Targeted Trolling and Trend Manipulation:
How Organized Attacks on Amber Heard and Other Women Thrive on Twitter

Report

07-18-2022

PREPARED BY:
BOT SENTINEL INC.
# TABLE OF CONTENTS

1. Key Findings
2. Disclaimer
3. Overview
4. Introduction
5. Hashtag Spamming
6. Copypasta
7. Bragging and Encouraging
8. Abuse and Harassment
9. Trolling Prominent Academic
10. Tweet Data
11. Top 20 SPAM Accounts
12. Twitter Rules
13. Conclusion
14. Suggestions for Improvement
15. About Bot Sentinel Inc.
1. We identified 627 Twitter accounts focused predominantly on tweeting negatively about Amber Heard and her female supporters.
2. 3,288 accounts were tweeting #AmberHeardIsAnAbuser, #AmberHeardIsAnAbuser, #AmberHeardIsALiar, and #AmberHeardIsALiar, and 19% of those accounts were dedicated to spamming the hashtags.
3. Women who tweeted in support of Amber Heard were subjected to rampant abuse and targeted harassment on Twitter.
4. Someone used a photo of a woman's deceased child to create a fake account and troll the woman because she tweeted in support of Amber Heard.
5. Toxic trolls used hashtag spamming to trend anti-Amber Heard hashtags artificially.
6. Over 24% of the accounts tweeting anti-Amber Heard hashtags were created within the past seven months.
7. Tactics like copypasta (copying and pasting duplicative content) were used to amplify messaging artificially.
8. Toxic trolls continued to tweet anti-Amber Heard hashtags and attack women weeks after the Depp vs. Heard trial ended.
In 2020, Amber Heard's legal team contacted Bot Sentinel after we published our findings on the coordinated attack targeting Lisa Page. Amber Heard's legal team hired us to determine whether the social media activity against Ms. Heard was organic or if there was some other explanation.

We determined that a significant portion of the activity wasn't organic and put our findings in a report. In June 2022, we began reexamining the activity after the Depp v. Heard verdict. Neither Amber Heard nor anyone from her team hired Bot Sentinel to review the activity. No one hired Bot Sentinel to compile and publish this report.
During the Depp vs. Heard trial, we observed anti-Amber Heard hashtags regularly trending on Twitter. After the trial ended, the anti-Amber Heard hashtags continued trending. We researched the activity to determine whether the trending hashtags resulted from organic activity or platform manipulation.

We immediately observed dozens of newly created accounts spamming negative anti-Amber Heard hashtags. Many accounts were replying to tweets with hashtags unrelated to the tweet they were responding to. Some accounts encouraged others to get the hashtags trending, and the trolls were successful on multiple occasions.

Our research also uncovered rampant abuse and widespread targeted harassment. Abusive trolls who identified as Johnny Depp supporters had subjected women to verbal abuse and targeted harassment. The trolls didn't limit their abuse to women; they also targeted their family members. In at least one case, the trolls doxxed a woman's family and created a fake Twitter account using a photo of the woman's deceased child to troll her.

This report will illustrate one of the worst cases of platform manipulation and flagrant abuse from a group of Twitter accounts. It will show how trolls used tactics to manipulate conversations and trends across Twitter while targeting and abusing women to suppress any positive tweets supporting Amber Heard.

We gave a copy of this report to Twitter prior to publishing.
We analyzed 14,292 tweets with the hashtags #AmberHeardIsAnAbuser, #AmberHeardLsAnAbuser, #AmberHeardIsALiar, and #AmberHeardLsALiar. We identified 627 Twitter accounts focused predominantly on tweeting negatively about Amber Heard and her female supporters. Approximately 24.4% of the Twitter accounts tweeting negative Amber Heard hashtags were created within the past seven months. The number of new accounts tweeting about Amber Heard was significantly higher than accounts tweeting about other topics.

Some would argue the accounts were created because individuals were interested in the Depp vs. Heard trial. We agree a significant number of anti-Amber Heard accounts were created in response to the trial. Still, the new accounts focused mainly on discrediting and attacking Amber Heard and anyone who supported her.

