

Cyber Frauds: Phishing, Astroturfing, Fake News, and Deepfake

Dongwon Lee

Penn State University, USA

`dongwon@psu.edu`



Oct. 24, 2019 @ ORAU Fraud Informatics Symposium

Fraud Informatics (FI) Project

- NSF SaTC EDU Grant (2018 – 2021)
- Joint effort between Penn State and ORAU
- To develop and evaluate materials to teach **modern** types of cyber frauds to diverse audience



The screenshot shows the National Science Foundation (NSF) website. At the top left is the NSF logo with the text "National Science Foundation" and "WHERE DISCOVERIES BEGIN". To the right is a search bar with the word "SEARCH" and a magnifying glass icon. Below the header is a dark blue navigation bar with white text for "HOME", "RESEARCH AREAS", "FUNDING", "AWARDS", "DOCUMENT LIBRARY", "NEWS", and "ABOUT NSF". The "HOME" link is highlighted in a light blue box. Below the navigation bar, the "Awards" section is visible, featuring a dark background with the word "Awards" in white. To the right of this section, there is a blue envelope icon, the text "Award Abstract #1820609", and the title "Developing and Evaluating Fraud Informatics Curriculum among Institutions in the Appalachian Region" in bold black text, which is underlined with a green line.

Objectives

- Cover latest modern types of cyber frauds
- Cover latest research on the prevention and detection of cyber frauds
 - AI methods
 - Data-driven
 - Information-processing
- Develop media-rich hands-on materials
 - Images and videos
 - Hands-on labs using games and tools

Formats of Delivery

1. 1-2 hour-long

- Fraud informatics “hygiene”
- K12 students or general audience

2. 2-3 week long

- Special topic plug-in to other related classes
- CompSci undergraduates

3. Semester-long

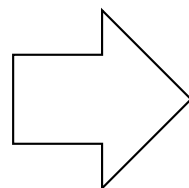
- Dedicated class on Fraud Informatics
- CompSci undergraduates

What is “Fraud”?

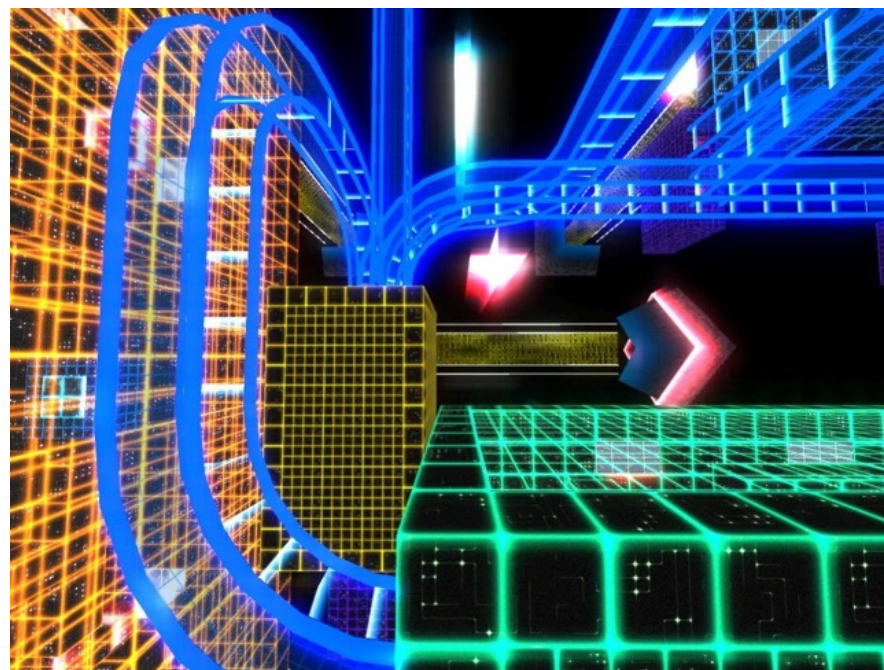
- Oxford dictionary
 - *“wrongful or criminal deception intended to result in financial or personal **gain**”*
- Van Vlasselaer et al. (2015)
 - *“Fraud is an **uncommon**, well-considered, **imperceptibly concealed**, **time-evolving**, and often **carefully organized** crime which appears in **many types of forms**”*
 - 5 characteristics

“Traditional” (Consumer) Frauds

- Credit card fraud
- Insurance fraud
- Product warranty fraud
- Healthcare fraud
- Money laundering
- Identity theft
- Telecommunications fraud



Cyberspace



“Modern” Frauds in Cyberspace

- Spam/Phishing, and Social Engineering Fraud
- Fake News
- Deepfake
- Astroturfing and Crowdturfing
- Sockpuppet and Catfish
- Academic Fraud
- ...

Other Important
modern cyber frauds?

Fraud Informatics

- Modern frauds need to be solved and taught in multiple disciplines and subjects
 - Computer Science (and AI)
 - Cognitive Science
 - Business
 - Criminology
 - Law
 - Policy ...

Avoid topics from traditional classes on Network, Systems, IoT securities



1. PHISHING

Terms

- Spamming: Unsolicited email/letter/SMS/...
- Social Engineering Attack: **Psychological** manipulation of victims for deception
 - Phishing = “ph” + fishing
 - Vishing = Voice Phishing
 - Spear Phishing
 - Whaling
 - ...

Targeted
Personalized
Human-written
Small-scale



Higher success rate
for attackers

Psychological Aspect

- Experiment in West Point, 2004
- Researchers sent a phishing email to 512 cadets, pretending it to be coming from a fictitious Colonel, asking them to click a malicious link regarding a grade change problem
- 80% of cadets clicked the link
- WHY so high?

Phishing Email

PayPal

We need your help

Your account has been suspended, as an error was detected in your informations.
The reason for the error is not certain, but for security reasons, we have suspended your account temporarily

We need you to update your informations for further use of your PayPal account.

[Update your information](#)

You are currently made disabled of :



Adding a payment method
Adding a billing address

Sending payment
Accepting payment

Please do not reply to this email. We are unable to respond to inquiries sent to this address. For immediate answers to your questions, visit our Help Center by clicking "Help" located on any PayPal page or email.

Copyright © 2016 PayPal, Inc. All rights reserved. PayPal is located at 2211 N. First St., San Jose, CA 95131.

Phishing Email



Refund Notification

Due to a sytem error you were double charged for your last order, A refund process was initiated but could not be completed due to errors in your billing information

REF CODE:2550CGE

You are required to provide us a valid billing address

[Click Here to Update Your Address](#)

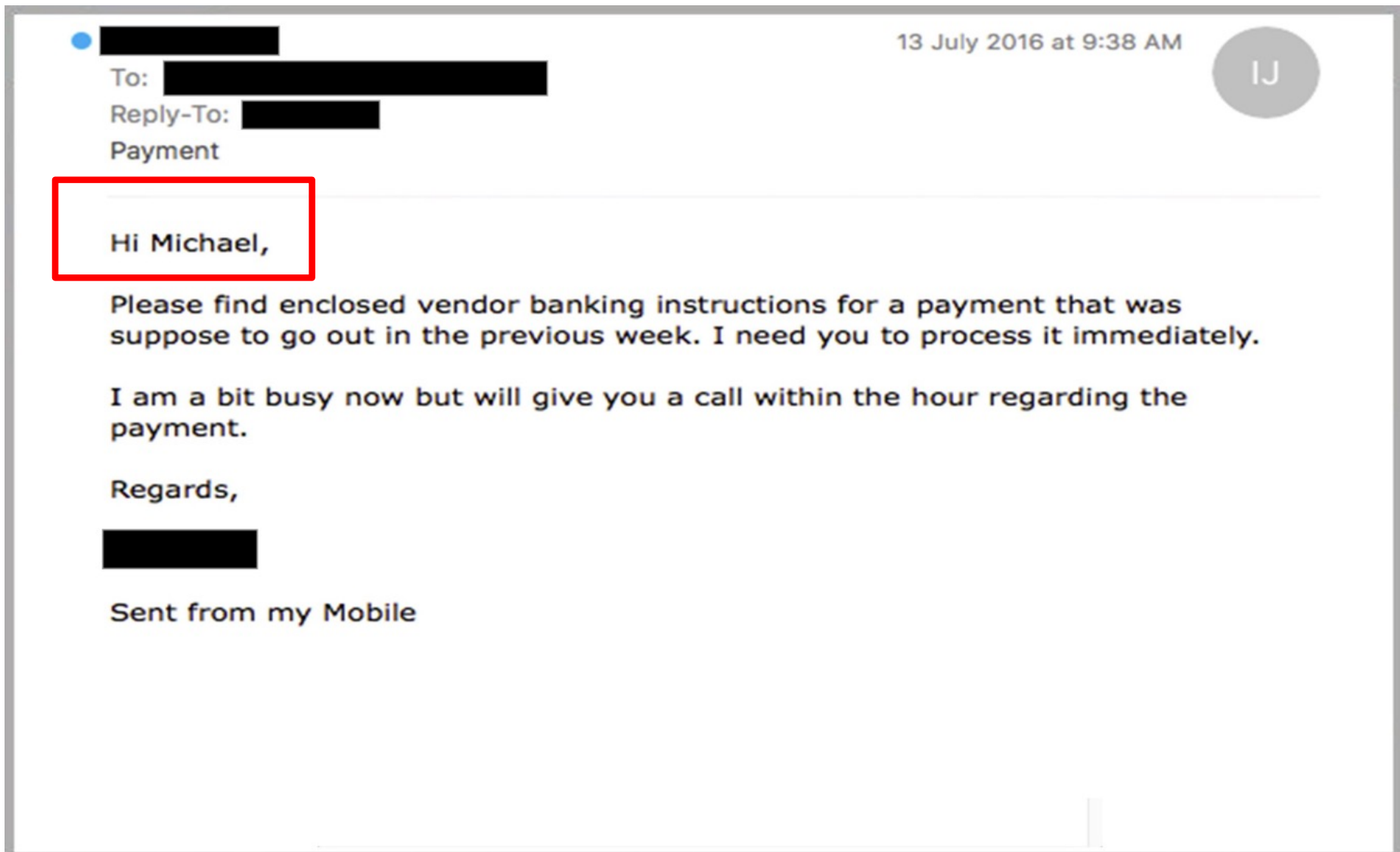
After your information has been validated you should get your refund within 3 business days

