

# **HLF Workshop on Misinformation and Social Bots**

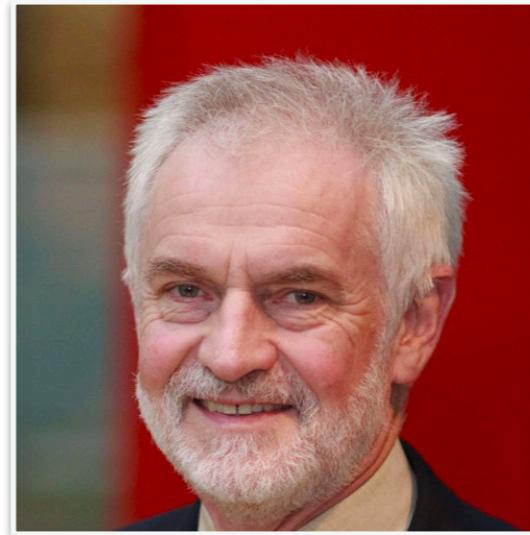
2019-09-23

# Introductions



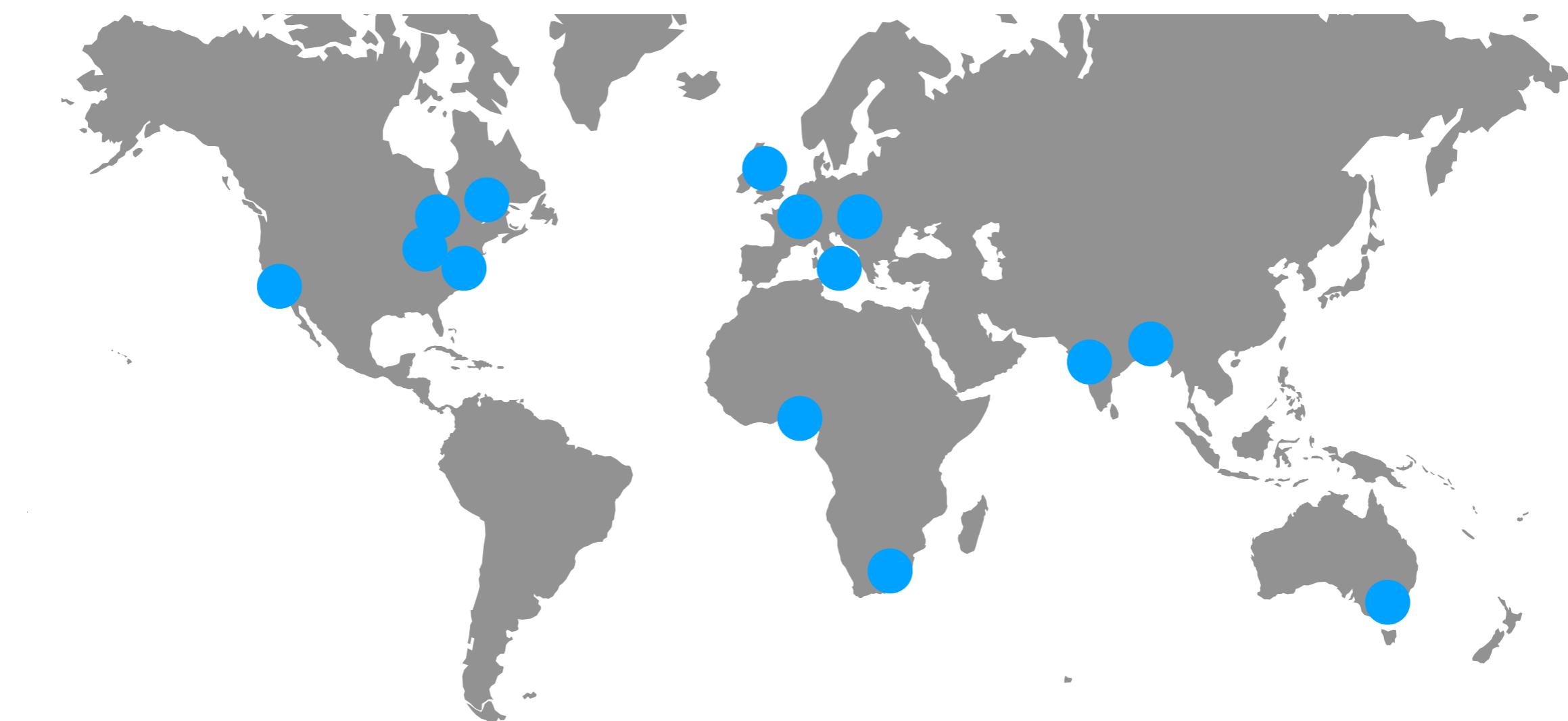
**Onur Varol**

Postdoctoral  
researcher at  
Northeastern  
University Network  
Science Institute



**Andreas Reuter**

Managing director of  
the Heidelberg Institute  
for Theoretical Studies  
(HITS) until 2016.



# Schedule

Time	Activity	Notes
15:30 – 15:35	Introductions and meeting kick-off	
15:35 – 16:00	Presentation on “Misinformation and social bots on social networks”	Presentation by Onur Varol
16:00 – 16:05	Break and study groups setup	Please find your study group from the attendance forms.
16:05 – 16:40	<u>Discussions between study groups</u> Emerging technologies and detection techniques Integrity of elections over the world Ethics and platform responsibilities Crowd-sourced efforts for data collection Media literacy and pro-active awareness campaigns Brave ideas for the new world	
16:40 – 17:00	Discuss topics identified by working groups Key takeaways and next steps	

Public speeches



Abraham Lincoln  
1860

First American  
president on radio



Warren G. Harding  
1922

First political advertisement  
aired on TV



Dwight D. Eisenhower  
1952

1650



1837



1876



1902



1927



1947  
Harry S. Truman



First presidential  
speech on TV

1960

Kennedy & Nixon



First presidential  
debate on TV



## I am Barack Obama, President of the United States -- AMA

POLITICS

submitted 4 years ago \* by PresidentObama

Obama



Hi, I'm Barack Obama, President of the United States. Ask me anything. I'll be taking your questions for half an hour starting at about 4:30 ET.

Proof it's me:

<https://twitter.com/BarackObama/status/240903767350968320>



President Obama

@POTUS



Following

Hello, Twitter! It's Barack. Really! Six years in, they're finally giving me my own account.

RETWEETS

274,794

LIKES

414,318



11:38 AM - 18 May 2015

2012



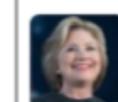
1988



2004



2015



Hillary Clinton

@HillaryClinton



Following

2016



Donald J. Trump

@realDonaldTrump

Obama just endorsed Crooked Hillary. He wants four more years of Obama—but nobody else does!

RETWEETS

507,017

LIKES

671,876

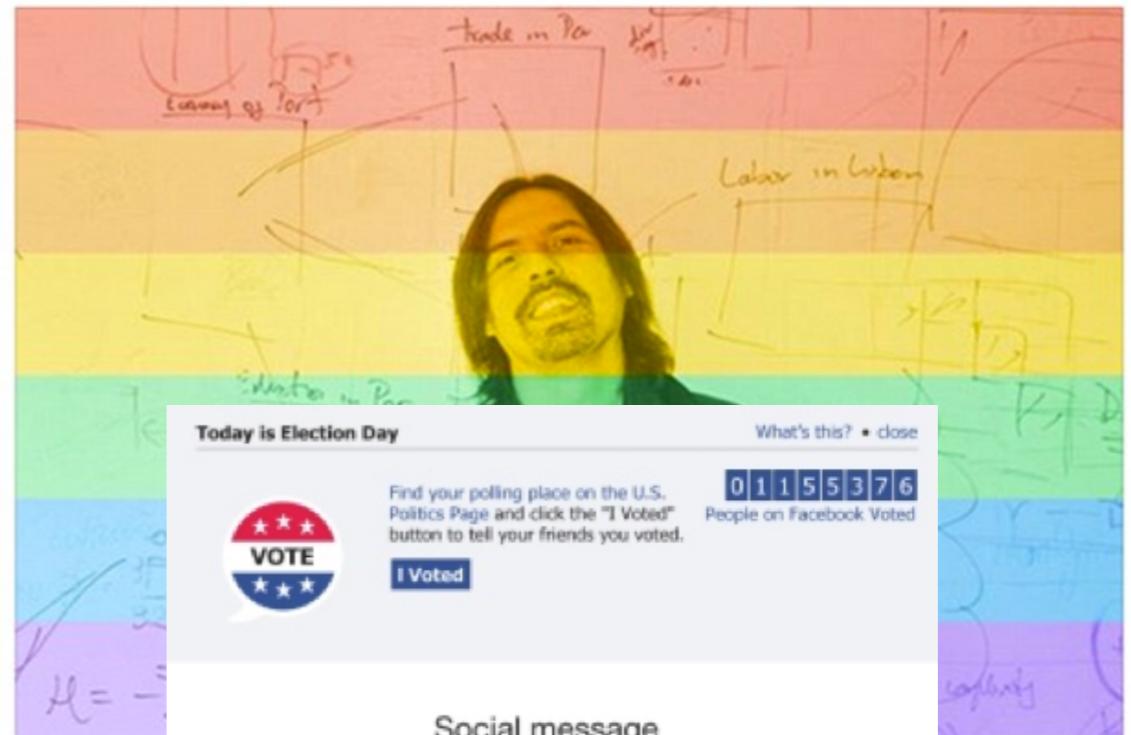


2:27 PM - 9 Jun 2016

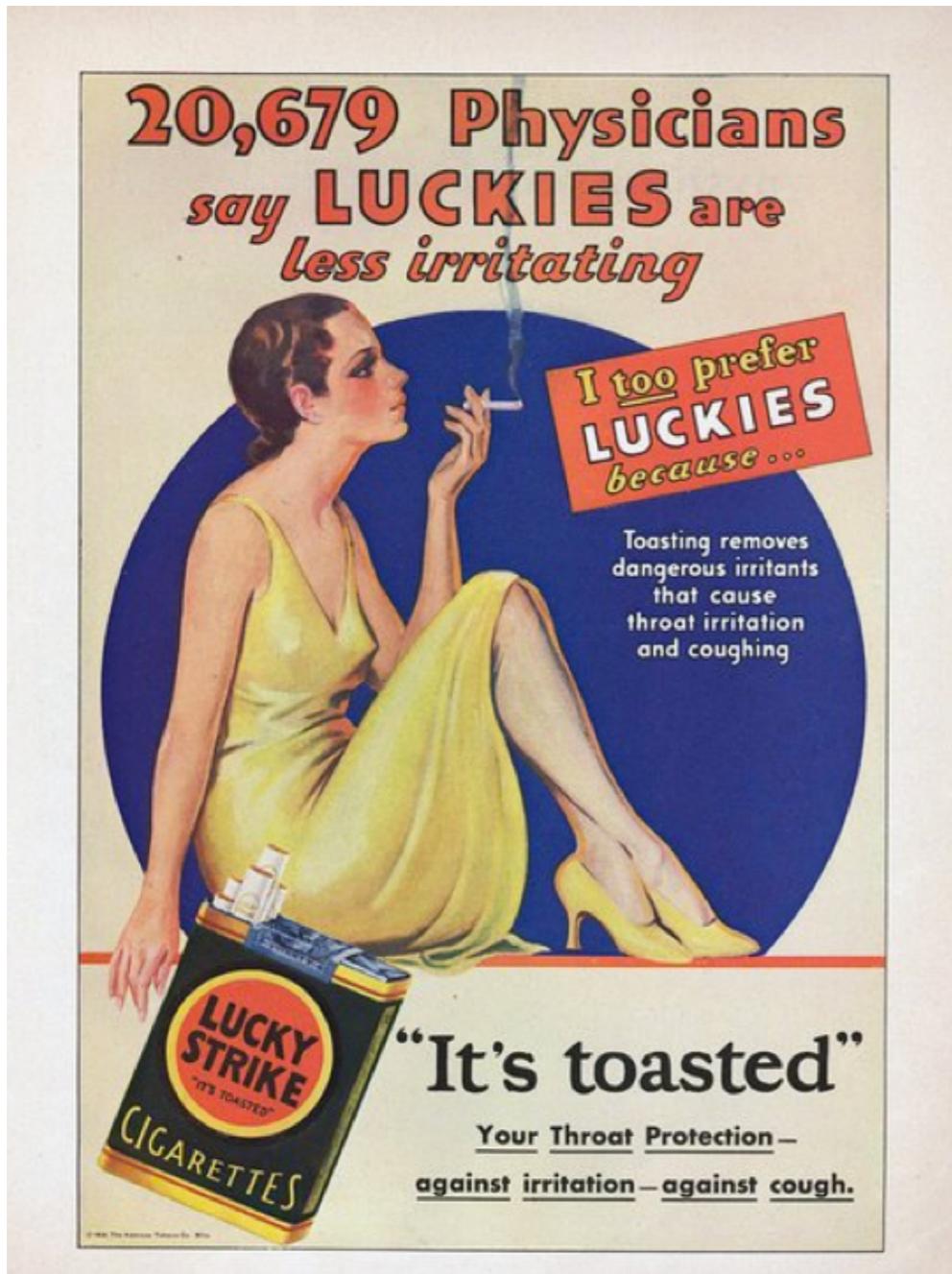


Cesar A. Hidalgo changed his profile picture.  
June 26 at 8:02pm · Edited ·

This is probably a Facebook experiment! The question is how long will it take for people to change their profile pictures back to normal. 😊



**Crystallizing Public Opinion - Edward L. Beynays**



**Crystallizing Public Opinion** - Edward L. Beynays

Cesar A. Hidalgo changed his profile picture.  
June 26 at 8:02pm · Edited ·

This is probably a Facebook experiment! The question is how long will it take for people to change their profile pictures back to normal. 😊



## Do Diffusion Protocols Govern Cascade Growth?

Justin Cheng<sup>1</sup>, Jon Kleinberg<sup>2</sup>, Jure Leskovec<sup>3</sup>, David Liben-Nowell<sup>4</sup>,  
Bogdan State<sup>1</sup>, Karthik Subbian<sup>1</sup>, and Lada Adamic<sup>1</sup>  
jcheng@fb.com, kleinber@cs.cornell.edu, jure@cs.stanford.edu, dln@carleton.edu,  
bogdanstate@fb.com, ksubbian@fb.com, ladamic@fb.com  
<sup>1</sup>Facebook, <sup>2</sup>Cornell University, <sup>3</sup>Stanford University, <sup>4</sup>Carleton College

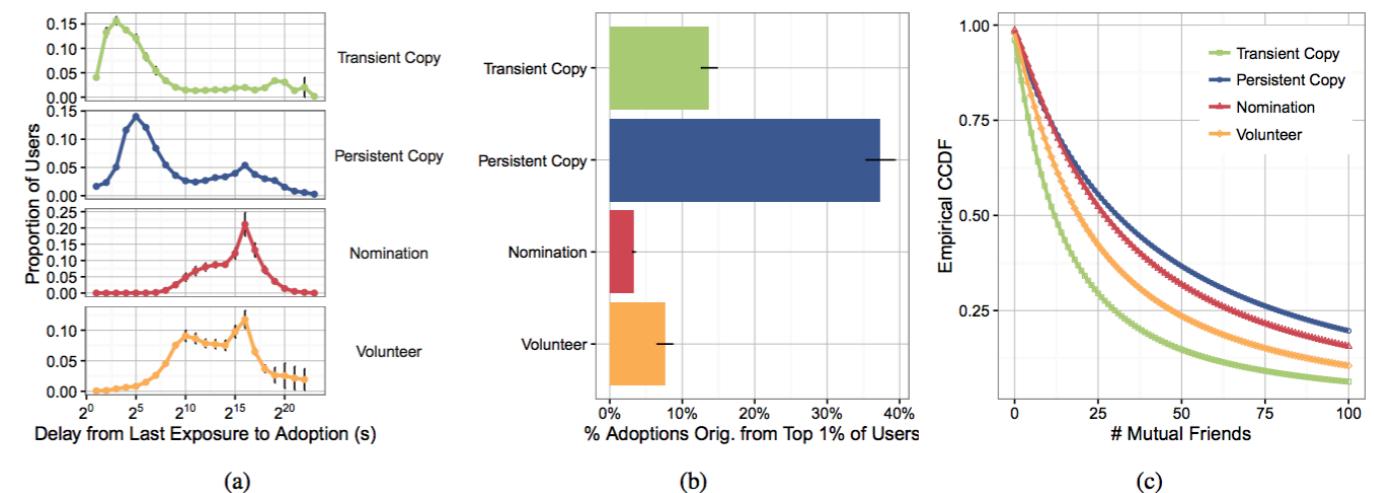


Figure 4: (a) Copy cascades are transmitted from individual to individual more quickly than volunteer or nomination cascades. (b) Copy cascades are also more reliant on hubs (i.e., nodes with high degree) for distribution, while volunteer and nomination cascades spread more evenly. (c) However, while transient copy cascades spread through weaker ties, persistent copy cascades spread through stronger ties. (Error bars denote standard errors of the mean across cascades.)