We observed Twitter trolls using various platform manipulation tactics. For example, the trolls used hashtag spamming and copypasta (copying and pasting duplicative content) to amplify content, manipulate trends, and suppress pro-Amber Heard content. Accounts openly boasted about platform manipulation tactics and urged others to help get anti-Amber Heard hashtags trending.

Twitter trolls didn't limit their abuse to platform manipulation. We observed the trolls abusing and harassing Amber Heard and multiple other women and swarming the tweets of their victims. Offensive tweets, misogyny, doxxing, and death threats were rampant. The same group of toxic trolls also harassed and abused women who tweeted positively about Amber Heard.

Twitter trolls that attacked Amber Heard and her allies recently began attacking Evan Rachel Wood and Cassidy Hutchinson. The trolls compared the women to Amber Heard and used the same abusive tactics against the women.

The accounts in this report violated multiple rules and policies of the platform, including but not limited to violent threats, abuse/harassment, hateful conduct, private information, and platform manipulation and spam.

To reduce the number of pages in this report, we have uploaded a separate list of accounts violating Twitter's rules.

Hashtag spamming was essential to amplifying anti-Amber Heard content and trending anti-Amber Heard hashtags artificially. Hashtag manipulation tactics gave the false impression of overwhelming opposition to Amber Heard.
Trolls used tactics like copypasta (copying and pasting duplicative content) to amplify messaging and manipulate Twitter’s trends. The goal, in this case, was to paint Amber Heard as a liar while painting Johnny Depp as an intelligent man who was telling the truth.
Trolls that participated in spamming anti-Amber Heard hashtags bragged about their ability to get hashtags to trend. They also encourage others to spam anti-Amber Heard hashtags.

Wonderful to see these trending at the same time! 😂

#JohnnyDeppIsASurvivor
#AmberHeardIsAnAbuser

LET'S KEEP IT GOING!
RETWEET & TWEET! 👉👇

When these hashtags are trending together, you know the 1st one had been trolled again 😂😂😂

#Ibelieveambertheturd #Ibelieveamberhadtheturd
#Ibelieveamberistheturd 😣😭😭

Hey fam, I suggest a new hashtag #IbelieveAmberHadTheTurd what do you think? How funny would that be to make it trend next to #AmberIsA liar #AmberHeardIsAnAbuser ...?

I trended #AmberHeardIsAnAbuser without telling to trend it. Just commented that everywhere until people started to use that also.

The several times i asked to trend it was when they spammed TUG and Popcorned planet

Why she need to announce?
Amber Heard and women who tweeted in support of Amber were attacked relentlessly. Twitter trolls would swarm the tweets of women who tweeted positively about Amber Heard and often used vulgar and threatening language.

---

Harion Demon
@DemonHarlon

Replying to @LindseyBoylan and @a_mberheard

you’re right let’s go fuck amber heard together anyway she likes these things

Darrel Johnson
@DarrelPatriot

Replying to @LindseyBoylan

Imao the bitch heard got what she deserved, cant wait to hear her dying in overdose after her life falls apart. #AmberTurd

Sydney Blaque
@SydneyBlaque

@realamberheard wow you are a dumb f*cking cunt. Get the hell over yourself. Leave the man alone and stop being the trash we all know you are. 👍🏻👏🏻👏🏻

Cloudier
@CloudierClouds

@realamberheard cunt

Translate Tweet

Kim
@Kim49210384 - Jun 10

@Eve_Barlow log off shitbag

Kim
@Kim49210384 - Jun 10

@Eve_Barlow hey dumb dumb you were in the court room making a mess you will be judged ya slimy cunt enjoy life behind bars.

Kim
@Kim49210384 - Jun 10

@Eve_Barlow you’re so clumsy and DUMBSY so fucking delusional you have an axe to grind. You’re a sell out shitbag who I’d love to slap around

Kim
@Kim49210384 - Jun 10

@Eve_Barlow YOU ARE VILE DISGUSTING DISGRACE TO THE JEWS! We don’t need you as our voice ya sad slimy junkie now fuck off

Kim
@Kim49210384 - Jun 10

@Eve_Barlow Desperate waste of flesh, feeble minded low level wanna be something or somebody so bad you jealous. Shit bag, I dare you to keep stalking Johnny ya fucking smelly silverback gorilla

Rando Who Drinks Lote Of Tea
@Marge56381403

Well of course @realamberheard lied. That’s all she’s done probably her whole life. Her parents should be ashamed of the scum where they produced and created. Good job just what the world needed. Mote trash
Trolling didn't stop with vulgar and abusive tweets. In one case, someone created a fake account using the image of a deceased child of a prominent academic. The troll used the fake account to taunt and torment the person because she tweeted in support of Amber Heard. Trolls escalated their harassment and doxxed the academic's family members. This criminal activity led to the target fearing for the safety of her family and herself.

