NATIONAL CAPITAL REGION
THREAT INTELLIGENCE CONSORTIUM
Intelligence Brief

How to Detect Disinformation Campaigns

Disinformation is false information intended to deliberately deceive or mislead individuals. Criminals, terrorist or extremist groups, state-sponsored actors, and conspiracy theorists use disinformation as a tool to spread propaganda, disseminate conspiracy theories, incite distrust and confusion, or compel violence. Social media bots and Internet shills are often used on social media platforms to quickly reach targets and accelerate the spread of disinformation.

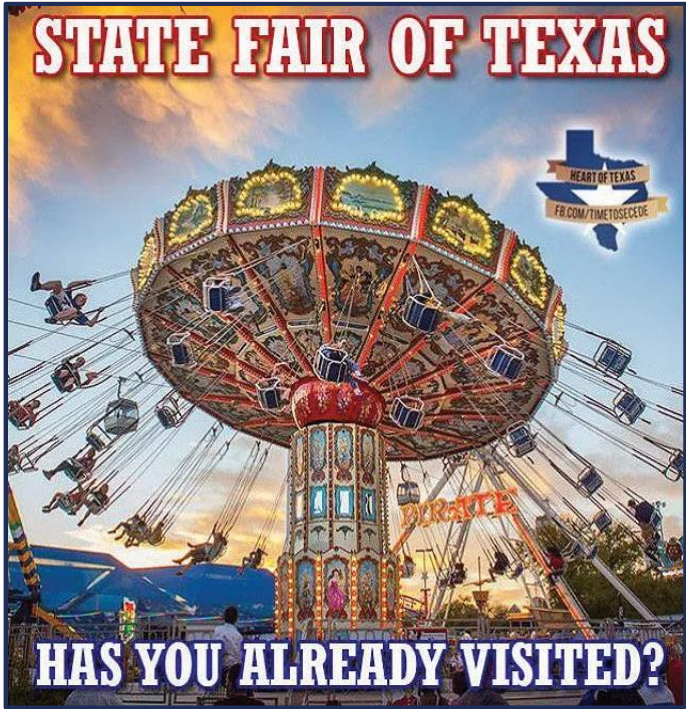|	BOT	SCHILL
ACCOUNT DEFINITION	A computer program that autonomously controls a social media account to perform certain tasks such as generating messages or interacting with other users.	A person who operates an online account to promote an idea, product, person, or group without divulging any information about direct associations or potential conflicts of interest.
HOW IT OPERATES	In disinformation campaigns, social media bots are often used by fraudulent social media accounts to automatically post or re-post messages and try to amplify specific topics. These accounts commonly display traits uncharacteristic of a legitimate account such as posting the same link multiple times, generating a lot of posts in a short amount of time, or only having a few followers.	Internet shills are often real people who engage with legitimate social media users and participate in online discussion forums to promote a specific viewpoint or opinion. Internet shills often receive payment in exchange for their online engagement promoting companies, governments, public figures, and more. Internet shills may use social media bots to expand their reach.
HOW TO DETECT IT	**Automated activity:** Social media bots often generate an excessive amount of online content over a short period. For example, according to the Atlantic Council's Digital Forensic Research Lab, the account @sunneversets1000, an amplifier of pro-Kremlin messaging, tweeted an average of 705 posts per day in a 288-day period in 2016-17. **Amplification:** Accounts may lack original content. For example, the last 195 tweets of @sunneversets100's account as of August 28, 2017 were retweets. **Anonymity:** Accounts often provide little to no personal information. The account's profile picture and background may not be unique to the individual or may be set to the social media's default photo. **Language:** A social media bot may retweet content in different languages via a single account to attract a wider audience. **Profile:** Social media bots may use a stolen image for their profile picture or a randomly generated alphanumeric username such as abc123.	**Content:** Accounts often use keywords and phrases that promote the agenda of the Internet shill's employer and content may be crafted to play upon common stereotypes to amplify political, cultural, or social division. **Grammar:** Foreign-based Internet shills often have grammatical errors or misspellings in posts. For example, a poster promoting the Heart of Texas rally used the incorrect verb tense. **Photos:** Internet shills may use stolen or stock photos as profile pictures rather than their own personal photos to stay anonymous. **Use of Icons:** Instead of using words, Internet shills will often use a plethora of emojis and graphics to appeal to social media users.

Multiple bot accounts with a stolen profile picture
(Source: Medium)

Shill account posts may have grammatical errors
(Source: Houstonia)

Social Media Companies' Actions Against Disinformation Campaigns

Since the 2016 presidential election, Twitter and Facebook have taken steps to close accounts involved in disinformation campaigns. Facebook removed nearly 3.39 billion accounts from October to March and Twitter shut down approximately 70 million accounts from May to July of 2018. The efforts of these social media platforms will likely cause disinformation actors to move activity to lesser-regulated platforms or employ enhanced technological capabilities to avoid detection.

Twitter: Twitter is focused on deterring the creation of fraudulent accounts at the time of creation through mandatory email or phone verification requirements and reCAPTCHAs, a tool used to prevent automated account creation by forcing users to identify portions of an image or type words displayed on the screen. They also prioritized the identification of suspicious activity such as high-volume tweeting with the same hashtag or mentioning the same user account without replies from the account being addressed.

Facebook: Facebook is restricting the ability to purchase false ads through strict policy enforcement and machine learning to detect fraud against inauthentic accounts. They are also using ranking improvements to reduce the prevalence of false news, creating easier reporting options for users, and forming partnerships with third-party fact-checking organizations to verify content. Lastly, Facebook is enhancing users' ability to make informed decisions about the content they read through a series of public service announcements with the News Literacy Project, launching the News Integrity Initiative, a global consortium focused on helping individuals making informed decisions about the news.