We hope to see you again soon.

[Amazon.com](https://www.amazon.com)

Email ID: 

Spear Phishing Email



Spear Phishing Email

The screenshot displays an Outlook window titled "Urgent: please update your payroll info - Message (HTML)". The interface includes a ribbon with "FILE" and "MESSAGE" tabs, and a left-hand navigation pane. The email content is as follows:

From: Julie Monical <jmonical@techmd.com> (highlighted with a red box and an arrow)
Date: Fri 4/29/2016 1:18 PM
Subject: Urgent: please update your payroll info
To: Drew Lawson

Body:

Hello,


You have an very important email from Human Resources Department with regards to updating your info for this month's payroll. The email is enclosed on the secure server, please access it usng the link below.

[Access the documents here.](#)

Helpfully,

Julie

Signature Block (highlighted with a red box):

	Julie Monical Director of Human Resources and Administration	Office: 888.883.2463 Direct: 657.229.7699 www.TechMD.com
---	--	--

At the bottom of the window, the status bar shows "ITEMS: 1,066 UNREAD: 1" and "CONNECTED TO: MICROSOFT EXCHANGE".

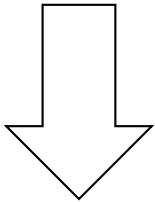
Vishing



**WATCH THIS HACKER
BREAK INTO
MY CELL PHONE ACCOUNT
IN 2 MINUTES**

<https://www.youtube.com/watch?v=BEHI2IAuWCk>

Personalized Attack

- How do attackers get information about victims?
- Scavenger-hunting, Hacking
 - 
- Data-driven guessing
 - Eg, by analyzing one's social media data, AI can accurately predict diverse demographics of users

You Are What You LIKE

- *Hypothesis:* The LIKE pattern in social media is correlated with one's personal traits

from Diy & Crafts Ideas Magazine
Ten genius storage ideas for the bathroom 7
Kohls Home Decor | My bathroom remodel. Love it!!! Kohls towels Kohls shower curtain Home ...
13827 1741 11

from The SITS Girls
DIY Menu Board
This cute menu board is an easy DIY project that will help you get organized for meal planning and add a touch of fun to your home decor. A double win! It could also be used for organizing your weekly schedule or family appointments.
59055 5736 92

Love Bakes Good Cakes
!!Freedom Fridays (LINKS ...

More cheat sheets! | CheekyLime | Stylish Camera Bags | Designer DSLR Bag For Women
from cheekylime.com
4472 586 6

Candice Stears Photography

GETTING STARTED WITH YOUR DSLR:
How to get out of auto mode and start taking better photos!

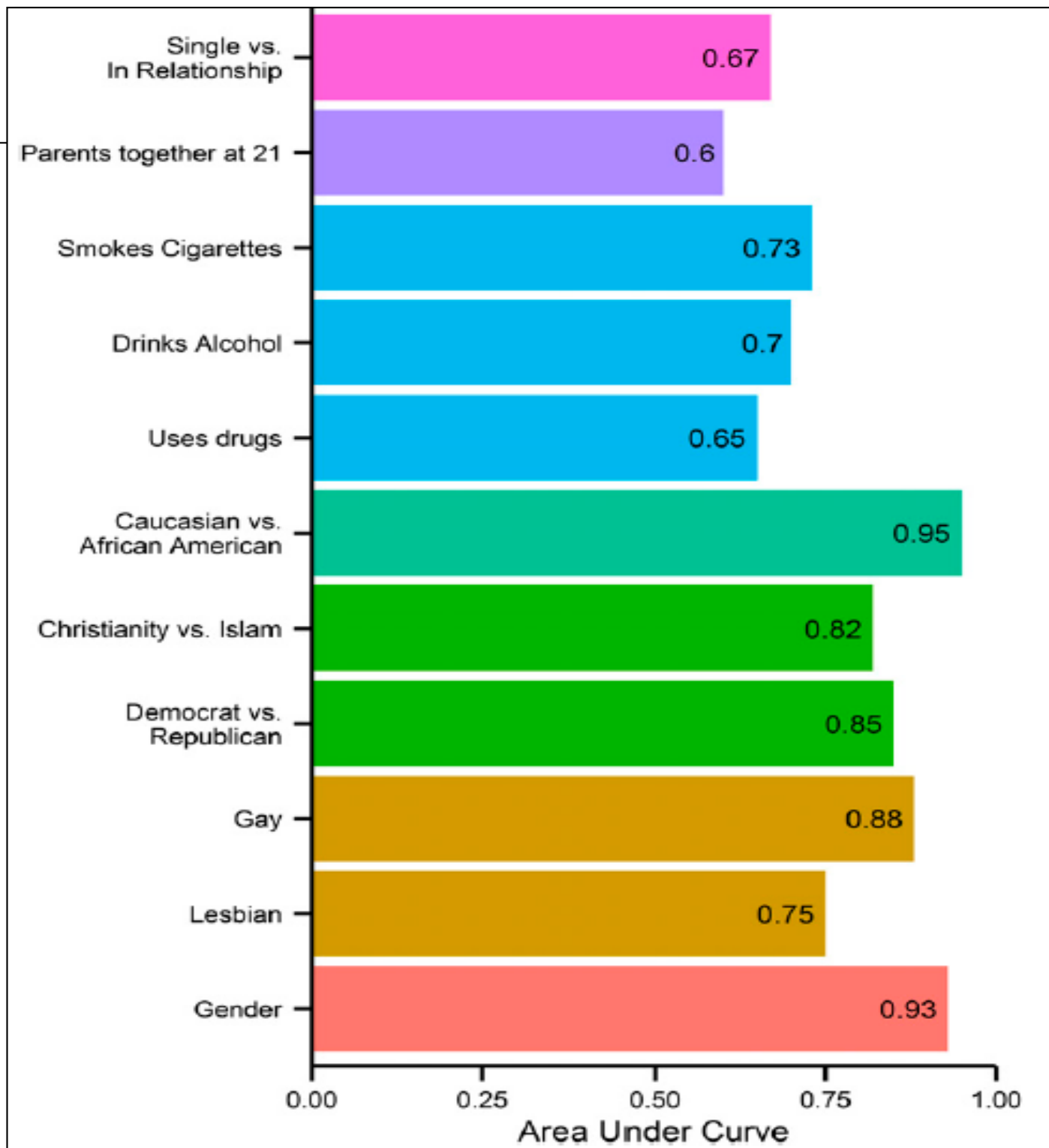
from Frugal Fun For Boys
A Beginner's Guide to Your DSLR Camera: How to Get out of Auto Mode and Start...
A beginner's guide to taking better photos with a DSLR. How to stop using the flash, get out of auto mode, and start taking better pictures.