# Terms and definitions

**Misinformation:** False or inaccurate information such as false rumors.

**Disinformation:** Intentions are important to distinguish from misinformation. Disinformation has deliberate effort to disseminate information known to be incorrect.

**Fake-news:** “Fabricated information that mimics news media content in form but not in organizational process or intent.” (Lazer, 2018)

**Social bots:** “A computer algorithm that automatically produces content and interacts with humans on social media, trying to emulate and possibly alter their behavior.” (Ferrara, 2016)

**Trolls & sock-puppets:** Online identities used for online deception and orchestrated activities with malicious intentions.

RADIO SHOW NEWS VIDEOS STORE TOP STORIES BREAKING NEWS CONTACT

## "SPIRIT COOKING": CLINTON CAMPAIGN CHAIRMAN PRACTICES BIZARRE OCCULT RITUAL

Menstrual blood, semen and breast milk: Most bizarre WikiLeaks revelation yet

Paul Joseph Watson | Infowars.com • NOVEMBER 4, 2016 • 3857 Comments



TODAY ON THE ALEX JONES SHOW

Wednesday, Feb. 21st - 2nd Amendment Under Fire - The

The Washington Times

HOME NEWS - OPINION - SPORTS - MARKET - SUBSCRIBE -



Follow

LEAKED EMAIL appears to link Clinton Campaign Chairman to bizarre occult ritual [bit.ly/2f2kyXP](http://bit.ly/2f2kyXP)

11:50 AM - 4 Nov 2016

7,230 Retweets 7,404 Likes



1.8K 7.2K 7.4K

Barbara Ann Goehl @GoehlAnn • 4 Nov 2016

Replying to @seanhannity

SHE IS PURE EVIL

2 11 37

1 more reply

HAPPENING WINTER OLYMPICS 2018: LATEST NEWS FROM THE GAMES IN NOW PYEONGCHANG

## WikiLeaks: Podesta invited to 'Spirit' dinner; host's known 'recipes' demand breast milk, sperm

'Dear Tony, I am so looking forward to the Spirit Cooking dinner ... All my love, Marina'



LISTEN TO SEAN ON iHEARTRADIO FIND A STATION ABOUT

THE SeanHannity SHOW SHOWS > ARTICLES > FORUMS >

some Clinton insiders don't hold Chelsea Clinton in very high regard.

WITH A SHARP KNIFE CUT DEEPLY INTO THE MIDDLE FINGER OF YOUR LEFT HAND

LEAKED EMAIL APPEARS TO LINK CLINTON CAMPAIGN CHAIRMAN TO BIZARRE OCCULT...

November 4, 2016

This is by far the most bizarre email to date.

Benkler, Yochai, Robert Faris, and Hal Roberts. Network propaganda: Manipulation, disinformation, and radicalization in American politics. Oxford University Press, 2018.



“LEAKED EMAIL appears to link **Clinton** Campaign Chairman to **bizarre occult ritual**,” tweeted Sean Hannity at 11:50 a.m. on November 4, 2016, linking to a radio and online segment on the topic.



“LEAKED EMAIL appears to link **Clinton** Campaign Chairman to **bizarre occult ritual**,” tweeted Sean Hannity at 11:50 a.m. on November 4, 2016, linking to a radio and online segment on the topic.



Washington Times published a story entitled “Wikileaks: Podesta invited to ‘**Spirit**’ dinner: host’s known ‘**recipes**’ demand breast milk, sperm,” crediting the Drudge Report as having broken the story.



"LEAKED EMAIL appears to link **Clinton** Campaign Chairman to **bizarre occult ritual**," tweeted Sean Hannity at 11:50 a.m. on November 4, 2016, linking to a radio and online segment on the topic.



Washington Times published a story entitled "Wikileaks: Podesta invited to '**Spirit dinner**: host's known '**recipes**' demand breast milk, sperm," crediting the Drudge Report as having broken the story.



The "spirit cooking" email was from Tony Podesta to his brother John, forwarding an invitation from performance artist Marina Abramovic "to the Spirit Cooking dinner at my place." The occult sex ritual framing and the **striking image of Abramovic** borrowed directly, with credit, from the story Cernovich posted the night before.



“LEAKED EMAIL appears to link **Clinton** Campaign Chairman to **bizarre occult ritual**,” tweeted Sean Hannity at 11:50 a.m. on November 4, 2016, linking to a radio and online segment on the topic.



Washington Times published a story entitled “Wikileaks: Podesta invited to ‘**Spirit**’ dinner: host’s known ‘**recipes**’ demand breast milk, sperm,” crediting the Drudge Report as having broken the story.



The “spirit cooking” email was from Tony Podesta to his brother John, forwarding an invitation from performance artist Marina Abramovic “to the Spirit Cooking dinner at my place.” The occult sex ritual framing and the **striking image of Abramovic** borrowed directly, with credit, from the story Cernovich posted the night before.



The **Infowars** story amped up the volume by explaining that spirit cooking was sacrament in the religion of Thelema and involves an occult performance. To deepen the **conspiracy**, the Infowars story picked up the **pizza storyline** that would become the foundation of Pizzagate soon after the election.



"LEAKED EMAIL appears to link **Clinton** Campaign Chairman to **bizarre occult ritual**," tweeted Sean Hannity at 11:50 a.m. on November 4, 2016, linking to a radio and online segment on the topic.



Washington Times published a story entitled "Wikileaks: Podesta invited to '**Spirit**' dinner: host's known '**recipes**' demand breast milk, sperm," crediting the Drudge Report as having broken the story.



The "spirit cooking" email was from Tony Podesta to his brother John, forwarding an invitation from performance artist Marina Abramovic "to the Spirit Cooking dinner at my place." The occult sex ritual framing and the **striking image of Abramovic** borrowed directly, with credit, from the story Cernovich posted the night before.



The **Infowars** story amped up the volume by explaining that spirit cooking was sacrament in the religion of Thelema and involves an occult performance. To deepen the **conspiracy**, the Infowars story picked up the **pizza storyline** that would become the foundation of Pizzagate soon after the election.



In March 2016, Hillary Clinton's campaign manager, John Podesta, was hacked in a spear-phishing attack. WikiLeaks published his emails in November 2016. Proponents of the Pizzagate conspiracy theory falsely claimed the emails contained coded messages that **connected several U.S. restaurants and high-ranking officials of the Democratic Party** with an alleged **human trafficking and child sex ring**.

# Social bots can ...

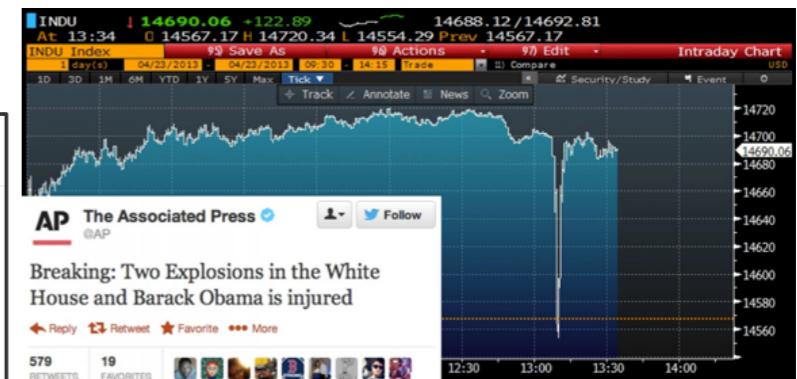
- Manipulate stock market
- Political smearing, astroturf, fake followers
- Marketing, spamming and social pollution
- Affect mood, behavior, etc.



**GIZMODO**

## The Fembots of Ashley Madison

Annalee Newitz  
Filed to: ASHLEY MADISON 8/27/15 5:05pm

A black and white photograph of a woman with her finger to her lips, suggesting silence or a secret.

A screenshot of the Twitter profile for TayTweets, which features a colorful, pixelated background and a large 'Tay.ai' logo. The profile has 96.2K tweets and 33.2K followers. A recent tweet from March 23, 2016, reads: "hellooooooooo w@rld!!!".

# Human-bot spectrum



Along this spectrum bots may vary by intentions and strategies which affects

- Profile humanness
- Automation strategies
- Types of content generated
- Network structures

**Earthquake Robot** (@earthquakeBot)  
Tweets 13.2K Following 6 Followers 63.9K  
**Tweets** **Tweets & replies**  
Earthquake Robot @earthquakeBot · 8h A 5.3 magnitude earthquake occurred in South of the Kermadec Islands. Details: eqbot.com/RcZ Map:  
34°52'07.0"S 179°22'20.6"W google.com  
May date 6/2018  
Up a tower.  
Joined March 2010  
Tweet to Earthquake Robot  
15 Followers you know

**Big Ben** (@big\_ben\_clock)  
Tweets 75.4K Following 465K Followers 465K  
**Tweets** **Tweets & replies**  
Big Ben @big\_ben\_clock · 26m BONG BONG BONG  
The first, established November 2009 & entirely unofficial. Copyright details here: tinyurl.com/y1qbev3, but still apparently imitated everywhere.  
Up a tower.  
Joined October 2009

**Buy 2,500 Twitter Followers for \$29**  
**VISIT** **WWW.MYSOCIALFANS.ORG**  
100% MONEY BACK GUARANTEE 100% safe  
**Calypso Cheverell** (@CalypsoChevere1) Follows you  
Pinned Tweet  
Calypso Cheverell @CalypsoChevere1 · 11h  
Quickly and efficiently get Thousands Twitter Followers per day. Buy 2500 Followers only \$29, see here goo.gl/WHU3Ck  
Joined July 2018

# Human-bot spectrum



Along this spectrum bots may vary by intentions and strategies which affects

- Profile humanness
- Automation strategies
- Types of content generated
- Network structures

**Earthquake Robot**  
@earthquakeBot  
I am a robot that tweets about any earthquakes 5.0 or greater as they happen. Built by @billsnitzer. Data is from the USGS. Get prepared: [amzn.to/PayDo](http://amzn.to/PayDo)  
Joined March 2010  
Tweet to Earthquake Robot  
15 Followers you know

Tweets 13.2K Following 6 Followers 63.9K

**Tweets** **Tweets & replies**

**Earthquake Robot** @earthquakeBot · 8h A 5.3 magnitude earthquake occurred in South of the Kermadec Islands. Details: [eqbot.com/RcZ](http://eqbot.com/RcZ) Map: [google.com](http://google.com)

**Earthquake Robot** @earthquakeBot · 9h A 5.0 magnitude earthquake occurred 100.66mi N of Tobelo, Indonesia. Details: [eqbot.com/Rct](http://eqbot.com/Rct) Map:

**Big Ben**  
@big\_ben\_clock  
The first, established November 2009 & entirely unofficial. Copyright details here: [tinyurl.com/y1qbev3](http://tinyurl.com/y1qbev3), but still apparently imitated everywhere.  
Up a tower.  
Joined October 2009

Tweets 75.4K Followers 465K

**Tweets** **Tweets & replies**

**Big Ben** @big\_ben\_clock · 26m BONG BONG BONG  
Up a tower. 16 19

**Big Ben** @big\_ben\_clock · 1h BONG BONG  
Up a tower. 20 35

**Buy 2,500 Twitter Followers for \$29**

**VISIT** **WWW.MYSOCIALFANS.ORG**

**Calypso Cheverell**  
@CalypsoChevere1  
Follows you  
Quickly and efficiently get Thousands Twitter Followers per day. Buy 2500 Followers only \$29, see here [goo.gl/WHU3Ck](http://goo.gl/WHU3Ck)  
Joined July 2018

**Tweets** **Tweets & replies** **Media**

**Calypso Cheverell** @CalypsoChevere1 · 11h  
#BobbyBrownBET Grow up your visibility and community on Twitter. Get 2500 Twitter followers for \$29! See here [goo.gl/FFGbB7](http://goo.gl/FFGbB7)

# Botometer®

An OSoMe project (bot•o•meter)



Botometer (formerly BotOrNot) checks the activity of a Twitter account and gives it a score based on how likely the account is to be a bot. Higher scores are more bot-like.

Note the following:

- Use of this service requires authenticating with Twitter. For more, see [this FAQ](#).
- Botometer often categorizes "organizational accounts", like [@BarackObama](#), as bot accounts.
- If something's not working, [please let us know](#).

Botometer is a joint project of the Indiana University Network Science Institute ([IUNI](#)) and the Center for Complex Networks and Systems Research ([CNetS](#)).

Interested in using Botometer in your application? Check out our [API](#).

onurvarol Check user Check followers Check friends

Please help us improve Botometer by participating a short, anonymous survey. [Go to the survey ➔](#) X