We didn't include the name of the academic to protect her identity and her family's safety. However, we felt it was important to include her story in this report to document the severity of the harassment.
We analyzed 14,292 tweets and identified 627 anti-Amber Heard Twitter accounts tweeting exclusively about Amber Heard and her female supporters. Approximately 24.4% of the Twitter accounts tweeting negative Amber Heard hashtags were created within the past seven months; the average for other topics is 8.6%. Accounts spammed #AmberHeardLsAnAbuser and #AmberHeardLsALiar with the letter "I" purposely replaced with the letter "L" to deceive Twitter's algorithms. The intentional misspelling demonstrates a calculated effort to manipulate hashtag trends.
The top 20 Twitter accounts spamming the hashtags #AmberHeardIsAnAbuser, #AmberHeardLsAnAbuser, #AmberHeardIsALiar, or #AmberHeardLsALiar.

<table>
<thead>
<tr>
<th>Handle</th>
<th>ID</th>
<th>Tweets</th>
<th>Joined</th>
<th>Account Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>@AttViolet</td>
<td>1539822661454397442</td>
<td>520</td>
<td>06-23-2022</td>
<td>21 day(s)</td>
</tr>
<tr>
<td>@TerryParker</td>
<td>361196381</td>
<td>288</td>
<td>08-24-2011</td>
<td>132 month(s)</td>
</tr>
<tr>
<td>@jaybee656677</td>
<td>1538654582712242178</td>
<td>278</td>
<td>06-19-2022</td>
<td>24 day(s)</td>
</tr>
<tr>
<td>@NormaisSantos</td>
<td>995214837300879360</td>
<td>255</td>
<td>05-12-2018</td>
<td>50 month(s)</td>
</tr>
<tr>
<td>@fishyfishfry</td>
<td>1523163116852121600</td>
<td>211</td>
<td>05-08-2022</td>
<td>2 month(s)</td>
</tr>
<tr>
<td>@monchxrtitetezza</td>
<td>1423159980237086241</td>
<td>170</td>
<td>08-05-2021</td>
<td>11 month(s)</td>
</tr>
<tr>
<td>@amtsstuff</td>
<td>1529641615119511552</td>
<td>158</td>
<td>05-26-2022</td>
<td>1 month(s)</td>
</tr>
<tr>
<td>@Havasuboo</td>
<td>1536562984470546866</td>
<td>146</td>
<td>06-14-2022</td>
<td>1 month(s)</td>
</tr>
<tr>
<td>@MarciaModenesse8</td>
<td>402001712</td>
<td>140</td>
<td>10-31-2011</td>
<td>130 month(s)</td>
</tr>
<tr>
<td>@emmastrange2702</td>
<td>740875465102958080</td>
<td>133</td>
<td>06-09-2016</td>
<td>74 month(s)</td>
</tr>
<tr>
<td>@ChelsyHebling</td>
<td>1513572922635571206</td>
<td>114</td>
<td>04-11-2022</td>
<td>3 month(s)</td>
</tr>
<tr>
<td>@dickwarlock13</td>
<td>1529517120095463168</td>
<td>109</td>
<td>05-25-2022</td>
<td>1 month(s)</td>
</tr>
<tr>
<td>@kermccoy</td>
<td>3031525901</td>
<td>98</td>
<td>02-11-2015</td>
<td>90 month(s)</td>
</tr>
<tr>
<td>@TheTypewriterQ</td>
<td>1442471158780100623</td>
<td>95</td>
<td>06-27-2021</td>
<td>9 month(s)</td>
</tr>
<tr>
<td>@aMegaDeppFan</td>
<td>1539058415262826499</td>
<td>91</td>
<td>06-21-2022</td>
<td>23 day(s)</td>
</tr>
<tr>
<td>@JUSTICE4DEPP01</td>
<td>1289362738393702402</td>
<td>82</td>
<td>08-01-2020</td>
<td>23 month(s)</td>
</tr>
<tr>
<td>@pacificcc</td>
<td>1418624095525146624</td>
<td>79</td>
<td>07-23-2021</td>
<td>11 month(s)</td>
</tr>
<tr>
<td>@JDs_Best_Life</td>
<td>1523337436264288258</td>
<td>70</td>
<td>05-08-2022</td>
<td>2 month(s)</td>
</tr>
<tr>
<td>@Rando_99999999</td>
<td>1519058134425210880</td>
<td>68</td>
<td>04-26-2022</td>
<td>2 month(s)</td>
</tr>
<tr>
<td>@Ranigill108</td>
<td>922570812929990556</td>
<td>67</td>
<td>10-23-2017</td>
<td>57 month(s)</td>
</tr>
</tbody>
</table>
The accounts in this report violated multiple rules and policies of the platform, including but not limited to violent threats, abuse/harassment, hateful conduct, private information, and platform manipulation and spam.