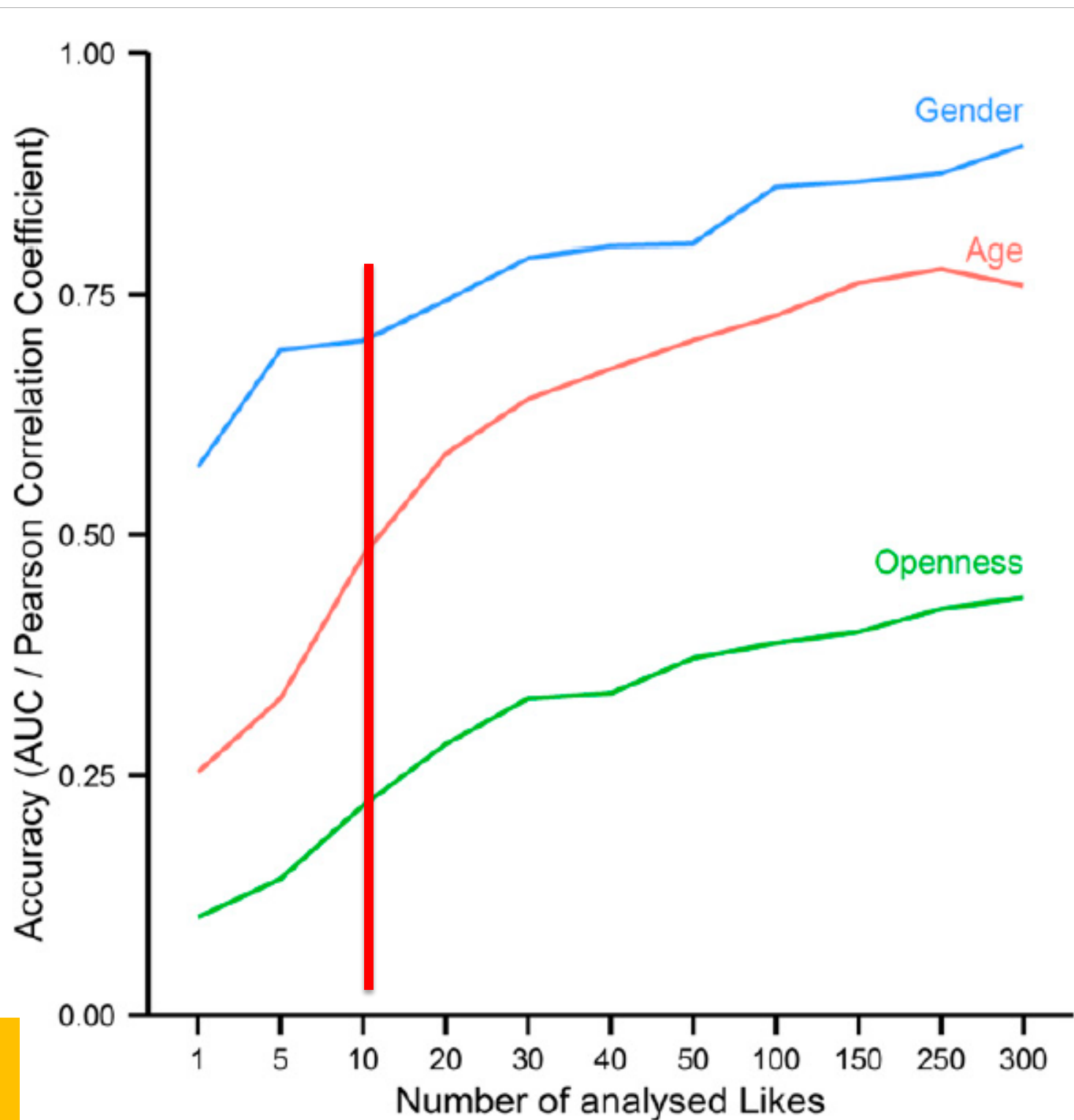
from Sweet C's Designs
Know your DSLR Lenses
dslr lenses- get to know all about telephoto, prime lenses, wide angle and kit lenses! Which to use for faces, which to use for places, and everything in between. (great info for beginner photographers)
#photography #DSLR #camera
1568 205 1

allison camille
Good to Know

Manual Modes of DSLR

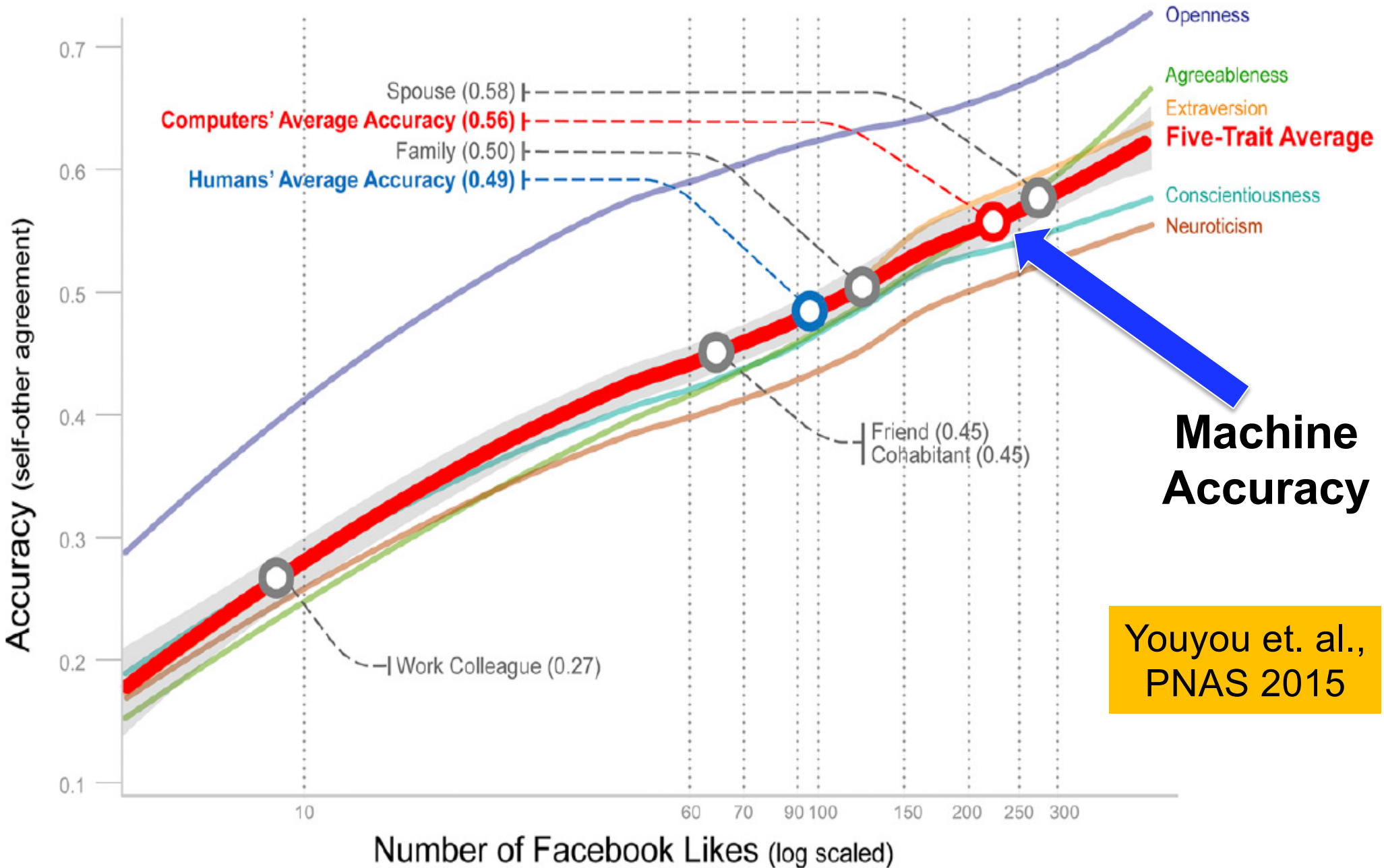
super beginner photography tips





Kosinski et. al.,
PNAS 2013

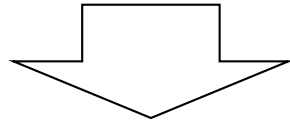
Personality Prediction



Youyou et. al.,
PNAS 2015

Scenario

- From LIKE data, an attacker predicted a victim to be:
 - An African American Christian female in her 20s living in NYC...



- More personalized spear phishing email can be written

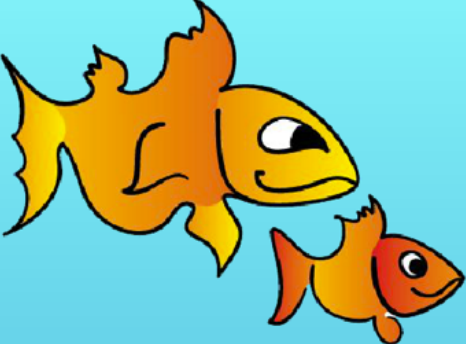
Dear **Ms.** Jane Doe, pardon for this interruption.

I am a **pastor** living in **Queens** ...

How to Spot Phishing Emails?