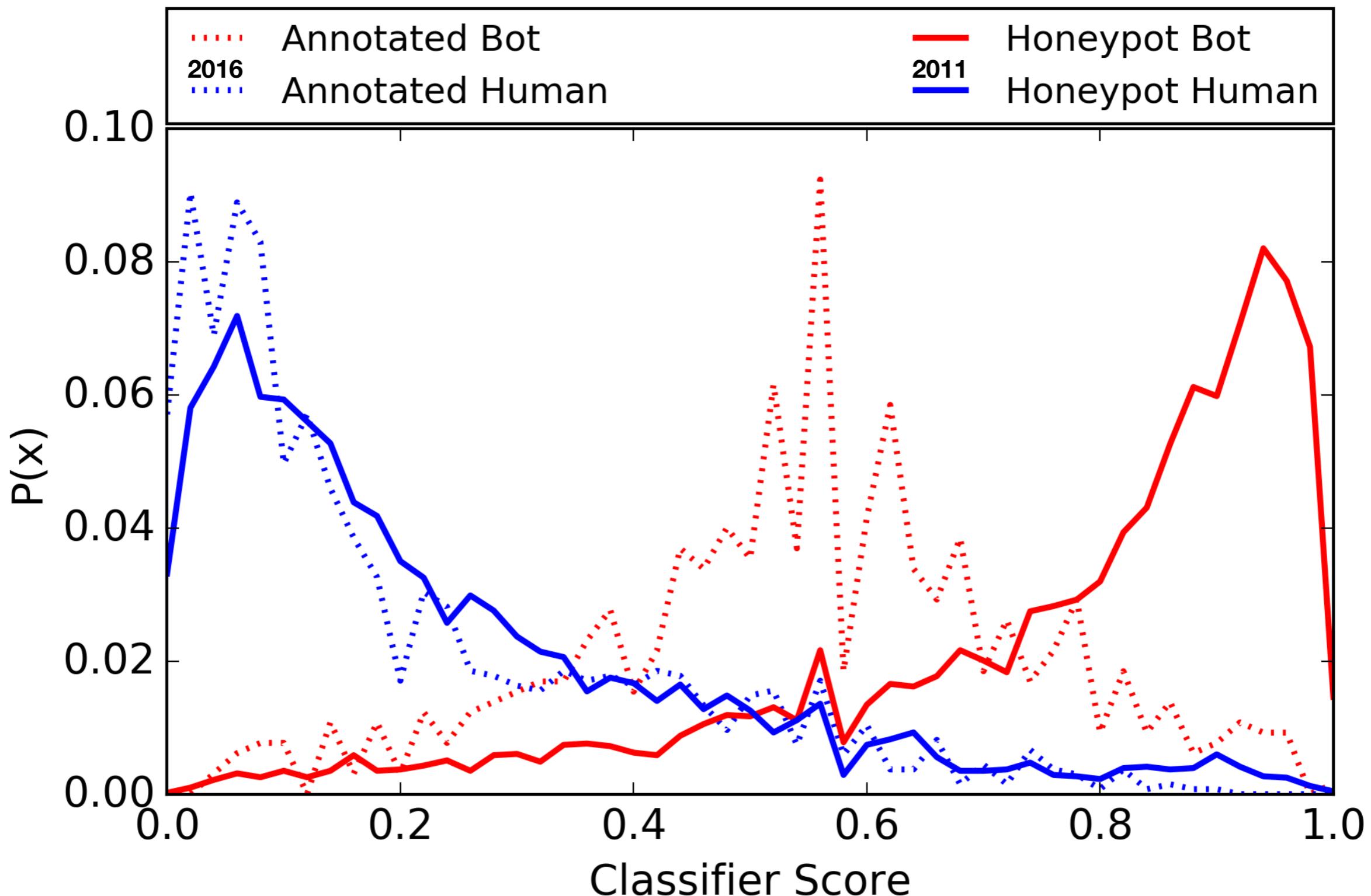
▼  @onurvarol

0.5 / 5  X

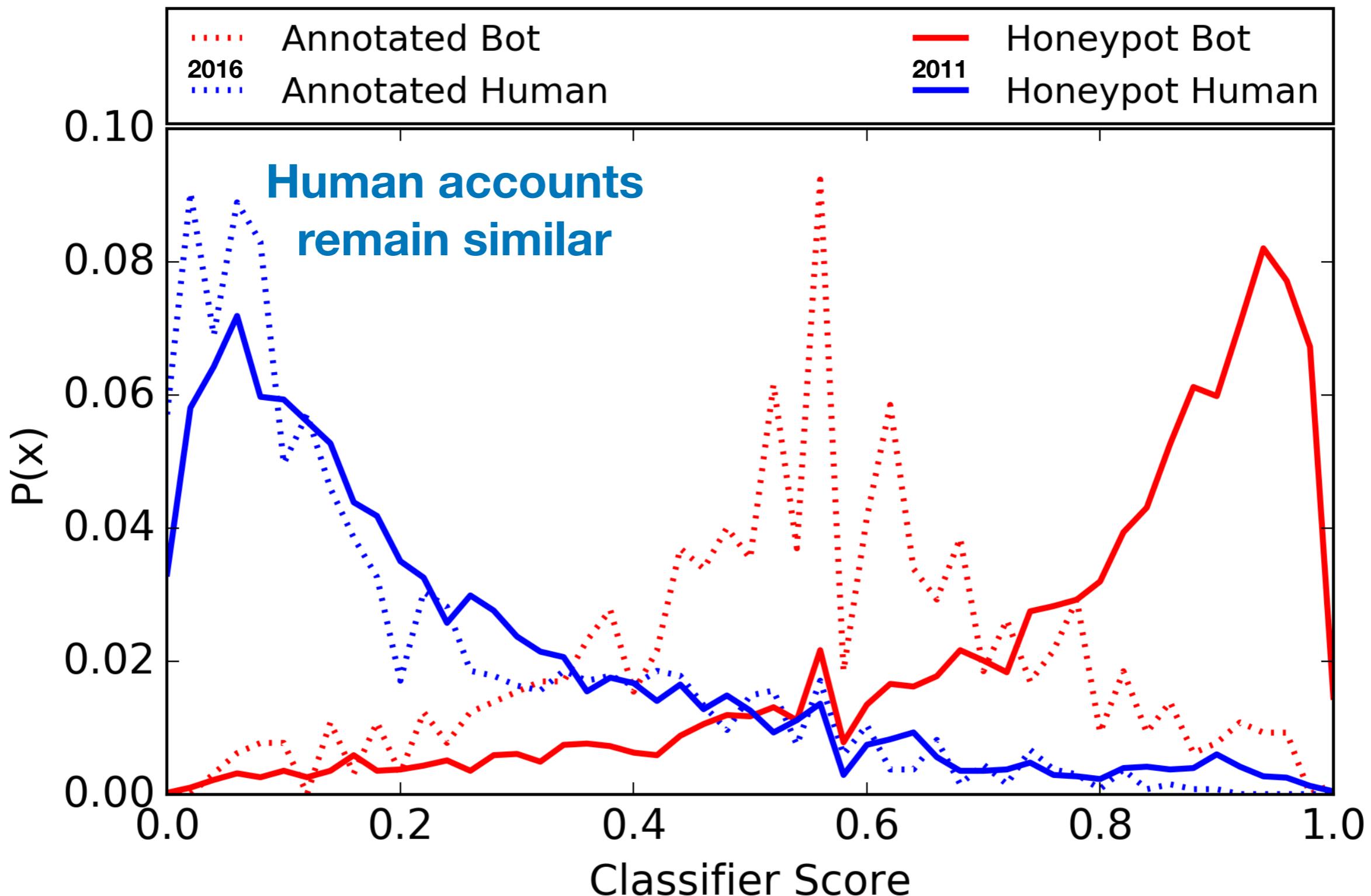
English-specific features	Language-independent features	Bot score based on	
Content: 1.1		All features: 0.5	<a href="#">Profile</a>
Sentiment: 1.0		Language-independent: 0.4	<a href="#">Tweet</a>
	Friend: 0.7	User profile language: en	<a href="#">Details</a>
	Network: 0.5	Complete Automation Probability: 0%	<a href="#">Feedback</a>
	Temporal: 0.4		
	User: 1.5		

[Export data](#)

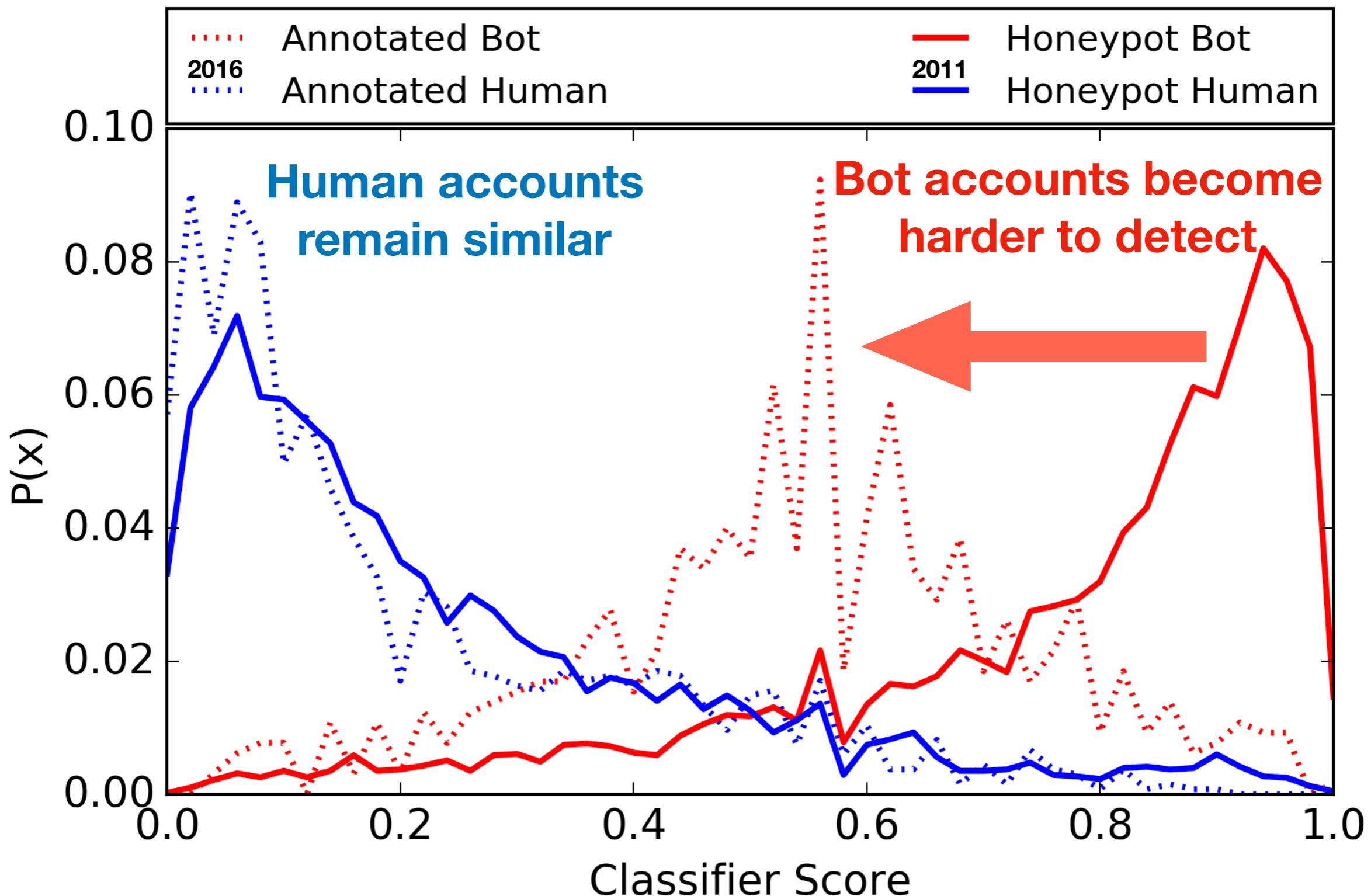
# Automated accounts are changing behavior over time

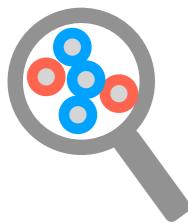


# Automated accounts are changing behavior over time



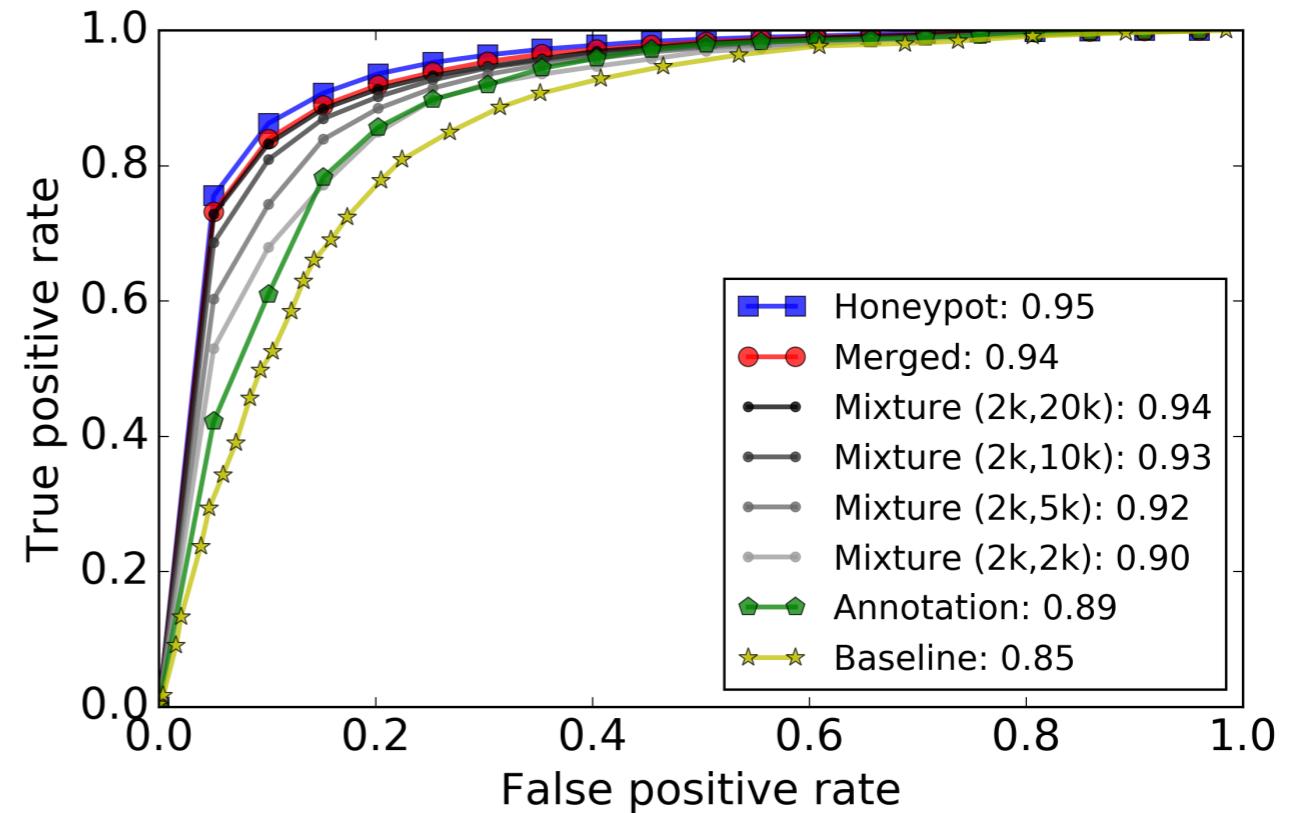
# Automated accounts are changing behavior over time

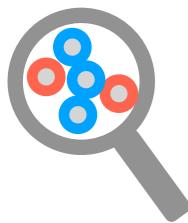




# Estimating bot population by tuning model sensitivity

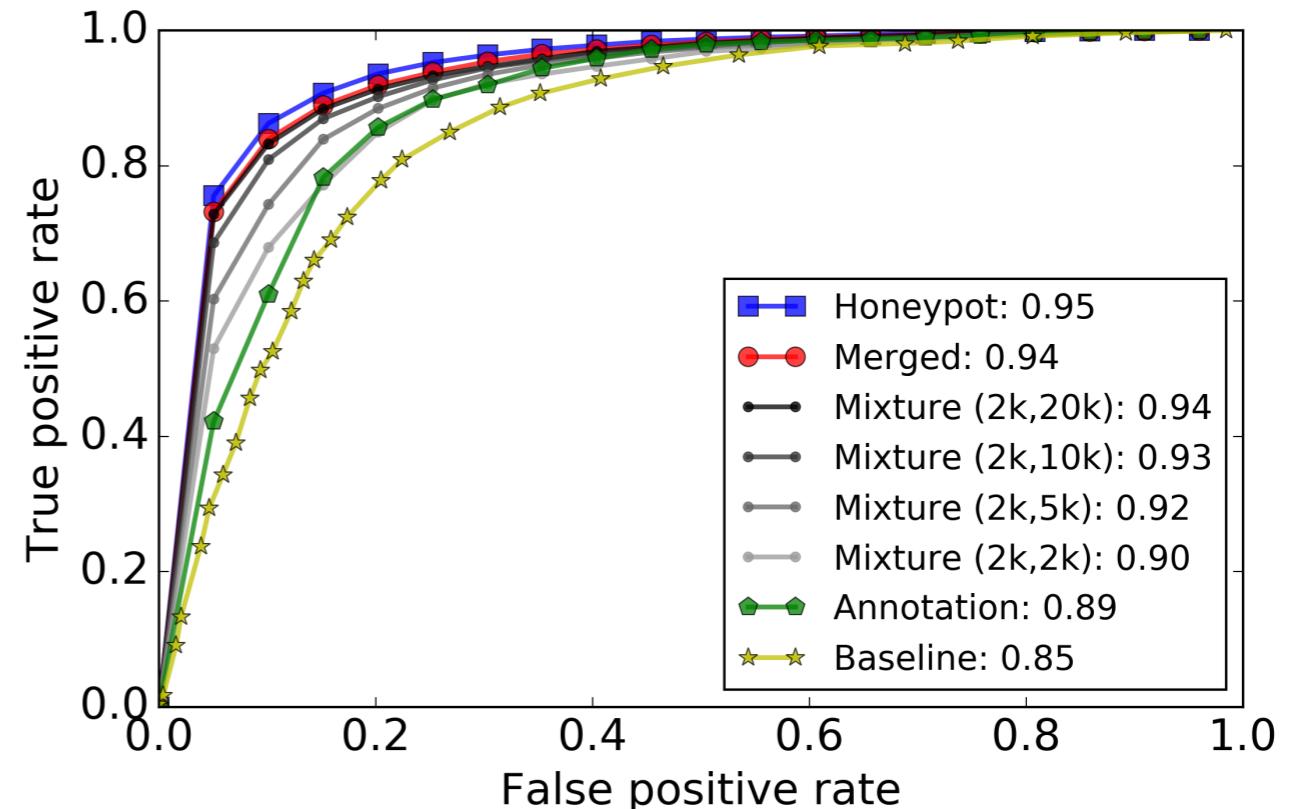
Trained on mixtures of honeypot and manually annotated data for tuning sensitivity to recent bot behaviors.





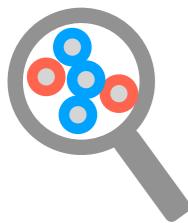
# Estimating bot population by tuning model sensitivity

Trained on mixtures of honeypot and manually annotated data for tuning sensitivity to recent bot behaviors.



