tr>
<td>They fought for our freedom, now its our time to fight for their rights!</td>
<td>112,737</td>
<td>15,163</td>
<td>13.4%</td>
<td>72,735.10</td>
<td>RUB</td>
</tr>
<tr>
<td>Join us if you know it’s just a car, NOT A SYMBOL OF HATE South United</td>
<td>300,815</td>
<td>16,587</td>
<td>5.5%</td>
<td>65,705.80</td>
<td>RUB</td>
</tr>
<tr>
<td>Brown Power is a platform designed to educate, entertain and connect Chicanos in the US.</td>
<td>968,768</td>
<td>56,405</td>
<td>5.8%</td>
<td>59,571.40</td>
<td>RUB</td>
</tr>
<tr>
<td>A page to unit all Muslim people living in the USA!</td>
<td>103,079</td>
<td>4,491</td>
<td>4.4%</td>
<td>59,453.80</td>
<td>RUB</td>
</tr>
<tr>
<td>Staying Woke, building-up our communities, uplifting our people. Join us! Woke Blacks</td>
<td>752,179</td>
<td>33,444</td>
<td>4.4%</td>
<td>58,193.60</td>
<td>RUB</td>
</tr>
<tr>
<td>We are proud to be black and stand for our community! Join to end the racism in the US.</td>
<td>371,295</td>
<td>16,442</td>
<td>4.4%</td>
<td>55,609.60</td>
<td>RUB</td>
</tr>
<tr>
<td>We are not against police, we against police brutality!</td>
<td>226,262</td>
<td>9,657</td>
<td>4.3%</td>
<td>54,729.50</td>
<td>RUB</td>
</tr>
</tbody>
</table>

Top ten IRA ads by spend, and associated click-through rates. Using 67.07 RUB = 1 USD (average 2016 exchange rate), that is $1649 USD spent on highest-spend ad.
Creation of a “media mirage”

A “media mirage” of interlinked Facebook Pages and Instagram accounts targeting Black Americans. Squares are IRA-owned Facebook Pages, parallelograms are IRA-owned Instagram accounts. Cloud shapes indicate non-IRA-attributed accounts including authentic Black media. A larger version with easily-readable page names is available online at newknowledge.com/IRAfigures
But what was their effect?


- 2016 YouGov survey, linked to social media behavior and data, 3 waves April-October 2016. Asked who they voted for after the election.
But what was their effect?

Fig. 2 | Average exposure to posts from domestic news media, politicians, and Russian influence campaign accounts during the final month of the election campaign. Panel a presents the mean exposure to posts from domestic news media (in red), political candidates (in orange), and Russian foreign influence accounts (in black) for the last month of the presidential campaign. Panel b presents an analogous time series per week for median exposure to domestic news media, political candidates, and Russian foreign influence accounts. Median exposure to Russian foreign influence accounts in the last month of the election is 0 for all days.
But what was their effect?

Fig. 5 | Regression results of the relationship between exposure to posts from Russian foreign influence accounts and voting behavior. These panels present OLS estimates (with 95% CIs) of the relationship between exposure to tweets from Russian foreign influence accounts and three vote choice outcomes. Each estimate is from a separate model (see Supplementary Methods E for full regression results). In the first row, the outcome “Vote choice (Clinton toward Trump)” is coded with three possible values: (+1) a shift from ranking Clinton preferable to Trump in the first wave to voting for Trump in the election; (-1) a shift from ranking Trump preferable to Clinton to voting for Clinton; or (0) no difference between first-wave ranked preferences and vote choice in the election. The second row outcome variable, “Vote choice (Clinton toward Trump, 3rd party, or not voting)” can also take 3 values: (+1) a shift from ranking Clinton above Trump in the first survey wave to voting for Trump, voting for a 3rd party candidate, or not voting; (-1) a shift from ranking Trump above Clinton to voting for Clinton, voting for a 3rd party candidate, or not voting; or (0) no difference between first-wave ranked preferences and for whom a respondent voted. The third row, “Rank (Trump > Clinton)” is coded as: (+1) a shift from ranking Clinton preferable to Trump in the first survey wave to preferring Trump to Clinton in the second wave; (-1) ranking Trump preferable to Clinton in the first wave to preferring Clinton to Trump in the second wave; or (0) no change between survey waves. The left panel represents a coding of the variable of interest as the log exposures + 1 number of posts in a respondent’s Twitter feed; the right panel, a binary variable indicating whether at least one such post appeared in a respondent’s feed. $n = 1,496$ survey respondents.
New York Times: Pro-U.S. influence Activities

- Stanford Internet Observatory
- Accounts taken down by Twitter and Facebook linked to U.S. government operation “Trans-Regional Web Initiative”
- Very few followers, likes, and retweets
- Target: Central Asia and Middle East
- Focused on Promoting/Defaming
Discussion

- Why invest so much money and effort into something that isn’t effective? Or is it effective?

- What about the possibility of backlash?

- How should we think about safeguarding elections from government-orchestrated foreign social media campaigns in the future? Or do we need to?