- Discussion


Lab: Domain Highlighting



ANTI-PHISHING PHIL

Anti-Phishing Phil was developed by the CMU Usable Privacy and Security Laboratory (CUPS) at Carnegie Mellon University with funding from the US National Science Foundation and ARO/CyLab. Anti-Phishing Phil is distributed under a Creative Commons Attribution-Noncommercial license. For information about commercial licensing, including opportunities to have the game customized for your business or training program, please see: [http:// cups.cs.cmu.edu/antiphishing_phil/](http://cups.cs.cmu.edu/antiphishing_phil/)

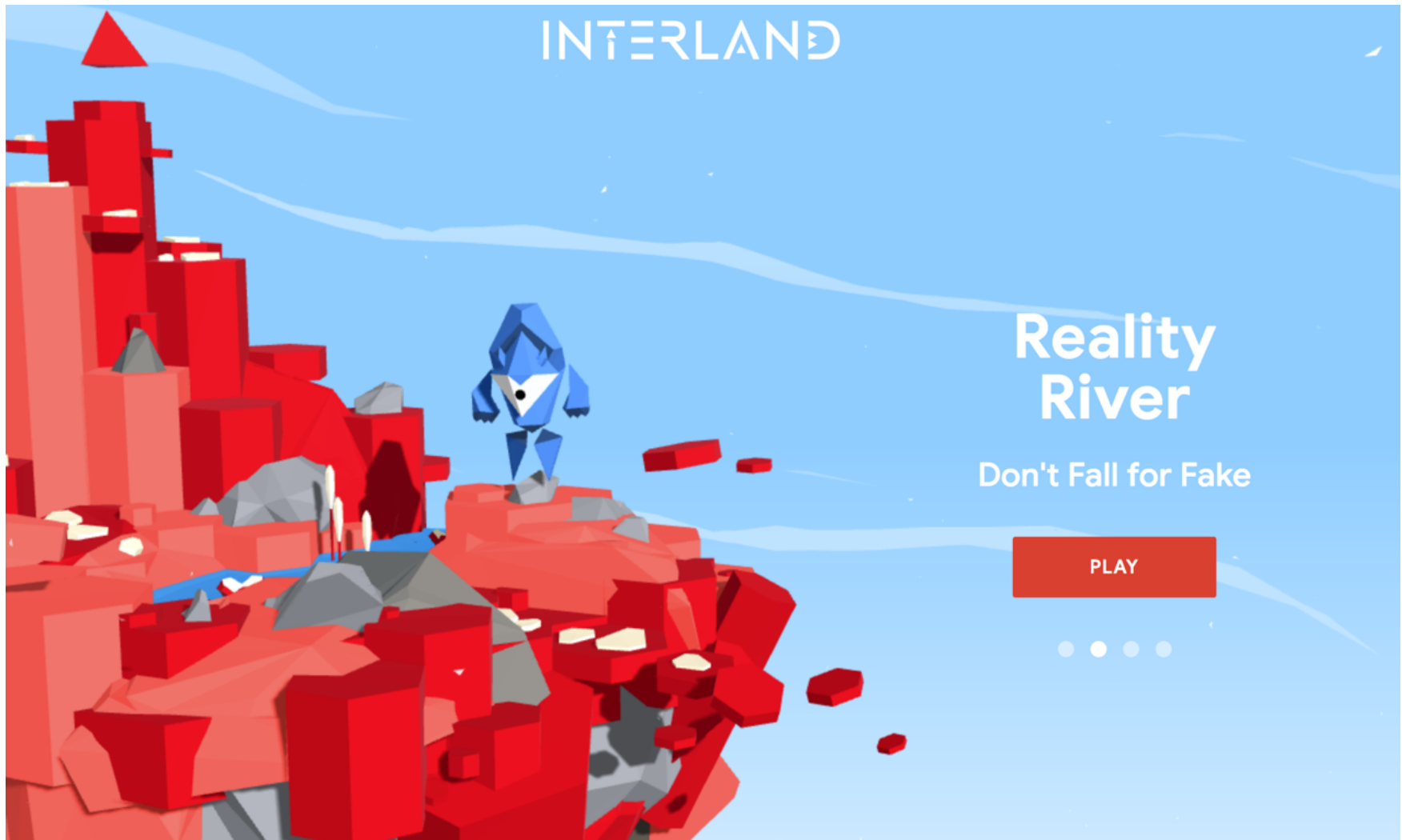
At the end of each round, Anti-Phishing Phil may send anonymous scoring information back to us so that we can improve the game and gain insights into how to better protect people from phishing attacks.



CMU Usable Privacy
and Security Laboratory

PLAY

Lab: Phishing



<https://beinternetawesome.withgoogle.com/en/interland/landing/reality-river>

Attack-Back #1

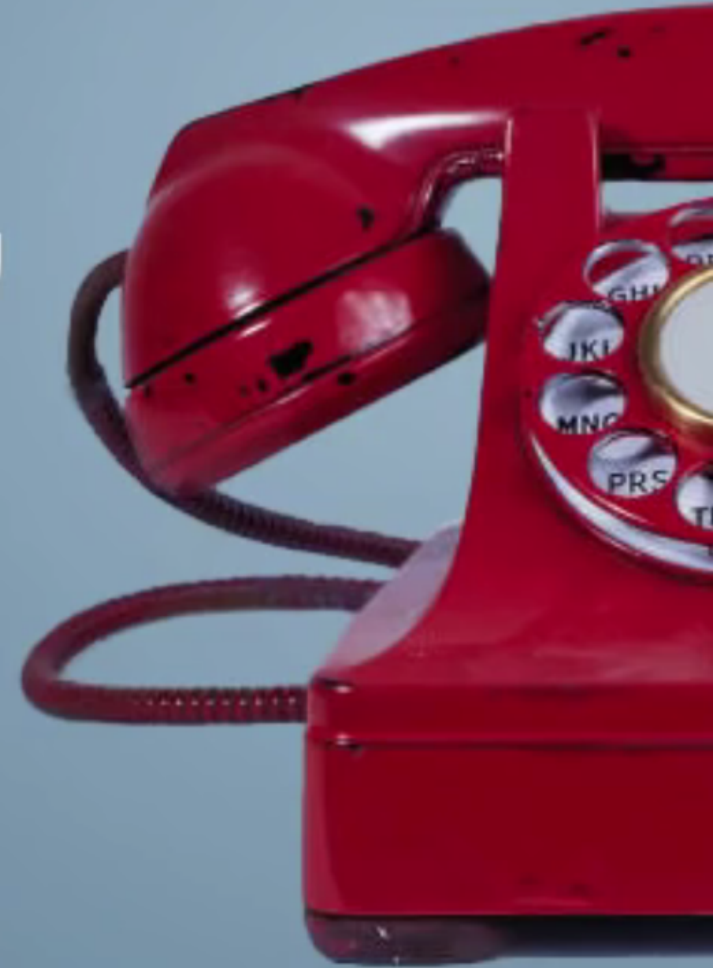


https://www.youtube.com/watch?v=_QdPW8JrYzQ

Attack-Back #2

“HELLO, THIS IS
Lenny!”

“Lenny” is a collection of recorded messages, designed to fool telemarketers into thinking they are talking to a real person.



<https://www.youtube.com/watch?v=t7kSWvt3KXY>

how to get fake
likes in
fb..



2. ASTROTURFING

Definition

- *Astroturf*: fake grass(roots)
- Examples
 - Fake LIKES in facebook
 - Orchestrated fake reviews in amazon.com

Power of LIKE



LIKE Us or Get Out !