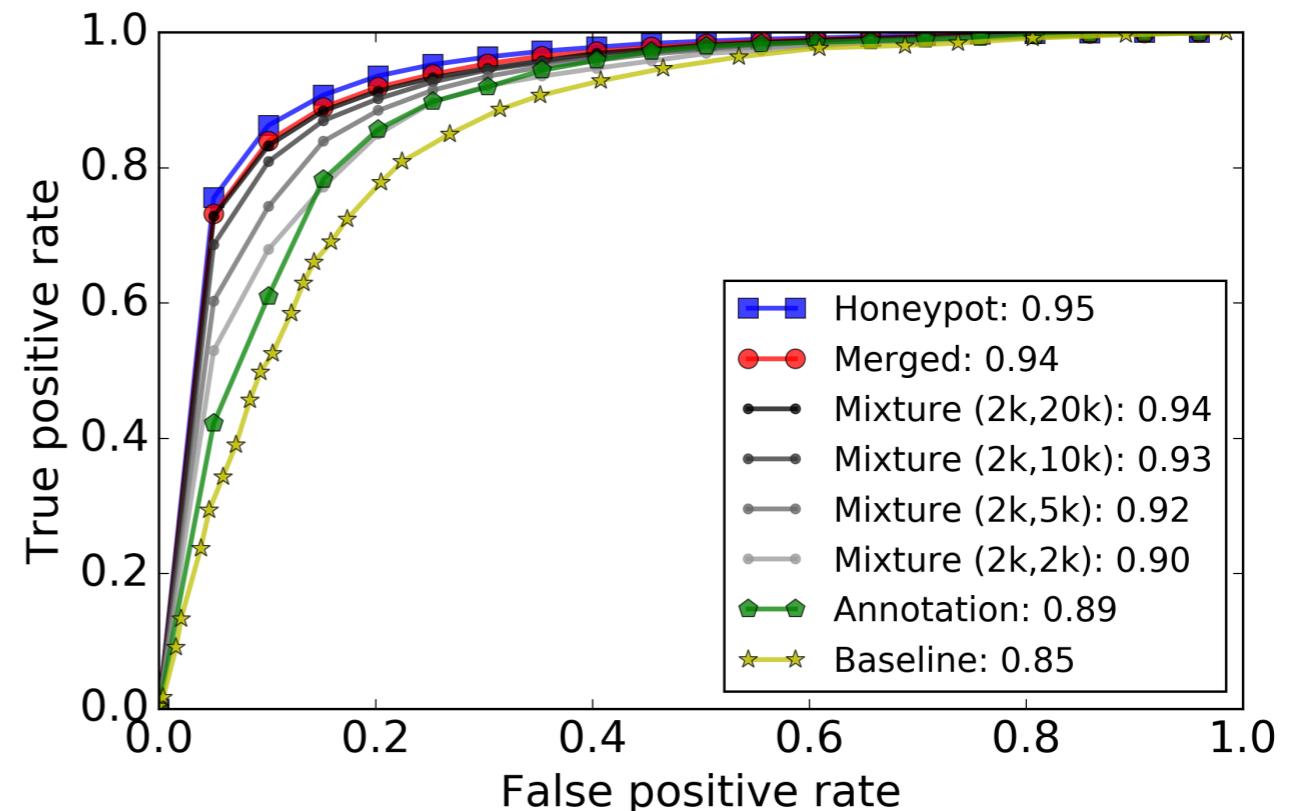
Adjusting sensitivity of detection system for sophisticated bots





# Estimating bot population by tuning model sensitivity

Trained on mixtures of honeypot and manually annotated data for tuning sensitivity to recent bot behaviors.



Adjusting sensitivity of detection system for sophisticated bots



9-15% of active accounts have automated behavior

**Botometer** is used carry out the largest study of Twitter human-bot ecosystem.

Between **9%-15% of active Twitter accounts estimated to have bot-like behavior.**

Social connectivity and information flow between humans and bots analyzed and behavior classes for accounts identified.

Varol, O., Ferrara, E., Davis, C. A., Flammini, A., & Menczer, F. **Online Human-Bot Interaction: Detection, Estimation, and Characterization** ICWSM 2017

**Senate Bill No. 1001**

**CHAPTER 892**

An act to add Chapter 6 (commencing with Section 17940) to Part 3 of Division 7 of the Business and Professions Code, relating to bots.

[ Approved by Governor September 28, 2018. Filed with Secretary of State September 28, 2018. ]

**LEGISLATIVE COUNSEL'S DIGEST**

SB 1001, Hertzberg. Bots: disclosure.

Existing law regulates various businesses to, among other things, preserve and regulate competition, prohibit unfair trade practices, and regulate advertising.

This bill would, with certain exceptions, make it unlawful for any person to use a bot to communicate or interact with another person in California online with the intent to mislead the other person about its artificial identity for the purpose of knowingly deceiving the person about the content of the communication in order to incentivize a purchase or sale of goods or services in a commercial transaction or to influence a vote in an election. The bill would define various terms for these purposes. The bill would make these provisions operative on July 1, 2019.

Vote: majority Appropriation: no Fiscal Committee: yes Local Program: no

**THE PEOPLE OF THE STATE OF CALIFORNIA DO ENACT AS FOLLOWS:**

**SECTION 1.** Chapter 6 (commencing with Section 17940) is added to Part 3 of Division 7 of the Business and Professions Code, to read:

**CHAPTER 6. Bots**

**17940.** For purposes of this chapter:

- (a) "Bot" means an automated online account where all or substantially all of the actions or posts of that account are not the result of a person.
- (b) "Online" means appearing on any public-facing Internet Web site, Web application, or digital application, including a social network or publication.
- (c) "Online platform" means any public-facing Internet Web site, Web application, or digital application, including a social network or publication, that has 10,000,000 or more unique monthly United States visitors or users for a majority of months during the preceding 12 months.
- (d) "Person" means a natural person, corporation, limited liability company, partnership, joint venture, association, estate, trust, government, governmental subdivision or agency, or other legal entity or any combination thereof.

**17941.** (a) It shall be unlawful for any person to use a bot to communicate or interact with another person in California online, with the intent to mislead the other person about its artificial identity for the purpose of knowingly deceiving the person about the content of the communication in order to incentivize a purchase or sale of goods or services in a commercial transaction or to influence a vote in an election. A person using a bot shall not be liable under this section if the person discloses that it is a bot.

(b) The disclosure required by this section shall be clear, conspicuous, and reasonably designed to inform persons with whom the bot communicates or interacts that it is a bot.

**17942.** (a) The duties and obligations imposed by this chapter are cumulative with any other duties or obligation imposed by any other law.

(b) The provisions of this chapter are severable. If any provision of this chapter or its application is held invalid, that invalidity shall not affect other provisions or applications that can be given effect without the invalid provision or application.

(c) This chapter does not impose a duty on service providers of online platforms, including, but not limited to, Web hosting and Internet service providers.

**17943.** This chapter shall become operative on July 1, 2019.

**Senate Bill No. 1001**

**CHAPTER 892**

An act to add Chapter 6 (commencing with Section 17940) to Part 3 of Division 7 of the Business and Professions Code, relating to bots.

[ Approved by Governor September 28, 2018. Filed with Secretary of State September 28, 2018. ]

**LEGISLATIVE COUNSEL'S DIGEST**

SB 1001, Hertzberg. Bots: disclosure.

Existing law regulates various businesses to, among other things, preserve and regulate competition, prohibit unfair trade practices, and regulate advertising.

This bill would, with certain exceptions, make it unlawful for any person to use a bot to communicate or interact with another person in California online with the intent to mislead the other person about its artificial identity for the purpose of knowingly deceiving the person about the content of the communication in order to incentivize a purchase or sale of goods or services in a commercial transaction or to influence a vote in an election. The bill would define various terms for these purposes. The bill would make these provisions operative on July 1, 2019.

Vote: majority Appropriation: no Fiscal Committee: yes Local Program: no

**THE PEOPLE OF THE STATE OF CALIFORNIA DO ENACT AS FOLLOWS:**

**SECTION 1.** Chapter 6 (commencing with Section 17940) is added to Part 3 of Division 7 of the Business and Professions Code, to read:

**CHAPTER 6. Bots**

**17940.** For purposes of this chapter:

(a) "Bot" means an automated online account where all or substantially all of the actions or posts of that account are not the result of a person.

(b) "Online" means appearing on any public-facing Internet Web site, Web application, or digital application, including a social network or publication.

(c) "Online platform" means any public-facing Internet Web site, Web application, or digital application, including a social network or publication, that has 10,000,000 or more unique monthly United States visitors or users for a majority of months during the preceding 12 months.

(d) "Person" means a natural person, corporation, limited liability company, partnership, joint venture, association, estate, trust, government, governmental subdivision or agency, or other legal entity or any combination thereof.

**17941.** (a) It shall be unlawful for any person to use a bot to communicate or interact with another person in California online, with the intent to mislead the other person about its artificial identity for the purpose of knowingly deceiving the person about the content of the communication in order to incentivize a purchase or sale of goods or services in a commercial transaction or to influence a vote in an election. A person using a bot shall not be liable under this section if the person discloses that it is a bot.

(b) The disclosure required by this section shall be clear, conspicuous, and reasonably designed to inform persons with whom the bot communicates or interacts that it is a bot.

**17942.** (a) The duties and obligations imposed by this chapter are cumulative with any other duties or obligation imposed by any other law.

(b) The provisions of this chapter are severable. If any provision of this chapter or its application is held invalid, that invalidity shall not affect other provisions or applications that can be given effect without the invalid provision or application.

(c) This chapter does not impose a duty on service providers of online platforms, including, but not limited to, Web hosting and Internet service providers.

**17943.** This chapter shall become operative on July 1, 2019.

- a) *The use and misuse of bots.* According to research conducted by the Center for Complex Networks and Systems Research at Indiana University and the Information Sciences Institute at the University of Southern California, “[i]ncreasing evidence suggests that a growing amount of social media content is generated by autonomous entities known as social bots.” (Onur Varol, et al., *Online Human-Bot Interactions: Detection, Estimation, and Characterization* (Mar. 27, 2017). Their estimates suggest that as many as 15 percent

Senate Bill No. 1001

CHAPTER 892

An act to add Chapter 6 (commencing with Section 17940) to Part 3 of Division 7 of the Business and Professions Code, relating to bots.

[ Approved by Governor September 28, 2018. Filed with Secretary of State September 28, 2018. ]

LEGISLATIVE COUNSEL'S DIGEST

SB 1001, Hertzberg. Bots: disclosure.

Existing law regulates various businesses to, among other things, preserve and regulate competition, prohibit unfair trade practices, and regulate advertising.

This bill would, with certain exceptions, make it unlawful for any person to use a bot to communicate or interact with another person in California online with the intent to mislead the other person about its artificial identity for the purpose of knowingly deceiving the person about the content of the communication in order to incentivize a purchase or sale of goods or services in a commercial transaction or to influence a vote in an election. The bill would define various terms for these purposes. The bill would make these provisions operative on July 1, 2019.

Vote: majority Appropriation: no Fiscal Committee: yes Local Program: no

THE PEOPLE OF THE STATE OF CALIFORNIA DO ENACT AS FOLLOWS:

**SECTION 1.** Chapter 6 (commencing with Section 17940) is added to Part 3 of Division 7 of the Business and Professions Code, to read:

**CHAPTER 6. Bots**

**17940.** For purposes of this chapter:

- (a) "Bot" means an automated online account where all or substantially all of the actions or posts of that account are not the result of a person.
- (b) "Online" means appearing on any public-facing Internet Web site, Web application, or digital application, including a social network or publication.
- (c) "Online platform" means any public-facing Internet Web site, Web application, or digital application, including a social network or publication, that has 10,000,000 or more unique monthly United States visitors or users for a majority of months during the preceding 12 months.
- (d) "Person" means a natural person, corporation, limited liability company, partnership, joint venture, association, estate, trust, government, governmental subdivision or agency, or other legal entity or any combination thereof.

**17941.** (a) It shall be unlawful for any person to use a bot to communicate or interact with another person in California online, with the intent to mislead the other person about its artificial identity for the purpose of knowingly deceiving the person about the content of the communication in order to incentivize a purchase or sale of goods or services in a commercial transaction or to influence a vote in an election. A person using a bot shall not be liable under this section if the person discloses that it is a bot.

(b) The disclosure required by this section shall be clear, conspicuous, and reasonably designed to inform persons with whom the bot communicates or interacts that it is a bot.

**17942.** (a) The duties and obligations imposed by this chapter are cumulative with any other duties or obligation imposed by any other law.

(b) The provisions of this chapter are severable. If any provision of this chapter or its application is held invalid, that invalidity shall not affect other provisions or applications that can be given effect without the invalid provision or application.

(c) This chapter does not impose a duty on service providers of online platforms, including, but not limited to, Web hosting and Internet service providers.

**17943.** This chapter shall become operative on July 1, 2019.

- a) *The use and misuse of bots.* According to research conducted by the Center for Complex Networks and Systems Research at Indiana University and the Information Sciences Institute at the University of Southern California, “[i]ncreasing evidence suggests that a growing amount of social media content is generated by autonomous entities known as social bots.” (Onur Varol, et al., *Online Human-Bot Interactions: Detection, Estimation, and Characterization* (Mar. 27, 2017). Their estimates suggest that as many as 15 percent

Some bots are virtually invisible as they carry out their work behind the scenes of search engines and other Internet Web sites. Some are easier to spot. However, whether or not it is apparent, bots and people regularly interact online. For instance, “[s]pambots show up in comment sections and Twitter bots clog people’s timelines with everything from marketing, to political campaigning, to social activism, to utter nonsense.” (*Id.*) “Social media make it easy for accounts controlled by hybrid or automated approaches to create content and interact with other accounts.” (Onur Varol, et al., *Online Human-Bot Interactions: Detection, Estimation, and Characterization* (Mar. 27, 2017) <<https://arxiv.org/pdf/1703.03107.pdf>> [as of Apr. 6, 2018].)

# Social bots activities during elections

The screenshot shows a report from the Pew Research Center's Internet & Technology section. The title is "Bots in the Twittersphere". It states that "An estimated two-thirds of tweeted links to popular websites are posted by automated accounts – not human beings". The author is Stefan Wójcik, Solomon Messing, Aaron Smith, Lee Rainie, and Paul Hitlin. A video thumbnail titled "BOTOMETER" is shown, featuring a man with glasses and a beard. The video duration is 4:25.

The screenshot shows an article from Science magazine. The title is "Social media 'bots' tried to influence the U.S. election. Germany may be next". The author is Kai Kupferschmidt. The date is Sep. 13, 2017, 3:45 PM. The image features a hand holding a pair of glasses over a smartphone, with Twitter birds floating around it. Below the image is the text "ADAPTED FROM PE-ART/ISTOCKPHOTO BY G. GRULLÓN/SCIENCE".

The screenshot shows a Poynter article. The title is "Bots spread a lot of fakery during the 2016 election. But they can also debunk it.". The image is a network graph showing connections between various news organizations like CNN, BBC, and Reuters, with red and blue nodes representing different bot networks.

The screenshot shows an article from nature communications. The title is "The spread of low-credibility content by social bots". The authors are Chengcheng Shao, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Alessandro Flammini & Filippo Menczer. The date is Published: 20 November 2018.

The screenshot shows an article from nature communications. The title is "Influence of fake news in Twitter during the 2016 US presidential election". The authors are Alexandre Bovet & Hernán A. Makse. The date is Published: 02 January 2019.

The screenshot shows an article from Science. The title is "Fake news on Twitter during the 2016 U.S. presidential election". The authors are Nir Grinberg<sup>1,2,\*</sup>, Kenneth Joseph<sup>3,\*</sup>, Lisa Friedland<sup>1,\*</sup>, Briony Swire-Thompson<sup>1,2</sup>, David Lazer<sup>1,2†</sup>. The date is 25 Jan 2019: Vol. 363, Issue 6425, pp. 374-378. DOI: 10.1126/science.aau2706

The screenshot shows an article from First Monday. The title is "Social Bots Distort the 2016 US Presidential Election Online Discussion". The date is Volume 21, Number 11 - 7 November 2016. The authors are Alessandro Bessi, Emilio Ferrara, Bence Kollanyi, Philip N. Howard, and Samuel C. Woolley. The date is Posted: 8 Jun 2017.

Bots and Automation over Twitter during the U.S. Election  
[COMPROP DATA MEMO 2016.4 / 17 NOV 2016](#)

Philip N. Howard  
Oxford University  
philip.howard@oii.ox.ac.uk  
@pnhoward

Samuel C. Woolley  
University of Washington  
samwooll@uw.edu  
@samuelwoolley

# The science of fake news

David M. J. Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, Jonathan L. Zittrain

*The list of author affiliations is provided in the supplementary materials.*

Email: [d.lazer@northeastern.edu](mailto:d.lazer@northeastern.edu)

- Hide authors and affiliations

Science 09 Mar 2018:  
Vol. 359, Issue 6380, pp. 1094-1096  
DOI: 10.1126/science.aao2998

- A new system of safeguards is needed.
- Interventions might be needed at stemming the flow and influence of fake news.
- Empowering individuals: Fact-checking, awareness unconscious biases, media literacy
- Platform-based detection and intervention: Reports on source quality, algorithmic biases leading echo-chambers, curb automated activities

# The spread of true and false news online

Soroush Vosoughi<sup>1</sup>, Deb Roy<sup>1</sup>, Sinan Aral<sup>2,\*</sup>

<sup>1</sup>Massachusetts Institute of Technology (MIT), the Media Lab, E14-526, 75 Amherst Street, Cambridge, MA 02142, USA.