**Safety**

*Violence.* You may not threaten violence against an individual or a group of people. We also prohibit the glorification of violence. Learn more about our **violent threat** and **glorification of violence** policies.

*Terrorism/violent extremism.* You may not threaten or promote terrorism or violent extremism. Learn more.

*Child sexual exploitation.* We have zero tolerance for child sexual exploitation on Twitter. Learn more.

*Abuse/harassment.* You may not engage in the targeted harassment of someone, or incite others to do so. This includes wishing or hoping that someone experiences physical harm. Learn more.

*Hateful conduct.* You may not promote violence against, threaten, or harass other people on the basis of race, ethnicity, national origin, caste, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease. Learn more.

*Perpetrators of violent attacks.* We will remove any accounts maintained by individual perpetrators of terrorist, violent extremist, or mass violent attacks, and may also remove tweets disseminating manifestos or other content produced by perpetrators. Learn more.

*Suicide or self-harm.* You may not promote or encourage suicide or self-harm. Learn more.

*Sensitive media.* Including graphic violence and adult content. You may not post media that is excessively gory or show violent or adult content within live video or in profile or header images. Media depicting sexual violence and/or assault is also not permitted. Learn more.

*Illegal or certain regulated goods or services.* You may not use our service for any unlawful purpose or in furtherance of illegal activities. This includes selling, buying, or facilitating transactions in illegal goods or services, as well as certain types of regulated goods or services. Learn more.

**What is Copypasta?**

"Copypasta" (a reference to copy-and-paste functionality to duplicate content) is an Internet slang term that refers to an attempt by multiple individuals to duplicate content from an original source and share it widely across social platforms or forums. On Twitter, copypasta or duplicative content can be a block of text, image or a combination of content that has been copied and pasted, or duplicated by any means across the platform.

While copypasta or duplicate content is a tactic for propagating a message, and is used for a wide range of purposes, it can be repetitive, spammy, and disruptive to people’s experience on Twitter. Duplicative content can also be used to artificially amplify content, suppress information or manipulate Twitter’s Trends, Top Search results and conversations across the platform.

**Hashtags**

- using a trending or popular hashtag with an intent to subvert or manipulate a conversation or to drive traffic or attention to accounts, websites, products, services, or initiatives; and
- Tweeting with excessive, unrelated hashtags in a single Tweet or across multiple Tweets.

**Privacy**

*Private information.* You may not publish or post other people’s private information (such as home phone number and address) without their express authorization and permission. We also prohibit threatening to expose private information or incentivizing others to do so. Learn more.

*Non-consensual nudity.* You may not post or share intimate photos or videos of someone that were produced or distributed without their consent. Learn more.

**Authenticity**

*Platform manipulation and spam.* You may not use Twitter’s services in a manner intended to artificially amplify or suppress information or engage in behavior that manipulates or disrupts people’s experience on Twitter. Learn more.

*Civic integrity.* You may not use Twitter’s services for the purpose of manipulating or interfering in elections or other civic processes. This includes posting or sharing content that may suppress participation or mislead people about when, where, or how to participate in a civic process. Learn more.