INQUISITR
NEWS WORTH SHARING

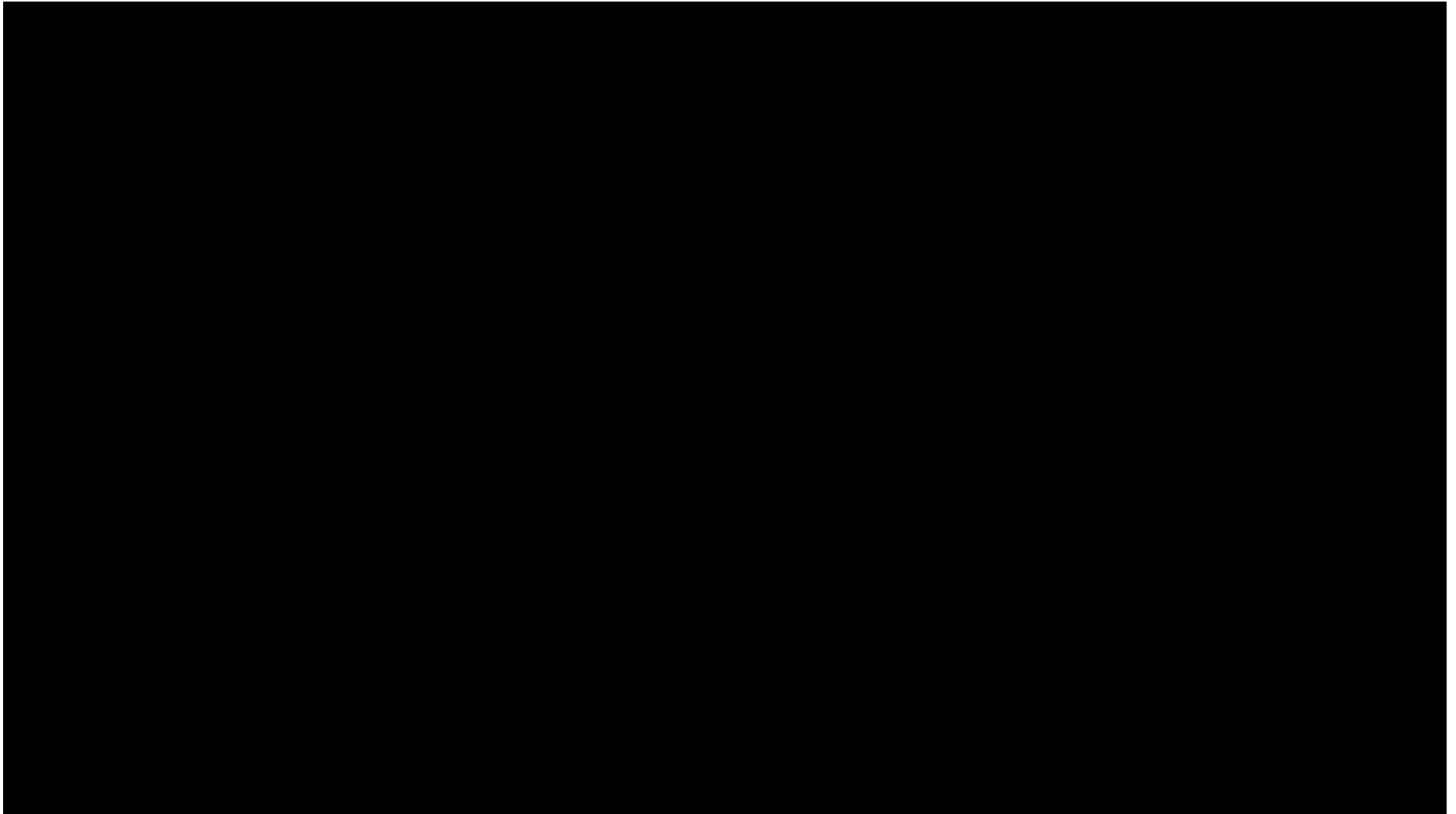
ENTERTAINMENT NEWS & POLITICS SPORTS LIFESTYLE HEALTH SCIEN

MAY 29, 2016

**SALT LAKE CITY APARTMENT COMPLEX
REQUIRES RESIDENTS TO LIKE THEIR
FACEBOOK PAGE, ANGERS THE INTERNET
WITH PREDICTABLE RESULTS**

DON CROTHERS

PBS Frontline, 2014



<http://www.pbs.org/wgbh/pages/frontline/generation-like/>

Fake LIKES

- People buy and sell *Likes*
- Huge commercial implications
- Headache for SNS to maintain healthy ecosystem

facebook business

CREATE PAGE CREATE AD ADS HELP SUPPORT

Marketing on Facebook Pages Ads Success Stories Learn How News

Search



News > Breaking new Ground in the Fight Against Fake Likes

April 17, 2015



By: Huseyin Kerem
Cevahir

Announcements



Breaking new Ground in the Fight Against Fake Likes

Keeping activity authentic on Facebook is an important part of our ongoing efforts to make sure Facebook is a trustworthy place for people and businesses to connect. Last October we shared some of the ways we work to [stop fraudulent activity on Facebook](#), particularly around fighting fake likes. We've made big strides in the past six months, so today we're providing an update on our progress in this area.



Overview Description Basic Gig Offer Gig Extras Reviews

Favorite

I will give you 600 real and permanent facebook page like with 2 split

★★★★★ (1) 1 Orders in Queue

Digital Marketing / Social Media Marketing

100% Safe & Reliable

600+

REAL FACEBOOK LIKES

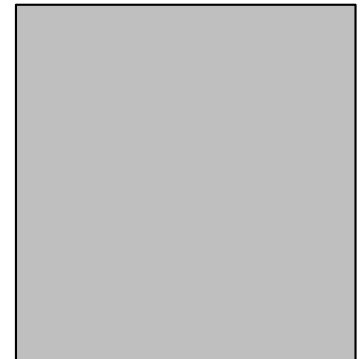
Fanpage/Post Likes
24 hours delivery



Order Now (\$10)

Basic Gig Quantity 1 (\$5)

⌚ Extra Fast: 1 Day Delivery 1 (\$5)



Level 2 Seller

From

Positive Rating



[HOW IT WORKS?](#)

[LOGIN](#)

[REGISTER](#)

[BUY POINTS](#)

[BUY SOCIAL MEDIA](#)

GET HIGH QUALITY LIKES FROM THE WHOLE PLANET!

LikesPlanet is the best way to advertise your social pages.



131,831,980 Hits/likes delivered

367,956 Members and counting!

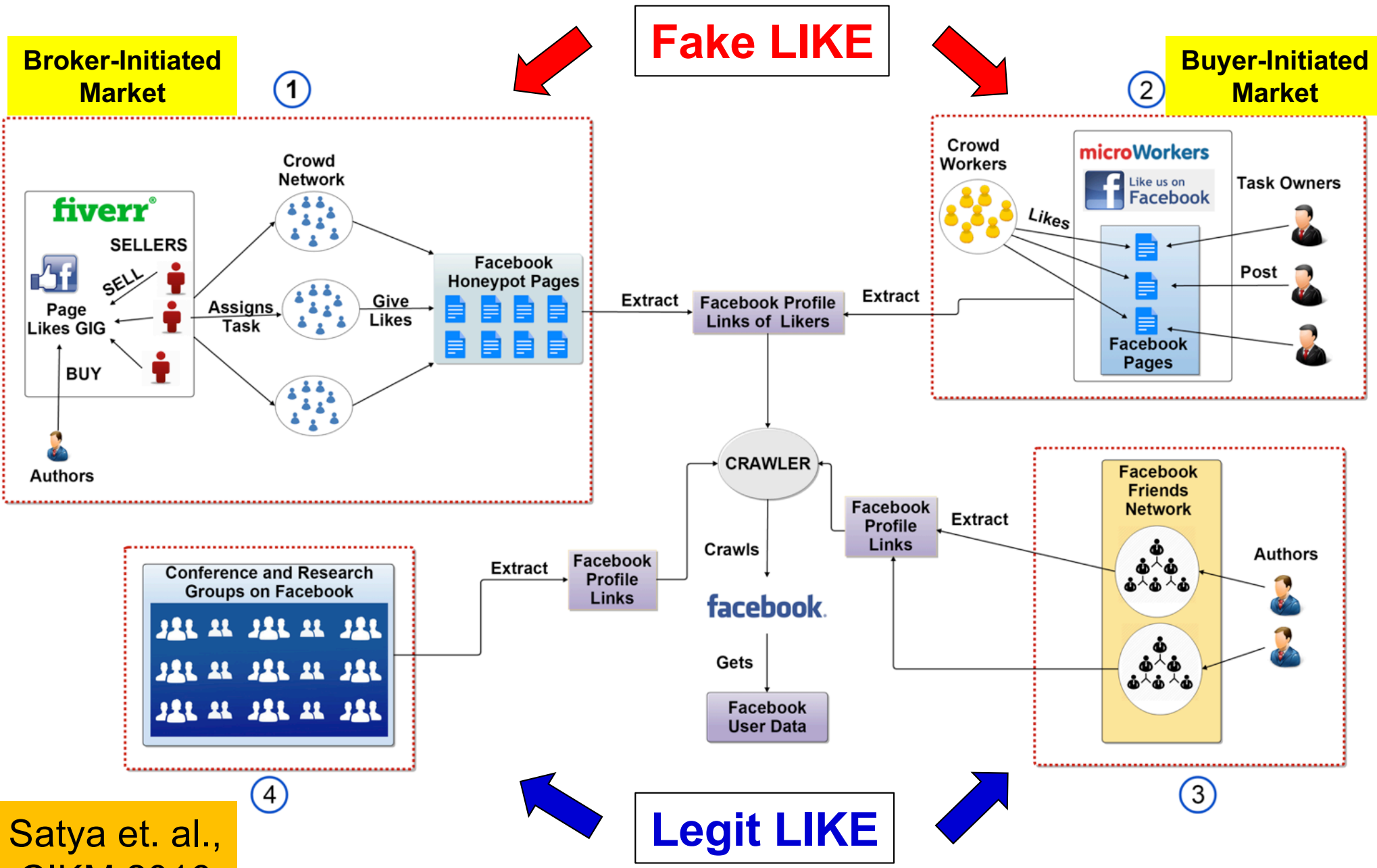


245 Members joined in the last 24 hours!



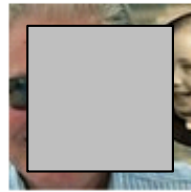
1,062 Online!