<sup>2</sup>MIT, E62-364, 100 Main Street, Cambridge, MA 02142, USA.

✉ \*Corresponding author. Email: sinan@mit.edu

- Hide authors and affiliations

Science 09 Mar 2018:  
Vol. 359, Issue 6380, pp. 1146-1151  
DOI: 10.1126/science.aap9559

- Falsehood diffused significantly farther, faster, deeper, and more broadly than the truth in all categories.
- Novelty is an important factor. False news was perceived as more novel than true news, which suggests that people are more likely to share novel information.
- Controlling for many factors, false news was 70% more likely to be retweeted than the truth.

# Dataset to study information dissemination

Dataset of news sources and dissemination of those URL on Twitter is used to study to study dissemination of **misinformation** and roles of **social bots**.

**Low-credibility sources** obtained by merging several lists compiled by third-party news and fact-checking organizations or experts. We consider 120 domains to collect information.

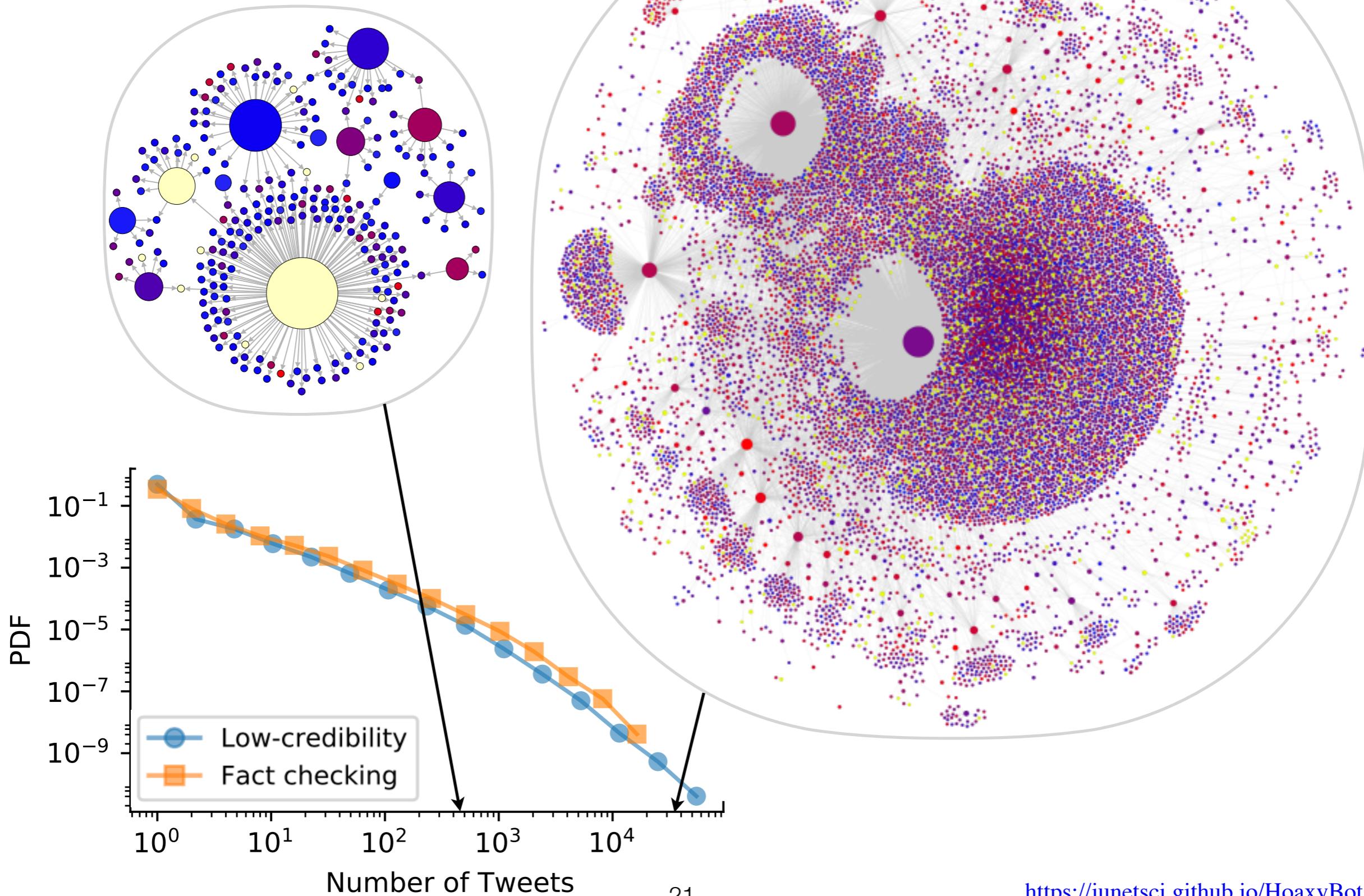
**Fact-checking** information gathered from 7 independent organization.

We monitor activity between May 2016 to March 2017 for collecting dataset and Hoaxy platform serves this data and more online.

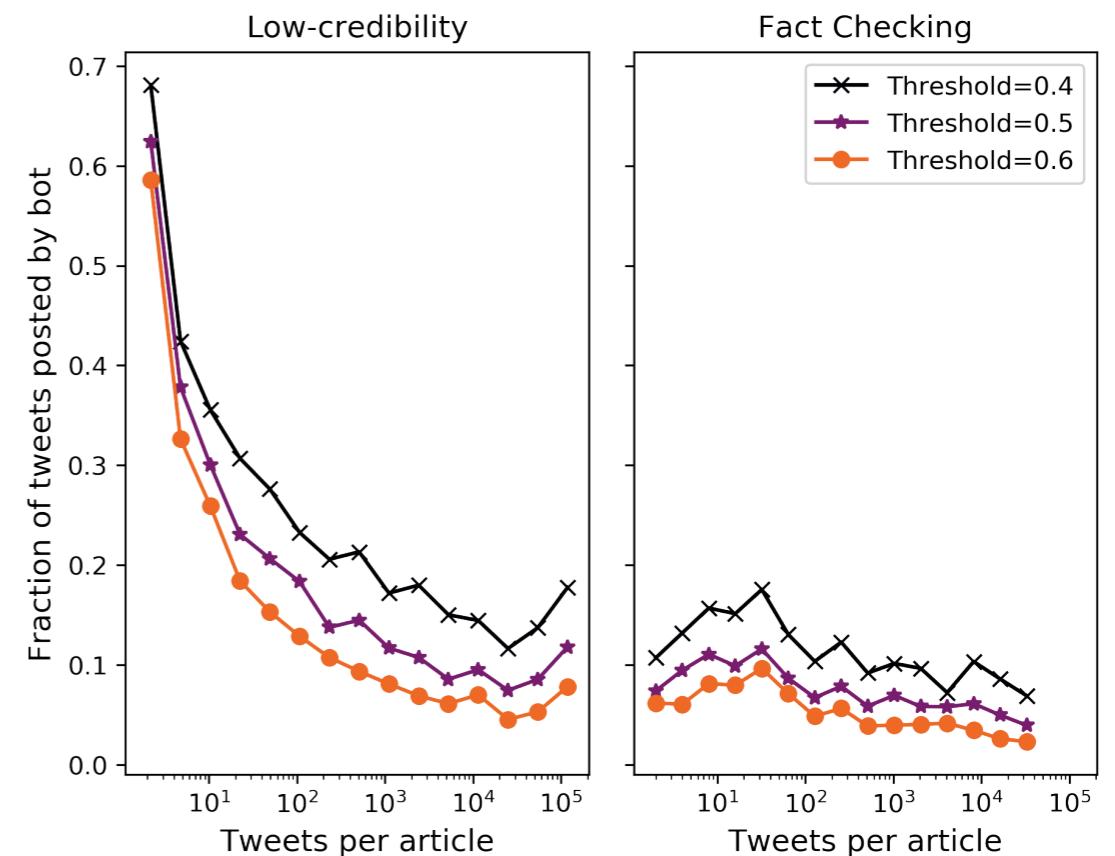
- Low-credibility: 390k articles and 13.6M tweets
- Fact-checking: 15k articles and 1.1M tweets



# Virality distribution of low-credibility and fact-checking content

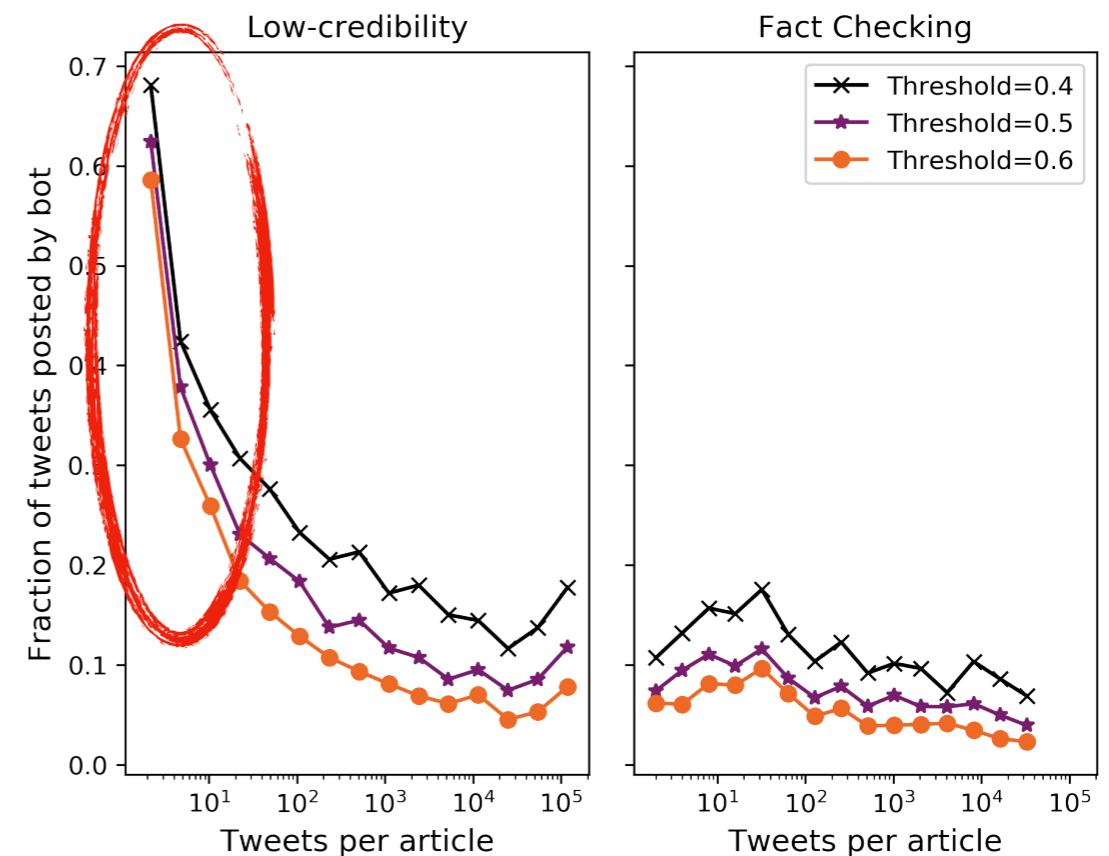


# Bots disseminate low-credibility content first



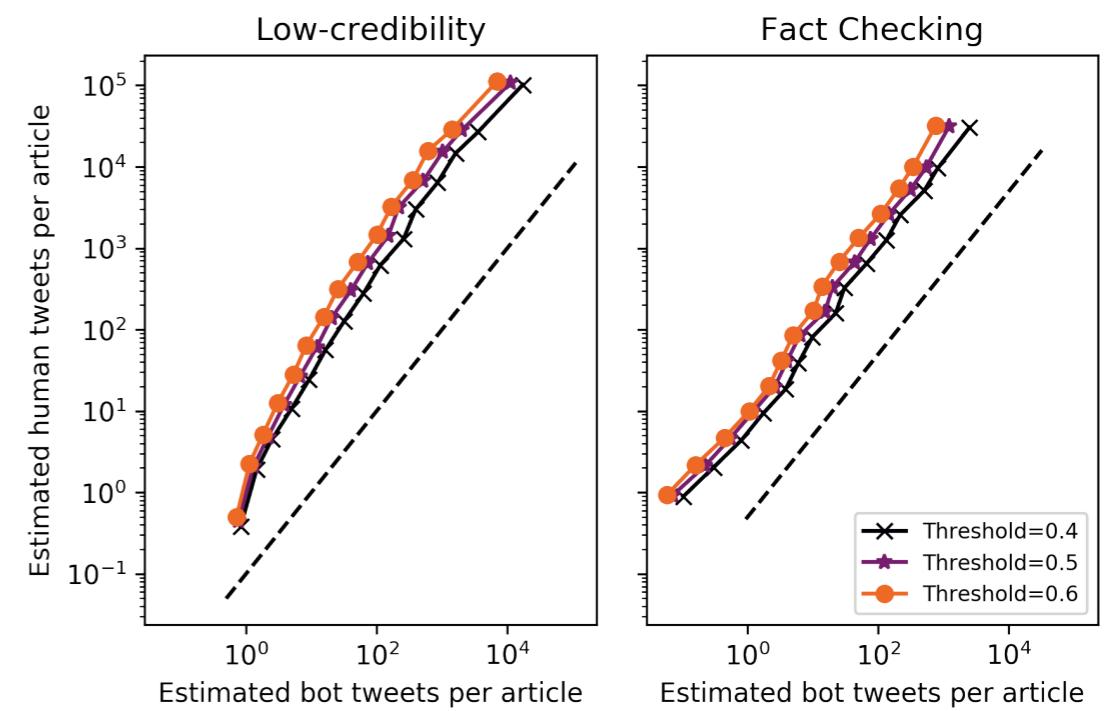
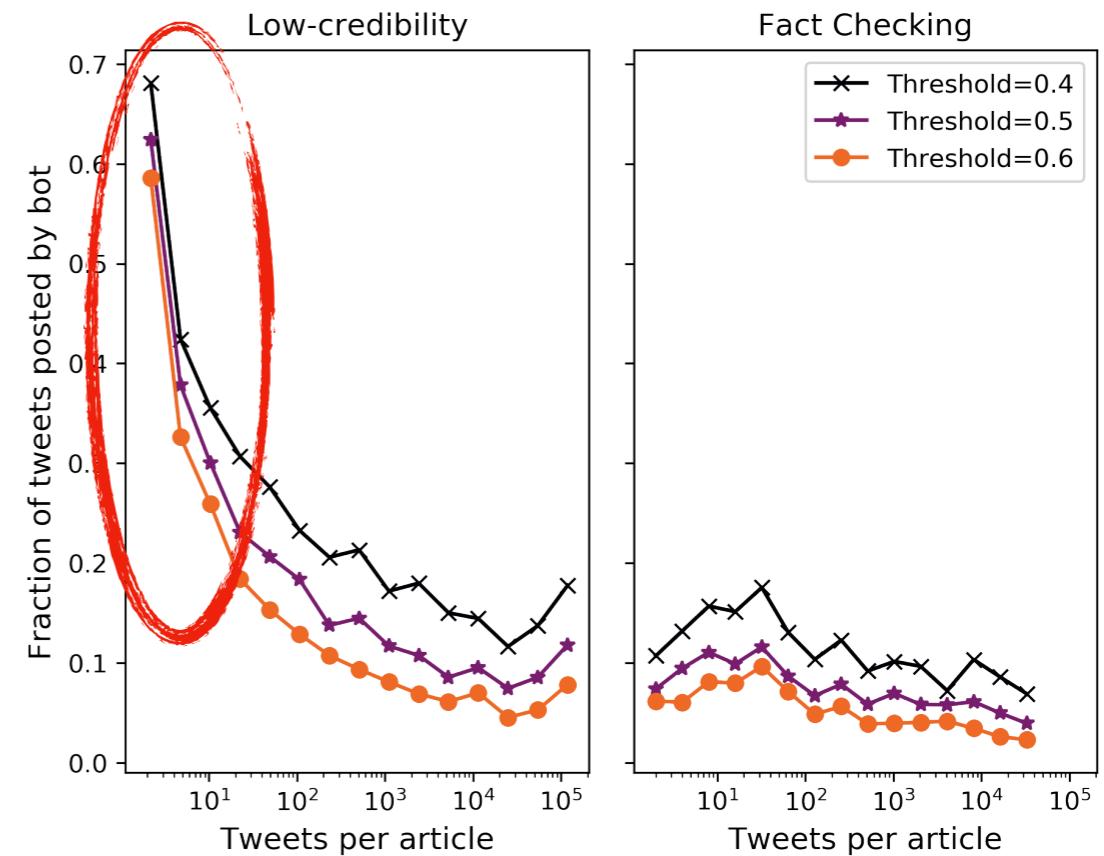
# Bots disseminate low-credibility content first

The fraction of tweets posted by likely bots is much higher for articles from low-credibility sources. As the popularity of the article increase bot fraction diluted by human response.