*Misleading and Deceptive identities.* You may not impersonate individuals, groups, or organizations to mislead, confuse, or deceive others, nor use a fake identity in a manner that disrupts the experience of others on Twitter. Learn more.

*Synthetic and manipulated media.* You may not deceptive share synthetic or manipulated media that are likely to cause harm. In addition, we may label Tweets containing synthetic and manipulated media to help people understand their authenticity and provide additional context. Learn more.

Copyright and trademark.* You may not violate others’ intellectual property rights, including copyright and trademark. Learn more about our **trademark policy** and **copyright policy.**

**Multiple accounts and coordination**

You can't mass-register Twitter accounts or use automation to create Twitter accounts.

You can't artificially amplify or disrupt conversations through the use of multiple accounts or by coordinating with others to violate the Twitter Rules. This includes:

- overlapping accounts – operating multiple accounts with overlapping use cases, such as identical or similar personas or substantially similar content;
- mutually interacting accounts – operating multiple accounts that interact with one another in order to inflate or manipulate the prominence of specific Tweets or accounts; and
- coordination – creating multiple accounts to post duplicative content or create fake engagement, including:
  - posting identical or substantially similar Tweets or hashtags from multiple accounts you operate,
  - engaging (Retweets, Likes, mentions, Twitter Poll votes) repeatedly with the same Tweets or accounts from multiple accounts that you operate,
  - coordinating with or compensating others to engage in artificial engagement or amplification, even if the people involved use only one account; and
  - coordinating with others to engage in or promote violations of the Twitter Rules, including violations of our **abusive behavior policy.**
A group of Twitter trolls repeatedly used platform manipulation tactics to trend anti-Amber Heard hashtags. The trolls encouraged hashtag spamming and openly boasted when they were successful. Trolls frequently used copypasta techniques to amplify pro-Johnny Depp content while spamming negative Amber Heard propaganda.

Twitter trolls regularly swarmed the tweets of women who supported Amber Heard, and abuse and targeted harassment were commonplace. Tweeting in support of Amber Heard without being swarmed by trolls was difficult at best.

Vocal supporters of Amber Heard received the most abuse, which included death threats and doxxing. Trolls targeted their victims' families, and the children of their victims weren't off-limits to the trolls. What we observed was one of the worst cases of cyberbullying and cyberstalking by a group of Twitter accounts that we've ever seen.

It's our opinion Twitter didn't do enough to mitigate the platform manipulation and did very little to stop the abuse and targeted harassment. Twitter essentially left the women to fend for themselves with little to no support from the platform. The abuse and targeted harassment continued well after the Depp vs. Heard trial ended, and unless Twitter takes action, the abuse and targeted harassment will still continue.
Bad actors have become adept at platform manipulation and targeted harassment, and Twitter has allowed itself to become weaponized against its users. Twitter must be more aggressive at suspending accounts that abuse hashtags and copypasta, but more importantly, adopt new policies to address accounts that target specific individuals or organizations.

Twitter must improve the way it moderates content. Far too many tweets that break Twitter's rules are incorrectly found not to have violated Twitter's rules, and this is frustrating for users. We believe the deficient current reporting system will have the unintended effect of users reporting abusive content less often, which in turn would allow abusive content to flourish.
Bot Sentinel Inc. is a research firm comprising data scientists, machine learning engineers, and researchers specializing in identifying and understanding disinformation and targeted attacks on social media platforms. Our mission is to detect, analyze, and effectively expose threats using machine learning/artificial intelligence, HUMINT (Human Intelligence), and other various methods and technologies.

The name Bot Sentinel means autonomous guard. In other words, we are using a good "bot" to help identify and guard against inauthentic accounts and toxic trolls. Bot Sentinel doesn't just classify automated accounts aka bads bots, we developed Bot Sentinel to classify several types of accounts including accounts operated by human beings.

Our free platform and tools help expose ongoing threats while helping to protect our users from targeted attacks. The platform we provide also allows journalists and other experts to conduct their own research using the free tools and data we offer.

Over 30 news outlets and publications have cited our research and data, and approximately 100,000 active users use our free platform and tools each month.

Press inquiries: press@botsentinel.com