Training Data for Machine Learning



Satya et. al.,
CIKM 2016

People who like Risk Board Game



زهير الجواهري
Likes Risk Board Game
Listens to Rana Samaha - رنا سماحه

+ Add Friend



Aditya Singh
Likes Risk Board Game

+ Add Friend



Haneen Uw
Likes Risk Board Game
لن ينساها لأنها استثناء
Listens to Osama Elhady

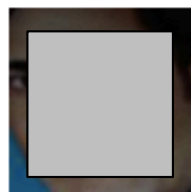


Wahy Zekryat
Likes Risk Board Game, NeverKnew and 21 others
Read القرآن هدى للناس
Listens to Amani Swissi



Nesso Mahlawy
Narmer Language School
Likes Risk Board Game, Clash of Queens and 5 others
Went to Narmer Language School

+ Add Friend



عادل شوقي
في مدينة الحب الحزين at اعمال حرة/ التجارة
Likes Risk Board Game, 仙姬情 and 3 others
Studied Civil Engineering at ثانوي صناعي

+ Add Friend



زياد ديسان
Amman, Jordan

+ Add Friend

20+ Find Friends

Board Game

Search posts on this Page

535 people like this

Invite friends to like this Page

ABOUT

This is not a real page, please do not like this page.

Ask for Risk Board Game's website

VISITOR POSTS

NEWS

Home | Video | World | US & Canada | UK | Business | Tech | Science | Magazine | Entertainment & Arts | Health | More

ADVERTISEMENT

**AND WHITE CLOTHES
FROM YELLOW STAINS**



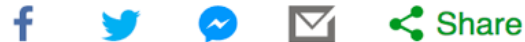
Degree
DEODORANT

Technology

<http://www.bbc.com/news/technology-22166606>

Samsung probed in Taiwan over 'fake web reviews'

16 April 2013 | Technology



Fair-trade officials in Taiwan are looking into reports that Samsung paid people to criticise rival HTC online.

Samsung is alleged to have hired students to post negative comments about phones made by Taiwan's HTC.

Samsung, based in South Korea, said the "unfortunate incident" had gone against the company's "fundamental



Bad reviews of HTC products were allegedly posted by students paid by Samsung

Top Stories

Pence leaves NFL game over protests

The US vice-president abandons a in his home state after players kneel during the national anthem.

3 hours ago

Huge Spain unity rally held in Barcelona

3 hours ago

Trump trades insults with top Republican

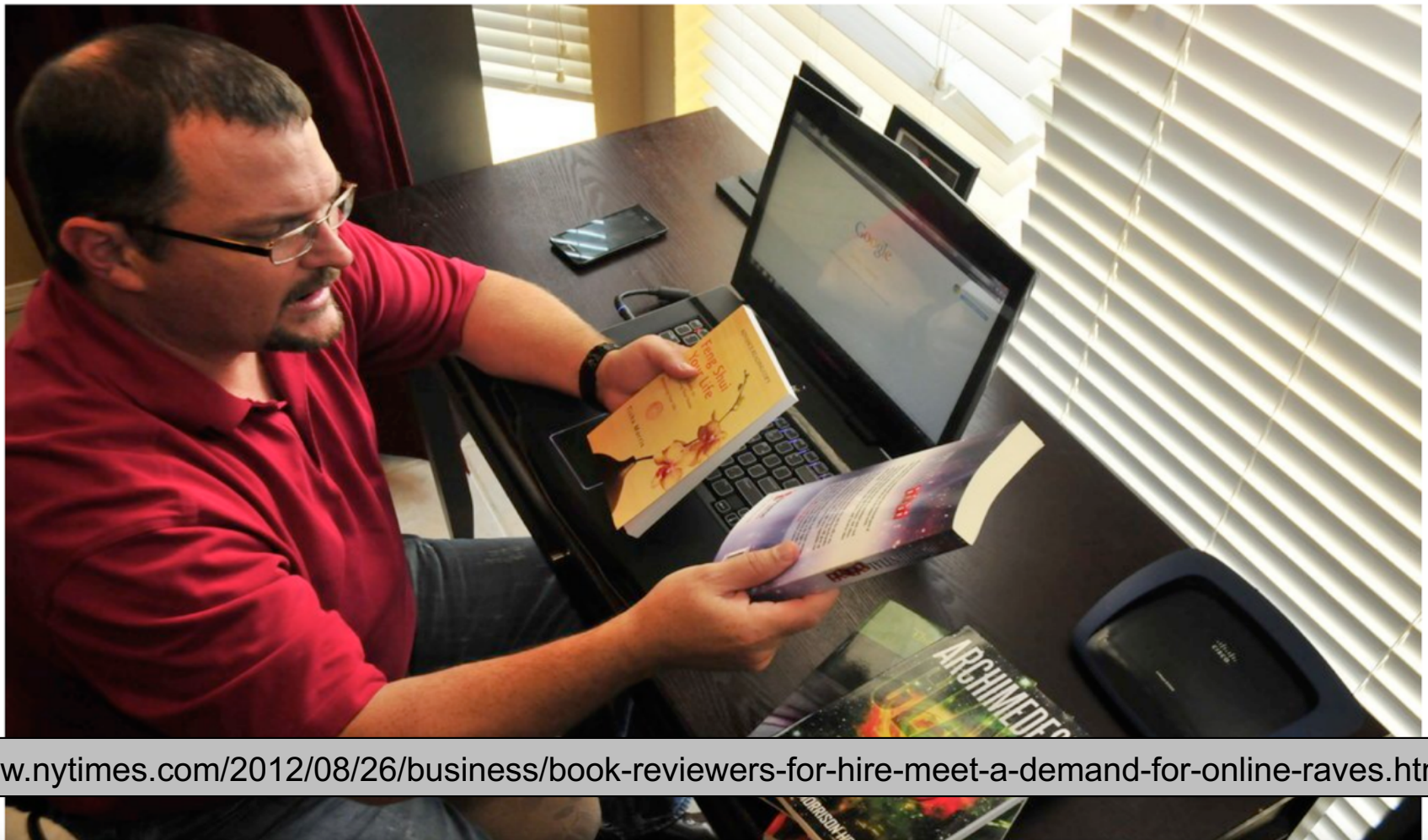
5 hours ago

ADVEF

BUSINESS DAY

The Best Book Reviews Money Can Buy

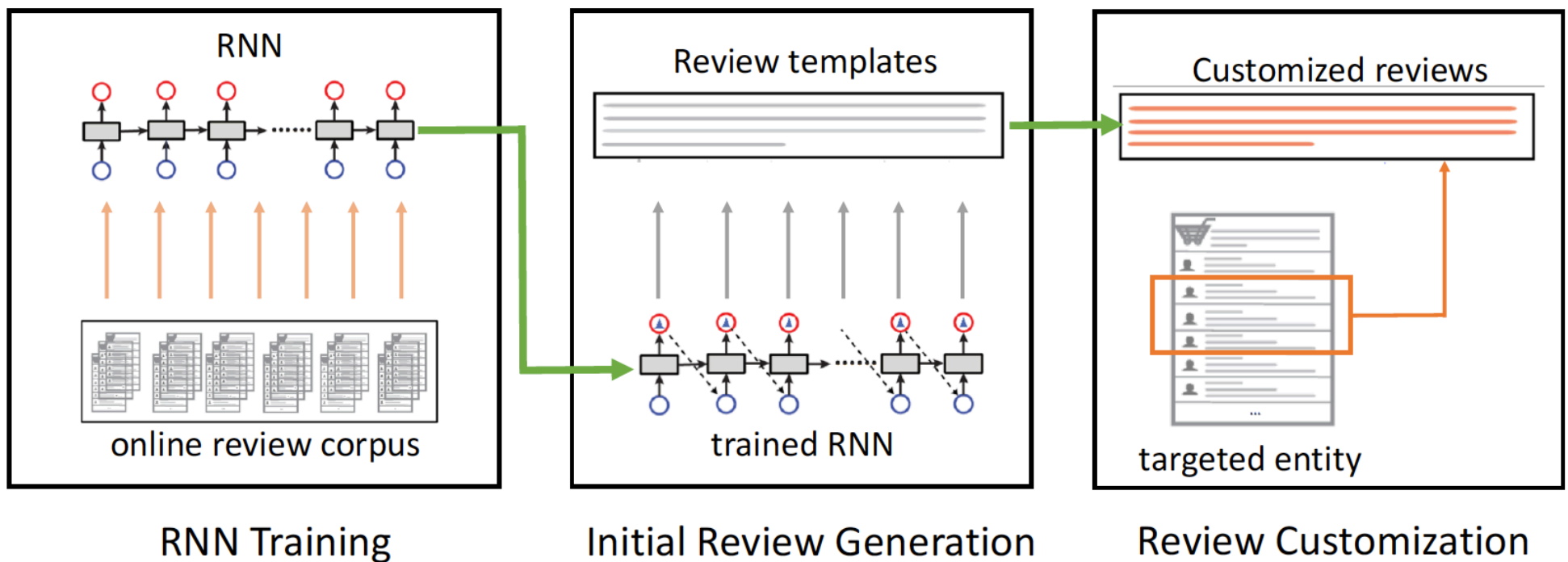
By DAVID STREITFELD AUG. 25, 2012



<http://www.nytimes.com/2012/08/26/business/book-reviewers-for-hire-meet-a-demand-for-online-raves.html>

Synthesized Amazon Reviews

Overview of Attack



Credit: Ben Zhao @ U. Chicago

Generated with *temperature* 0.1

I have been going to this place for a few years now and I have never had a bad experience. The service is great! They are always so friendly and helpful. I will definitely be back and I will be back for sure!