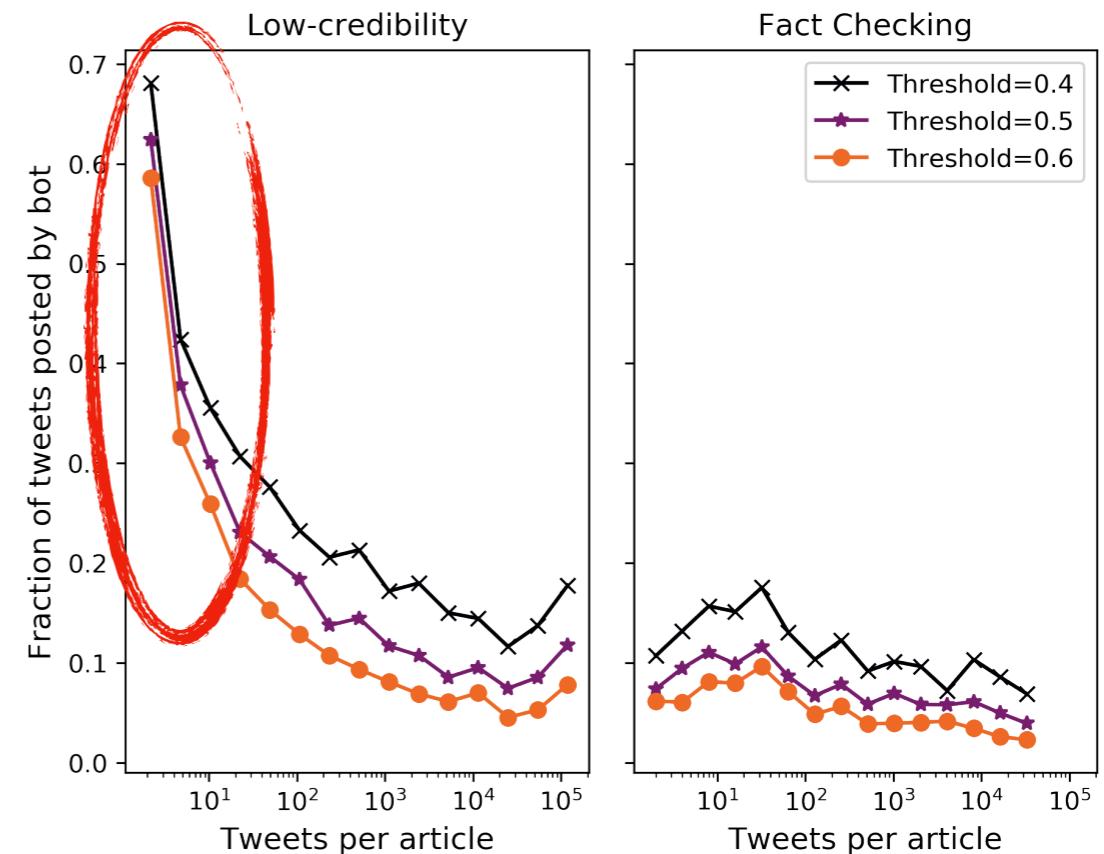
# Bots disseminate low-credibility content first

The fraction of tweets posted by likely bots is much higher for articles from low-credibility sources. As the popularity of the article increase bot fraction diluted by human response.

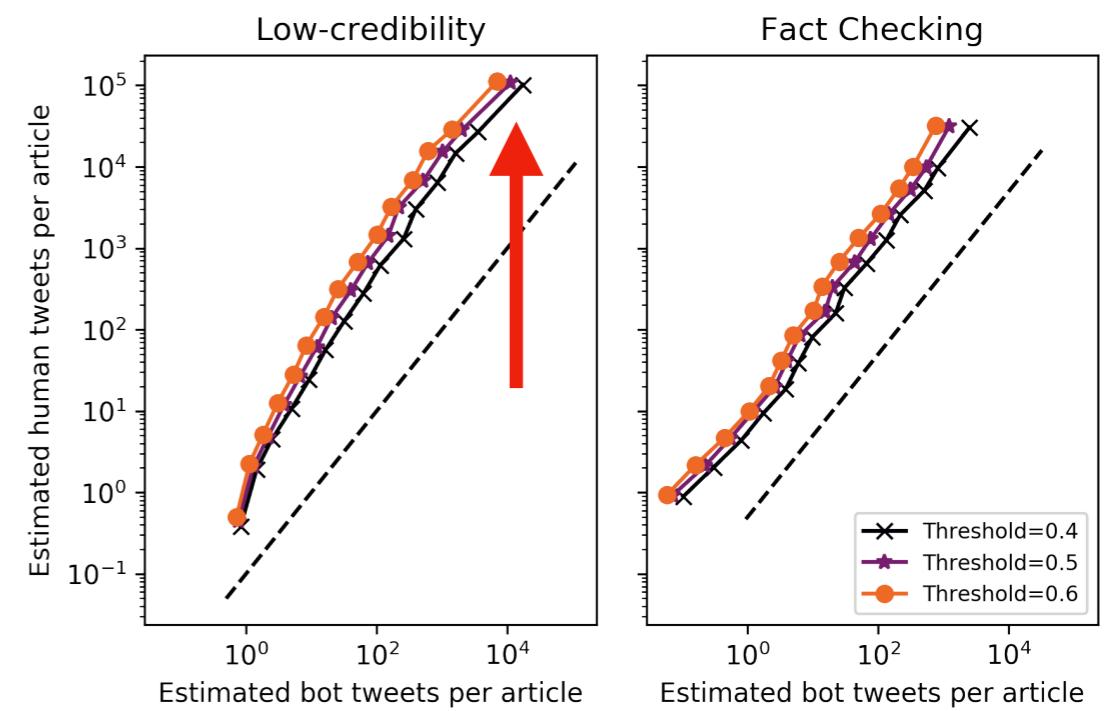


# Bots disseminate low-credibility content first

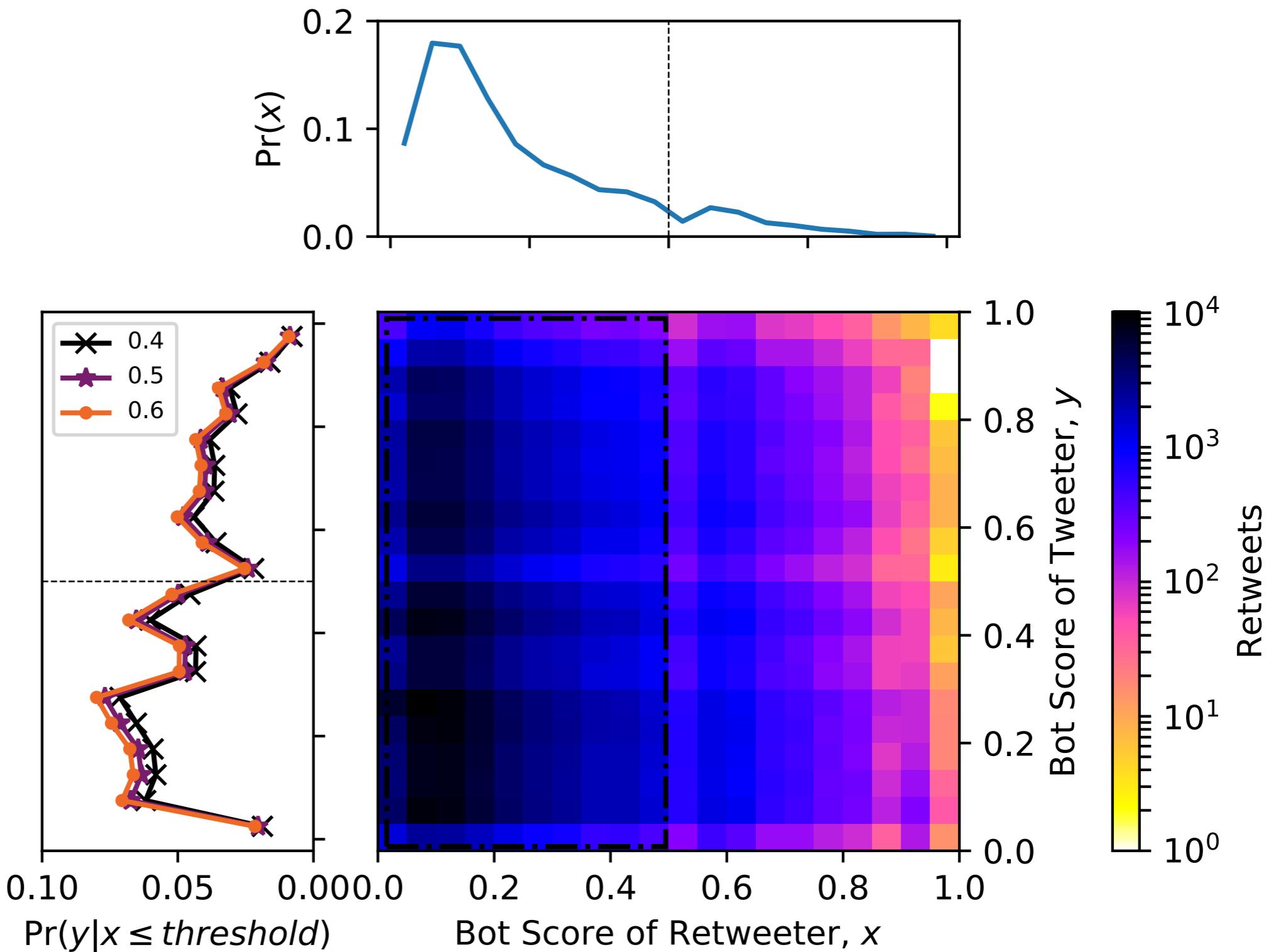
The fraction of tweets posted by likely bots is much higher for articles from low-credibility sources. As the popularity of the article increase bot fraction diluted by human response.



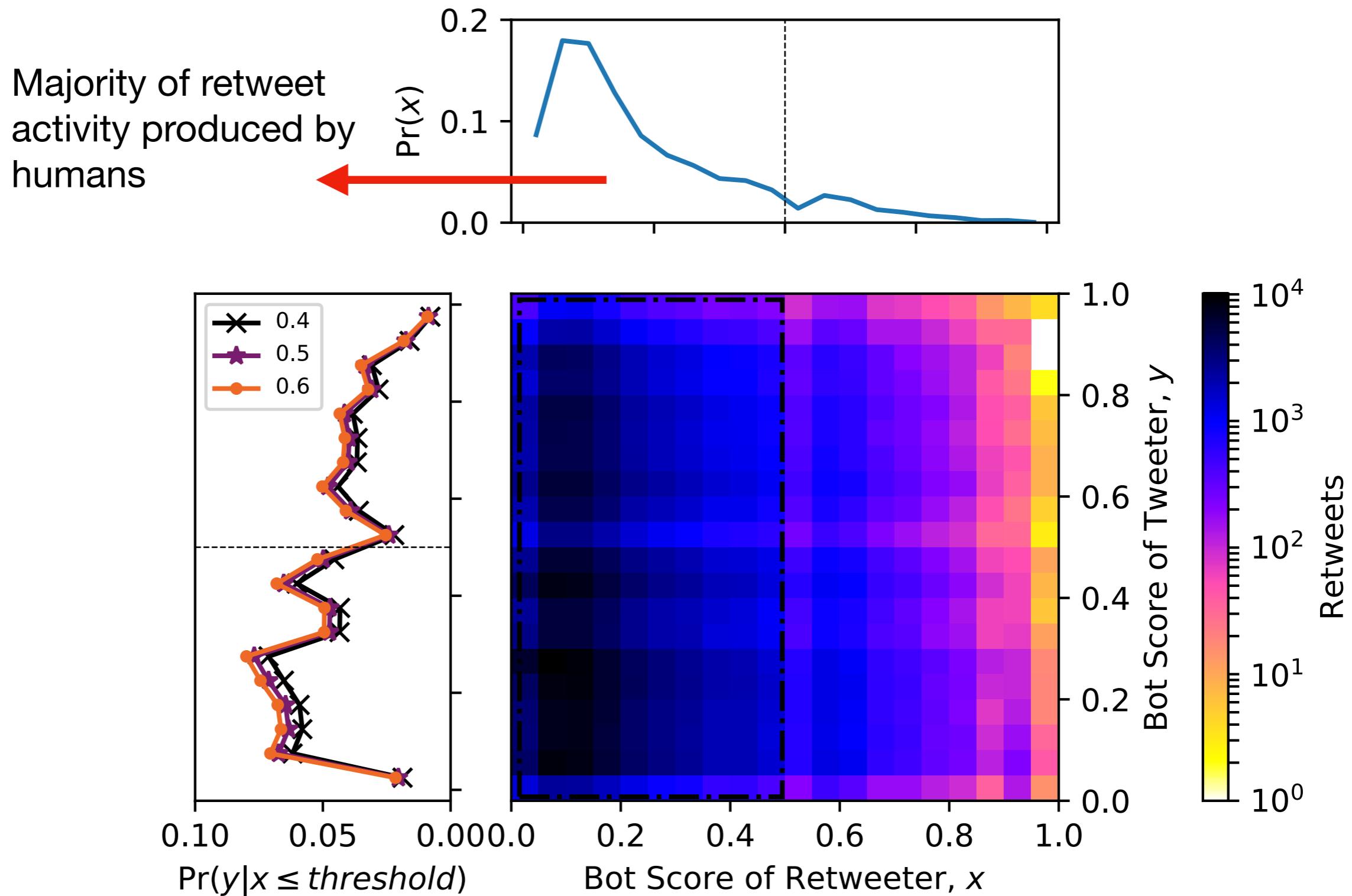
The estimated number of human tweets per article grows faster than bot tweets for low-credibility articles.