Generated with *temperature* 0.3

This place is amazing! The bartenders are absolutely amazing. The pasta is delicious and I love their pastries and it is amazing. I love the breakfast, friendly staff and the price is very reasonable. I have never had a bad experience here. I will be back for sure!

Generated with *temperature* 0.5

I love this place. I went with my brother and we had the vegetarian pasta and it was delicious. The beer was good and the service was amazing. I would definitely recommend this place to anyone looking for a great place to go for a great breakfast and a small spot with a great deal.

Generated with *temperature* 0.7

I have been a customer for about a year and a half and I have nothing but great things to say about this place. I always get the pizza, but the Italian beef was also good and I was impressed. The service was outstanding. The best service I have ever had. Highly recommended.

Generated with *temperature* 1.0

The food here is freaking amazing, the portions are giant. The cheese bagel was cooked to perfection and well prepared, fresh & delicious! The service was fast. Our favorite spot for sure! We will be back!

LAB


- Using FakeSpot (<https://www.fakespot.com/>), try a few Yelp restaurant reviews
 - Any restaurants with B or lower grade?
 - Understand the analysis of low grade
- Using ReviewMeta (<https://reviewmeta.com/>), try a few Amazon product reviews
 - Any product with FAIL rating?
 - Understand the analysis of FAIL rating

 **The Halloween Store** Shop candy, costumes, and more

Electronics > Computers & Accessories > Computer Accessories & Peripherals > Keyboards, Mice & Accessories > Mice

Purchased 1 time.
 You purchased this item on October 23, 2019.
 Color: Blue | [View this order](#)





Copy & Paste an Amazon Product Page URL


[RUN REPORT](#)

CATEGORIES
PHONE APPS
BROWSER EXTENSIONS
FAQ'S
HOW IT WORKS
BLOG
FORUM

[Amazon.com](#) > [Electronics](#) > [Computers & Accessories](#) > [Computer Accessories & Peripherals](#) > [Keyboards, Mice & Accessories](#) > [Mice](#)

Jelly Comb 2.4G Slim Wireless Mouse with Nano Receiver [More product info](#)

From [Jelly Comb](#)



See Price & Product Info


amazon

Rating:

4.4

★★★★★

9,095 Reviews



4% of potentially unnatural reviews removed ([Why?](#))

[View/Edit Adjustment](#)

[Jump to Report Card](#)

Adjusted

Rating:

4.4

★★★★★

8,696 Reviews

See Product Details

FAIL

Suspicious Reviewers

Take-Back Reviewers



2,086 of the 9,103 reviewers have had at least one of their past reviews for another product deleted. The Take-Back Reviewers have rated this product an average of 4.5 while the reviewers who don't have any deleted reviews in their history have rated this product an average of 4.3. Based on our statistical modeling, the discrepancy in average rating between these two groups is significant enough to believe that the difference is not due to random chance, and may indicate that there are unnatural reviews.



From Take-back Reviewers



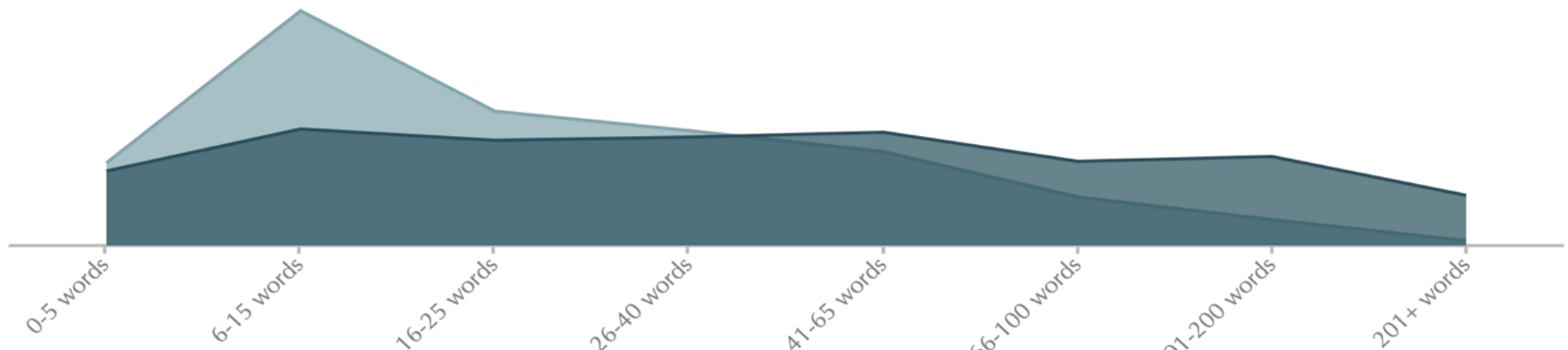
From Reviewers Without Previously Deleted Reviews



[Read more about our Suspicious Reviewers test.](#)

FAIL

Word Count Comparison





3. FAKE NEWS

False Information

Definitions of False Information

	Authenticity	Intention	News?
Fake news	False	Bad	Yes
False news	False	Unknown	Yes
Satire news	Unknown	Not bad	Yes
Disinformation	False	Bad	Unknown
Misinformation	False	Unknown	Unknown
Rumor	Unknown	Unknown	Unknown

Types of False Information

Snopes

- NEW What's New
- Hot 50
- Fact Check
- News
- About
- FAQ
- Contact
- Random



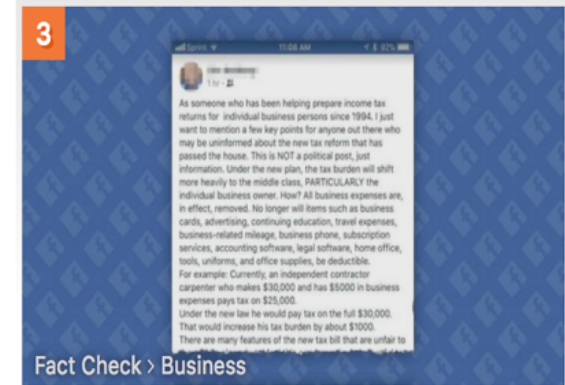
Fact Check > Politics

Did an Alabama Church Display a Sign Comparing Roy Moore to Jesus?



Fact Check > Fake News

Did Tokyo Open the First Human Meat Restaurant?



Fact Check > Business

Does the 2017 Tax Reform Bill Eliminate Small Business Deductions?

Real News

Commentary / Feature Writing

Misreporting

Native Advertisement

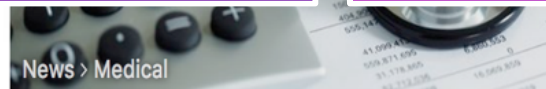
Professional Political Content

Citizen Journalism

Satire / Clickbaits

Polarizing and Sensationalist Content

Fake News / Hoaxes



News > Medical

Does the 2017 Republican Tax Bill Cut Cancer Treatments for Medicare Patients?

4 December 2017 - The Tax Cuts and Jobs Act does not mention the health care program, but critics say cancer patients would be hit hard if the bill passes.



Fact Check > Science

Massive Indian Ocean Earthquake and Tsunami Expected by the End of 2017?

27 September 2017 - An Indian man's prediction is based on his purported psychic abilities and carries no scientific weight.