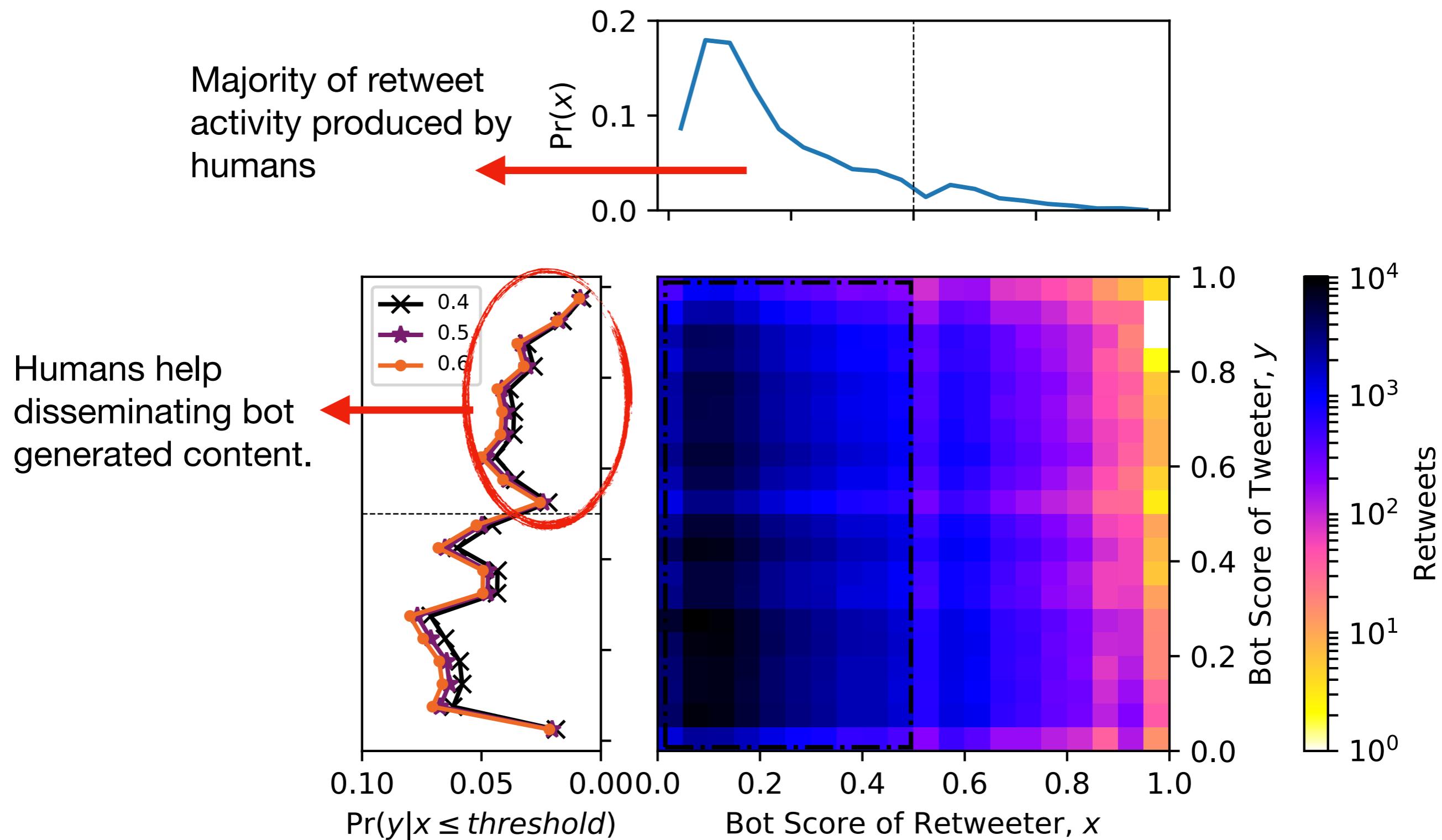
# Bot generated content retweet mostly by humans



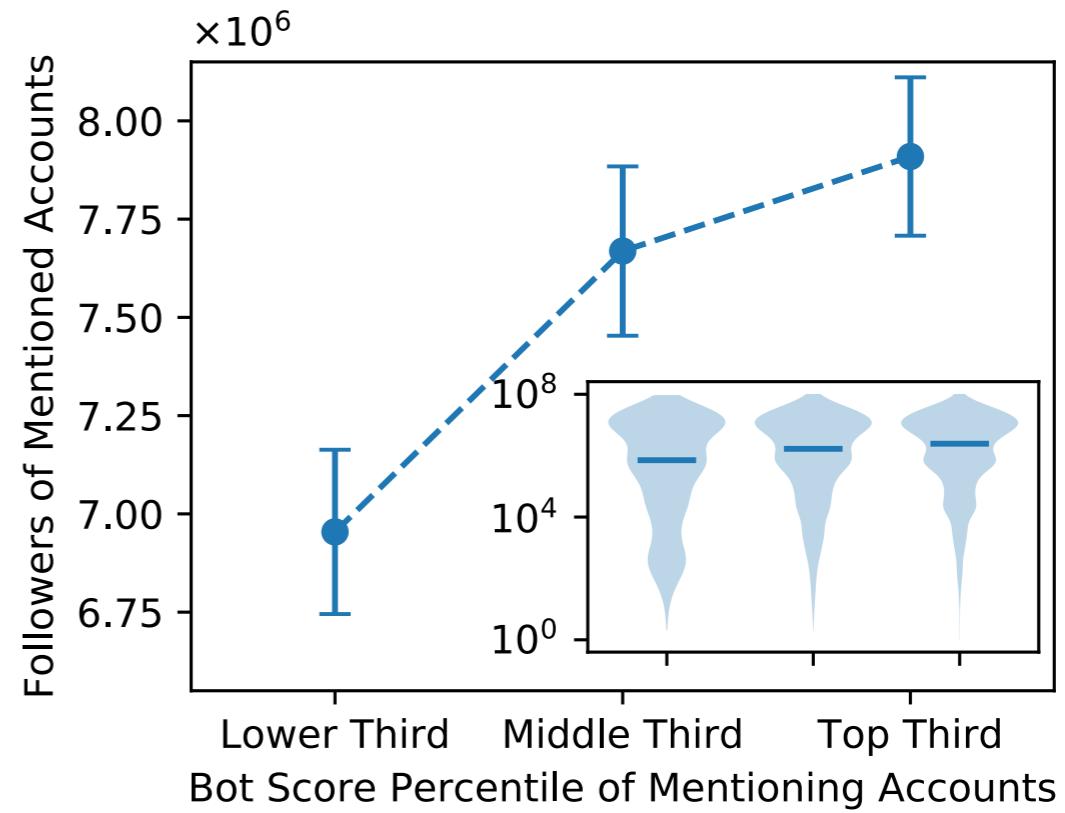
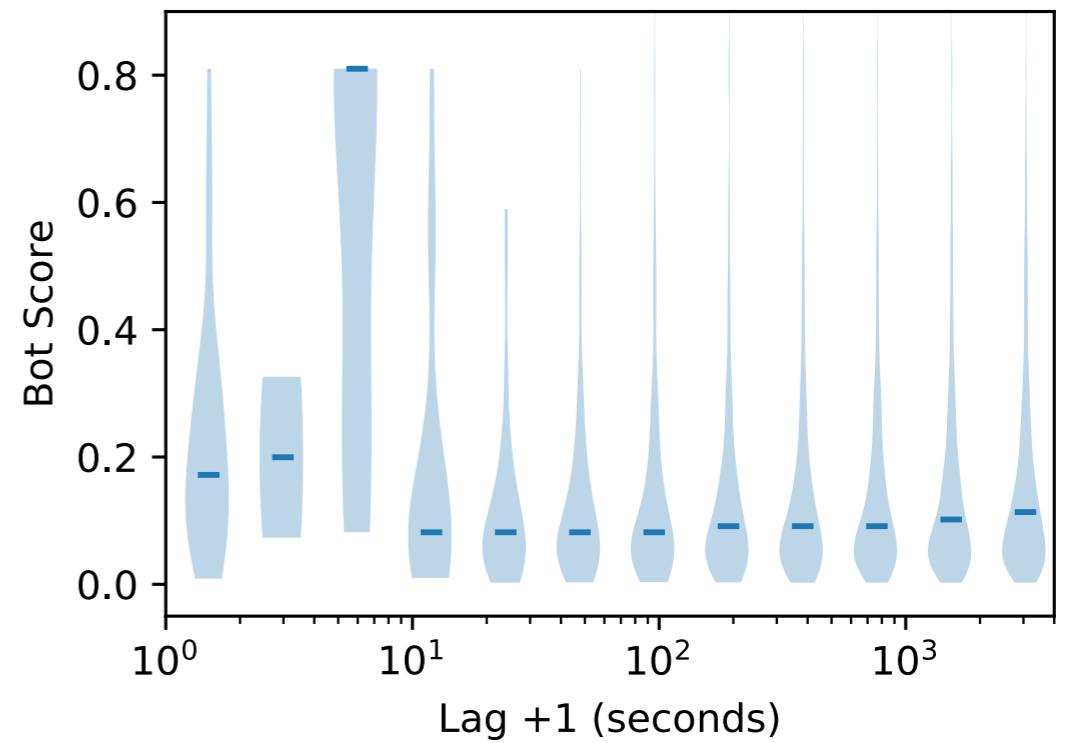
# Bot generated content retweet mostly by humans



# Bot generated content retweet mostly by humans

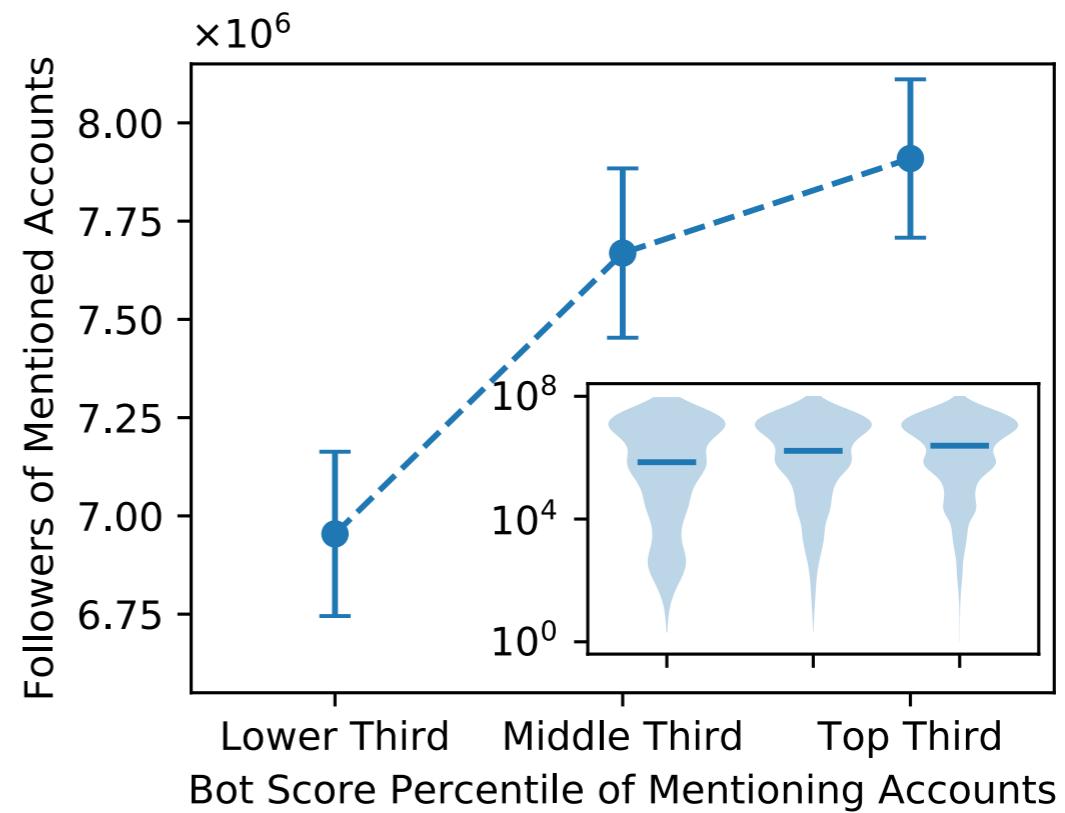
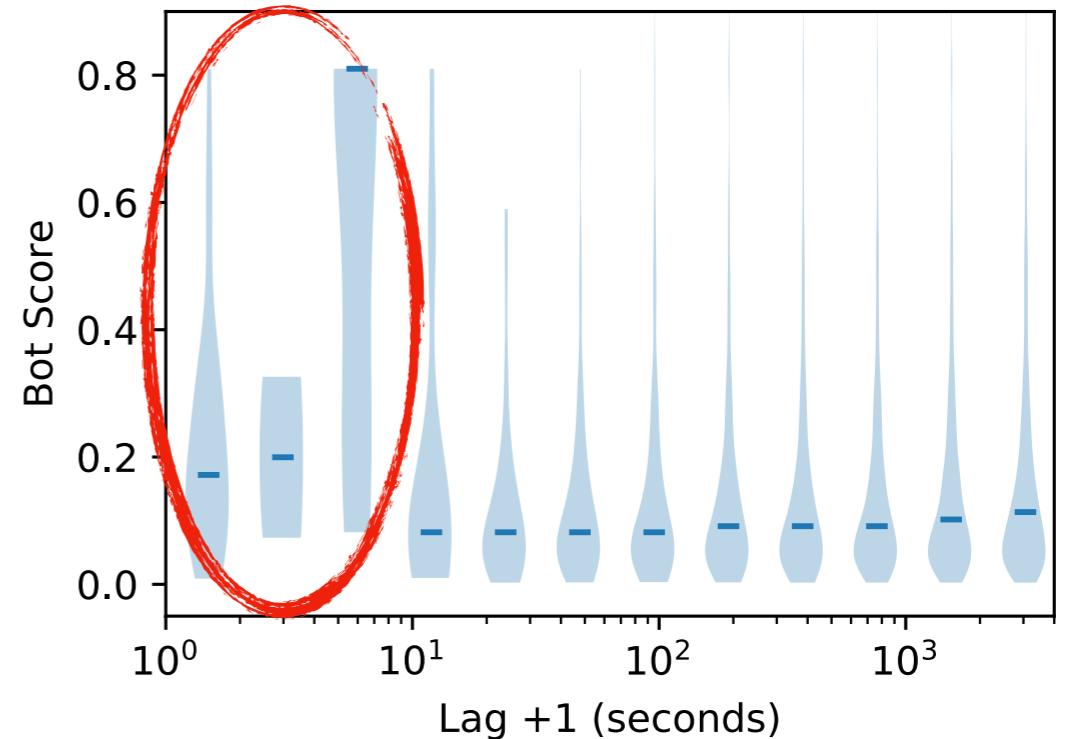


# Strategies employed by bots



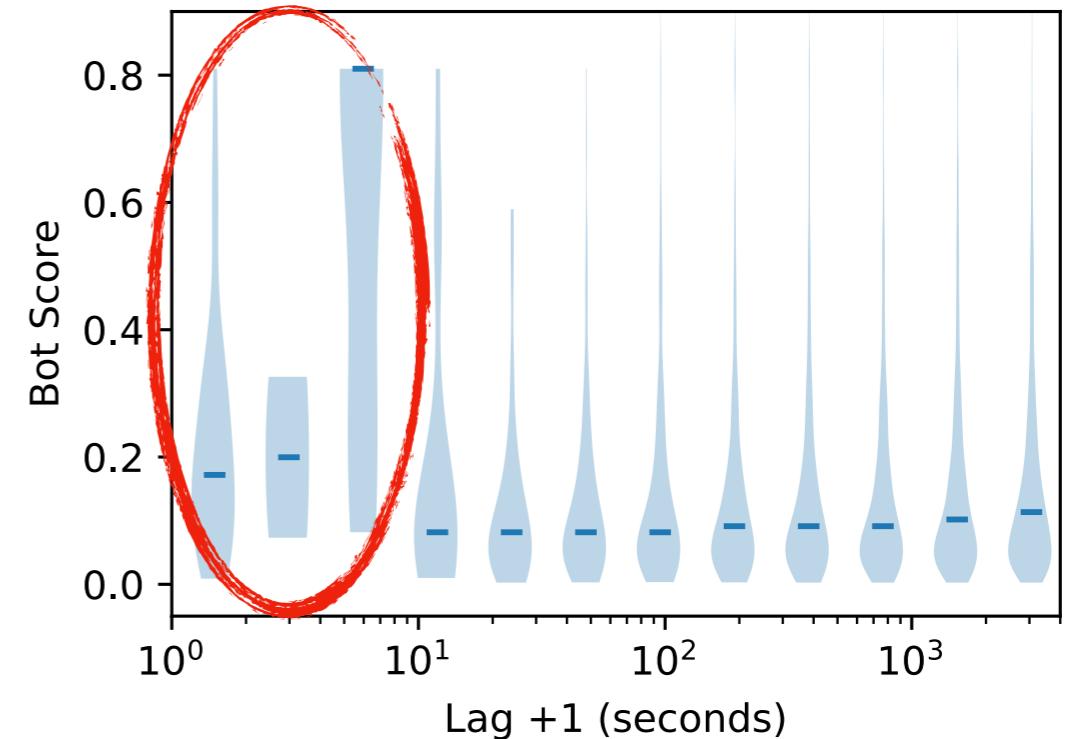
# Strategies employed by bots

Early bot support after a viral low-credibility article is first shared. Bots boost initial popularity to target human biases in retweeting behavior.

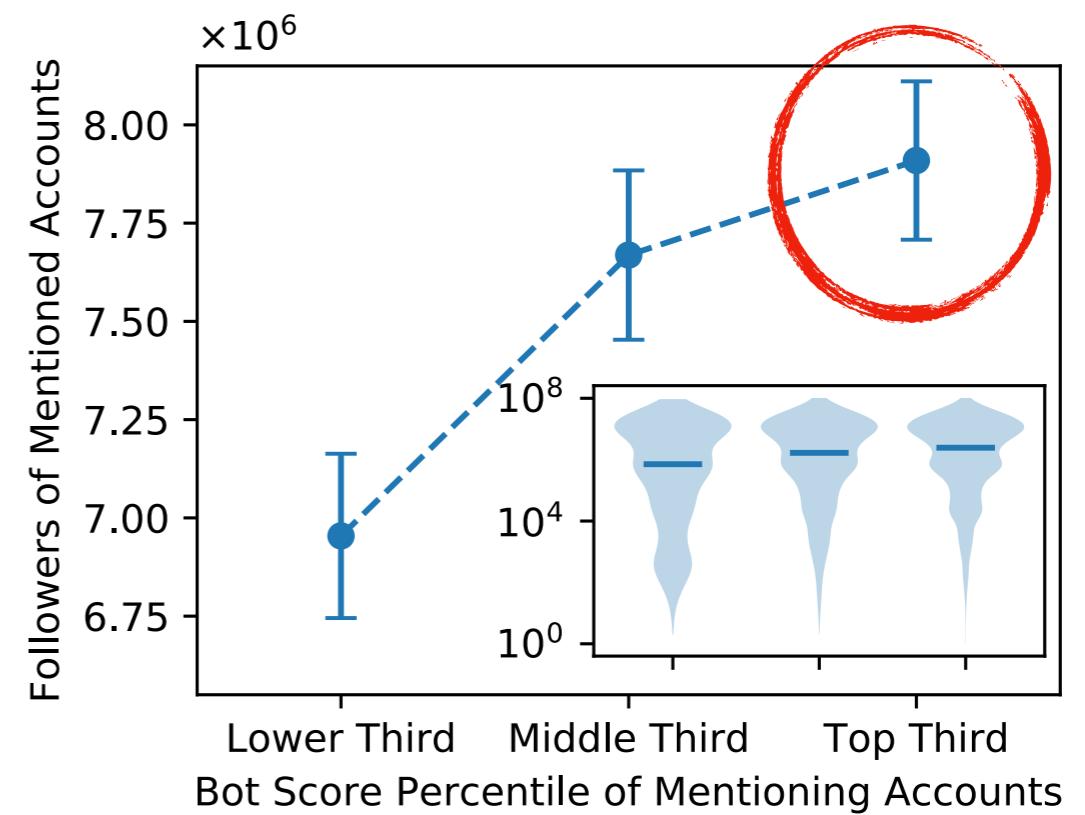


# Strategies employed by bots

Early bot support after a viral low-credibility article is first shared. Bots boost initial popularity to target human biases in retweeting behavior.



Targeting of influential accounts. Bots mention or reply popular accounts to catch their attention.





Low-credibility and fact-checking articles hardly distinguishable based on their popularity.

**Botometer** used to analyze accounts creating and disseminating news sources.

Social **bots were instrumental** during 2016 US Presidential Election.

Bots employ different **strategies to boost article popularity** and **engage with influencer** accounts on network.

Shao, C., Ciampaglia, G. L., Varol, O., Yang, K. C., Flammini, A., & Menczer, F. (2018). **The spread of low-credibility content by social bots.** *Nature communications*, 9(1), 4787.

# Bot Repository

Increasing evidence suggests that a growing amount of social media content is generated by autonomous entities known as social bots. Many social bots perform useful functions, but there is a growing record of malicious applications of social bots. We believe it is important to provide public datasets and tools that help identification of social bots, since deception and detection technologies are in an arms race.

**Bot repository** is a centralized place to share annotated datasets of Twitter social bots. We also provide list of available tools on bot detection.

## Datasets

Collection of annotated social bot datasets. Links to original source (webpage, publication, etc.) is provided in the details.

[View details »](#)

## Tools

Links to publicly available tools on social bot detection. Information on webpage, publication, etc. is provided in the details.

[View details »](#)

## Contribute

We always need more publicly available datasets and tools. Please provide information about your awesome dataset or tool. We would like to feature them in our website.

[View details »](#)

✉ Main contact for inquiries: [Onur Varol](#)

## Botometer

[FAQ](#)
[API](#)
[Publications](#)
[Bot Repo](#)


# API

So you want to use Botometer in your application? These client libraries, examples, and tools can help.

A free [Mashape Marketplace](#) account is required in order to use our API.

[Email us](#) if you're interested in licensing Botometer.

# Official

We provide and maintain the following client libraries. Please [let us know](#) if you find errors or have suggestions.

## HTTP API

All of the other API client libraries use this Mashape HTTP API behind the scenes. This is where you should start if you want to implement your own client library.

## Python client library

Our official Python client library works with Python 2.7 and 3.4+, possibly others.

# Community

The following are provided by community members. Direct questions and feedback to their respective authors.

## Botson Chrome extension

Uses Botometer to filter your Twitter timeline. Contributed by [Andy Jiang](#).

## NodeJS client library

Contributed by [Keith Collins](#).

## R client library

Contributed by [Joey Marshall](#).

[Tweet at us](#) if you'd like us to link to your project here.

## Tools (beta)

The who, what, where, when, and how of social media data

### Hoaxy



Visualize the spread of claims and fact checking.

### Botometer



Check how bot-like a Twitter user behaves.

### Fakey



Play this game to learn to recognize fake news on your social feed.

### BotSlayer



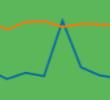
Detect bot-amplification on Twitter

### EchoDemo



Simulation demonstrating how two basic mechanisms of social media can lead to polarized social networks.

### Trends



Compare when memes gain and lose popularity.

### Networks



Explore who is discussing a meme and what memes are related.

### Movies



Generate movies of how

### Bot Electioneering Volume



Generate movies of how

### API



Query our data for your own

### Enhanced Data



Over 20B public tweets since

HOAXY<sup>beta</sup>

About Stats Data OSoMe

## HOAXY<sup>beta</sup>

Visualize the spread of claims and fact checking

three million votes illegal aliens

Relevant Recent

1 Search

2 Select claims and fact checking articles to visualize.

Select Top 20

3 Visualize

Tweet Share

Report: Three Million Votes in Presidential Election Cast by Illegal Aliens  
From infowars.com on Nov 14, 2016

17982 Tweets 52300 Facebook Shares

Did 3 million undocumented immigrants vote in 2016?  
From politifact.com on Nov 17, 2016

9415 Tweets 30920 Facebook Shares

Three Million Votes in Presidential Election Cast by 'Illegal Aliens'?  
From snopes.com on Nov 16, 2016

5699 Tweets 12638 Facebook Shares

FALSE: Obama Encouraged 'Illegal Aliens' to Vote  
From snopes.com on Nov 6, 2016

1911 Tweets 25327 Facebook Shares

UC Berkley President Janet Napolitano Played Key Role in Providing Sanctuary to Illegal Aliens  
From Breitbart.com on Feb 2, 2017

1574 Tweets 16910 Facebook Shares

Italy Announces 'Zero Tolerance' for Illegal Aliens  
From Breitbart.com on Dec 31, 2016

1060 Tweets 44483 Facebook Shares

Hillary Clinton Received Over 800,000 Illegal Votes, Research Claims  
From Breitbart.com on Jan 27, 2017

764 Tweets 20553 Facebook Shares

California Licenses 800K Illegal Aliens as Motor Voter Law Looms  
From Breitbart.com on Dec 30, 2016

744 Tweets 3726 Facebook Shares

NPR: 25 Million Votes For Clinton 'Completely Fake' – She Lost Popular Vote  
From yournewswire.com on Jan 29, 2017

687 Tweets 92438 Facebook Shares

# Bot Repository

Increasing evidence suggests that a growing amount of social media content is generated by autonomous entities known as social bots. Many social bots perform useful functions, but there is a growing record of malicious applications of social bots. We believe it is important to provide public datasets and tools that help identification of social bots, since deception and detection technologies are in an arms race.

**Bot repository** is a centralized place to share annotated datasets of Twitter social bots. We also provide list of available tools on bot detection.

## Datasets

Collection of annotated social bot datasets. Links to original source (webpage, publication, etc.) is provided in the details.

[View details »](#)

## Tools

Links to publicly available tools on social bot detection. Information on webpage, publication, etc. is provided in the details.

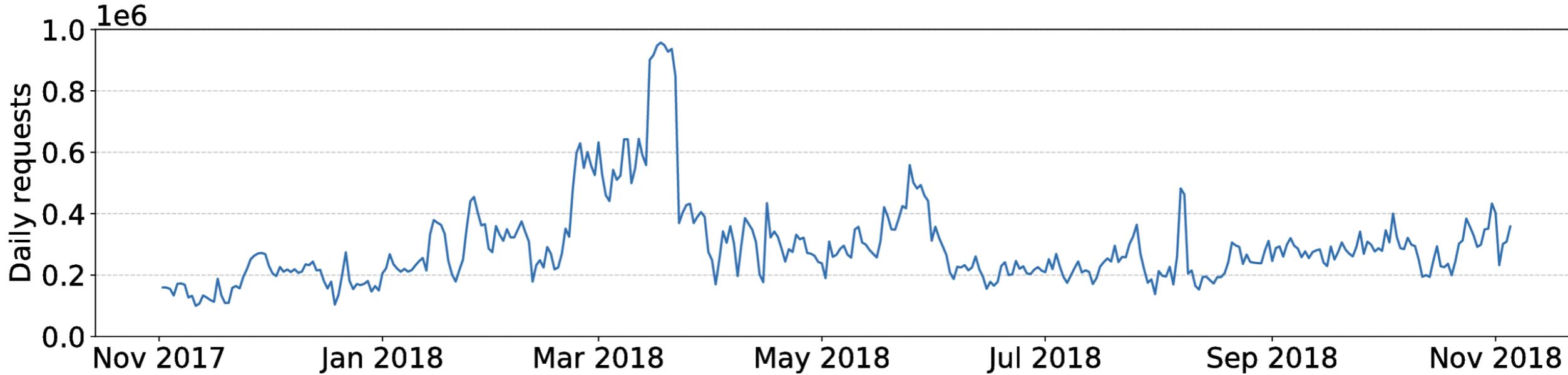
[View details »](#)

## Contribute

We always need more publicly available datasets and tools. Please provide information about your awesome dataset or tool. We would like to feature them in our website.

[View details »](#)

✉ Main contact for inquiries: [Onur Varol](#)



All of the other API client libraries use this interface to interact with the Botometer API behind the scenes. This is where you should start if you want to implement your own client library.

### Python client library

Our official Python client library works with Python 2.7 and 3.4+, possibly others.

## Community

The following are provided by community members. Direct questions and feedback to their respective authors.

### Botson Chrome extension

Uses Botometer to filter your Twitter timeline. Contributed by [Andy Jiang](#).

### NodeJS client library

Contributed by [Keith Collins](#).

### R client library

Contributed by [Joey Marshall](#).

[Tweet at us](#) if you'd like us to link to your project here.

## Tools (beta)

The who, what, where, when, and how of social media data

### Hoaxy



Visualize the spread of claims and fact checking.

### Botometer



Check how bot-like a Twitter user behaves.

### Fakey



Play this game to learn to recognize fake news on your social feed.

### BotSlayer



Detect bot-amplification on Twitter

### EchoDemo



Trends



### Networks



### Maps





Data Config About Help

Query: puppies,kittens

Show Only:

Filter results

Hashtags  User Mentions  Pics/Vids  Links

« < 1 2 3 4 ... > »



BS Level ↑	Entity	Research	Last Seen	Tweets	Accounts	Trendiness	Botness
██████	#middleeast	➡   ⚡   G	3 minutes ago	83	83	8300%	2.1 / 5
██████	#usdrone	➡   ⚡   G	3 minutes ago	82	82	8200%	2.1 / 5
██████	@carolyncorneli8	➡   ⚡   G	17 minutes ago	18	18	1800%	2.5 / 5
██████	@bluestatesos	➡   ⚡   G	17 minutes ago	20	20	2000%	2.5 / 5
██████	@tdlombard	➡   ⚡   G	14 minutes ago	20	20	950%	2.6 / 5

# When do we make mistakes?

Neither **human** nor **ML models** are perfect.

- Humans lack of computational power and memory
- ML algorithms can do certain jobs better but weaker in generalization

Possible causes for **false positives**: Very active users (fans etc.), connected apps, business professionals

Certain reasons for **false negatives**: Organizations, foreign language tweets

# Deep-fakes “Fake news on steroids”



Deep fakes: <https://www.youtube.com/watch?v=gLoI9hAX9dw>

# Protecting elections from social media manipulation

Sinan Aral<sup>1,2,3</sup>, Dean Eckles<sup>1,2</sup>

[+ See all authors and affiliations](#)

Science 30 Aug 2019:  
Vol. 365, Issue 6456, pp. 858-861  
DOI: 10.1126/science.aaw8243

## A 4-step plan for fighting social media manipulation in elections

- Catalogue exposures to manipulation: Asking W and H questions about content.
  - **What** are the content shared?
  - **Where** it was shared?
  - **How** it was disseminated?
  - **How much** advertisement was used?
  - **Who** paid for it?
- Combine exposure and voting behavior datasets: Matching mobility information with voter registration records.
- Assess the effectiveness of manipulative messages: Measuring exposure effects using A/B tests as companies have been doing for products.
- Calculate consequences of voting behavior changes: Modeling voter behavior using causal models and simulations.

NASA HAS CONFIRMED  
THAT THE ASTEROID IS  
HEADED DIRECTLY FOR US.

...YES, A QUESTION?



WHAT ROLE HAS SOCIAL  
MEDIA PLAYED IN THIS  
ASTEROID'S ORBIT?



\*SIGH\*



HAS TWITTER CHANGED THE WAY WE  
RESPOND TO ASTEROID THREATS?

WELL, IT'S MADE THE  
PRESS CONFERENCE  
QUESTIONS STUPIDER.



FASCINATING!

WHAT ABOUT  
FACEBOOK?

NASA HAS CONFIRMED  
THAT THE ASTEROID IS  
HEADED DIRECTLY FOR US.

...YES, A QUESTION?



WHAT ROLE HAS SOCIAL  
MEDIA PLAYED IN THIS  
ASTEROID'S ORBIT?



\*SIGH\*



HAS TWITTER CHANGED THE WAY WE  
RESPOND TO ASTEROID THREATS?

WELL, IT'S MADE THE  
PRESS CONFERENCE  
QUESTIONS STUPIDER.



**Social media can not solve problems directly but  
can be efficiently used for grassroots movements**



# Survey study

I propose to create a survey to be conducted during the HLF. Using the inputs from researchers and laureates, we can investigate attitudes toward these important subjects and how we can address them in the near future.

Survey will consist of basic background information and separate section capturing questions created by different study groups.

# Study groups

- ◆ Emerging technologies and detection techniques
- ◆ Integrity of elections over the world
- ◆ Ethics and platform responsibilities
- ◆ Crowd-sourced efforts for data collection
- ◆ Media literacy and pro-active awareness campaigns
- ◆ Brave ideas for the new world

# Study group interests

- [ 7 ]: Integrity of elections over the world
- [ 7 ]: Brave ideas for the new world
- [ 6 ]: Emerging technologies and detection techniques
- [ 5 ]: Ethics and platform responsibilities
- [ 4 ]: Crowd-sourced efforts for data collection
- [ 3 ]: Media literacy and pro-active awareness campaigns



# Study groups

Emerging technologies and detection techniques	Alshehri	Mohammad	Graduate PhD	Crowd-sourced efforts for data collection	Russo Russo	Gabriele	Graduate PhD
	Bharadwaj	Pooja	Undergraduate		Roitero	Kevin	Graduate PhD
	Dwivedi	Ashutosh Dhar	Graduate PhD		Nthetse	Oduetse	Undergraduate
	Galletta	Antonino	Graduate PhD		Akter	Mousumi	Graduate PhD
	Jain	Rahul	PostDoc		Desolda	Giuseppe	PostDoc
	Lahiri	Avisek	Graduate PhD		Li	Jie	PostDoc
	Shen	Zejiang	Undergraduate		Neeraj	Neeraj	PostDoc
Integrity of elections over the world	Neeraj	Neeraj	PostDoc	Media literacy and pro-active awareness campaigns	Thevis	Andrea Hannah	Graduate PhD
	Hermann	Alexander Benjamin Chr	PostDoc		Bharadwaj	Pooja	Undergraduate
	Hassan	Wajih UI	Graduate PhD		David	Marco	Undergraduate
	Hope	Tom	Graduate PhD		Jain	Lakshay	Undergraduate
	Randig	Marvin	Undergraduate		Moutamal	Maryse	Graduate PhD
	Sourav	Suman	Graduate PhD		Oshinubi	Kayode	Graduate PhD
	Wei	Yu-Lin	Graduate PhD		Pires Cardoso	Matheus	Undergraduate
	Yfantidou	Sofia	Graduate PhD		Rao	Vanishree	PostDoc
	Ahmed	Chuadhry Mujeeb	Graduate PhD				
Ethics and platform responsibilities	Nguyen Quang Do	Lisa	Graduate PhD	Brave ideas for the new world	Jindal	Anish	PostDoc
	Gray	Charles	Graduate PhD		Kerriou	Celine	Graduate PhD
	López Valdez	Hiram	PostDoc		Nwokoye	Chukwunonso	Graduate PhD
	Martinuzzi	Emanuele	Undergraduate		Salami	Behzad	PostDoc
	Mertz	Ian	Graduate PhD		Sharma	Dravyansh	Graduate PhD
	Porfirio	David	Graduate PhD		Wang	Jingbo	Graduate PhD
	Zaikovskii	Anatolii	Graduate PhD		Wu	Xinyu	Undergraduate