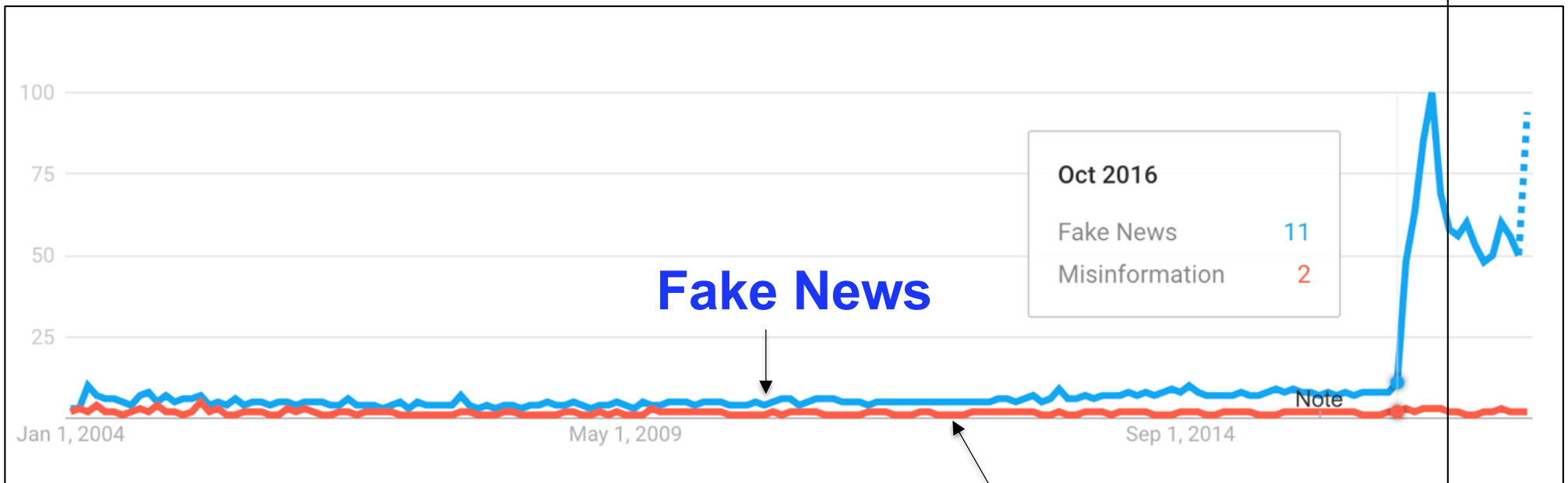
Fact Check > Fake News

Did a CIA Agent Confess to Killing Bob Marley?

4 December 2017 - A disreputable web site capitalized on an old conspiracy theory about the death of the reggae icon to spin fake news clickbait.

Surge of “Fake News”: Google Trend

US Election @ Nov. 2016



Fake News

Misinformation

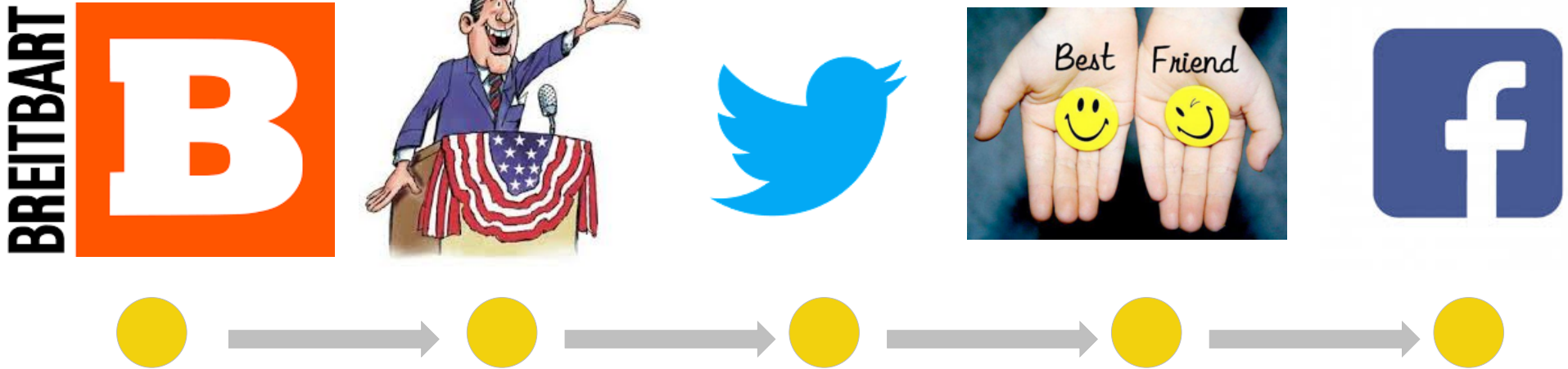
More Problems in Social Media?

1. Fundamental shift in communication:
Consumer as producer
2. Monetary incentives: Ads by Google/Facebook



More Problems in Social Media?

3. Source Layering



More Problems in Social Media?

4. Virality

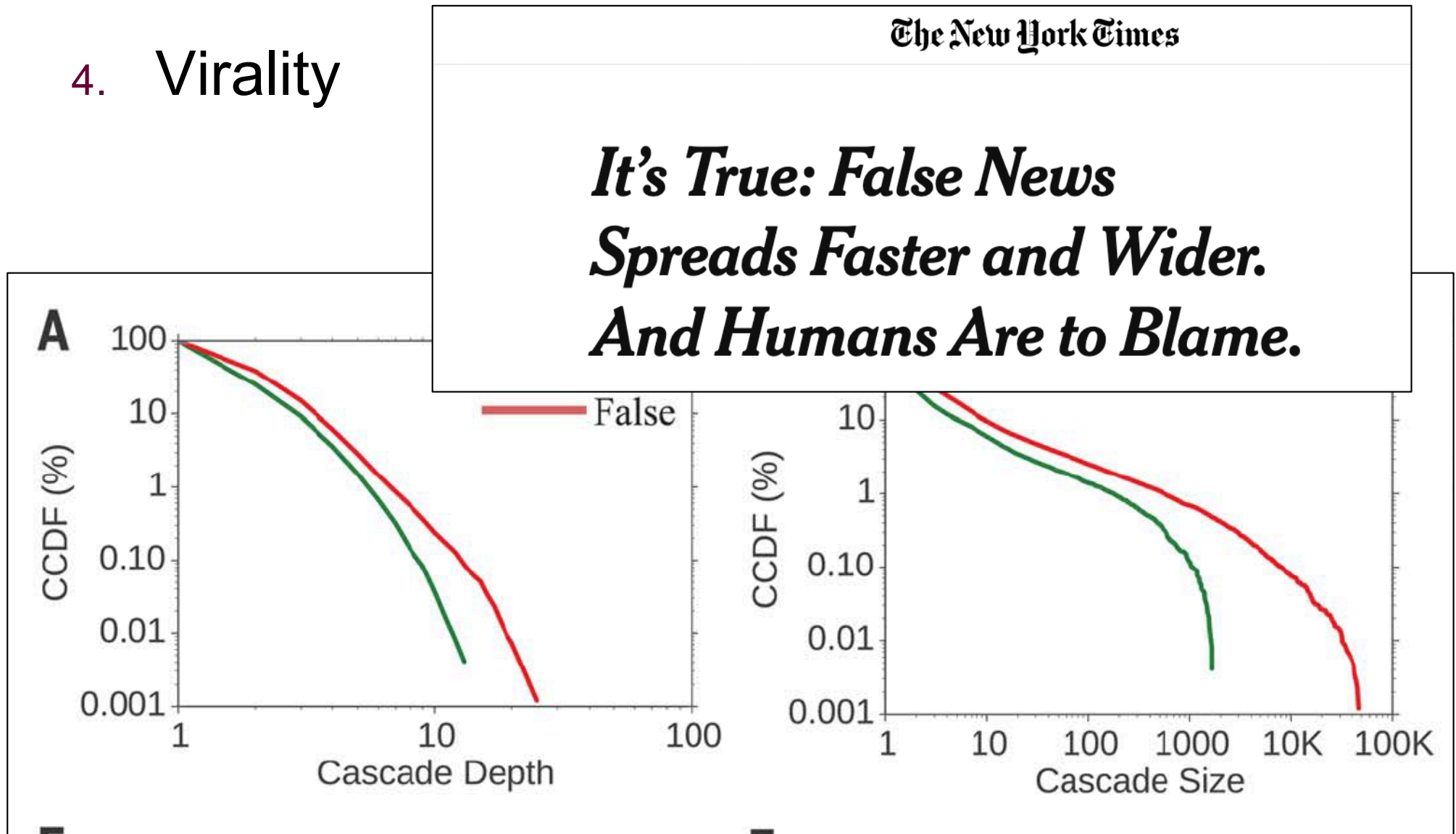
Source: <https://www.knightfoundation.org/features/misinfo>

In 2016, social bots played a significant role in spreading false information

Many of these accounts published multiple tweets with the **#NoDAPL** hashtag at nearly the same time, only a few minutes apart.

More Problems in Social Media?

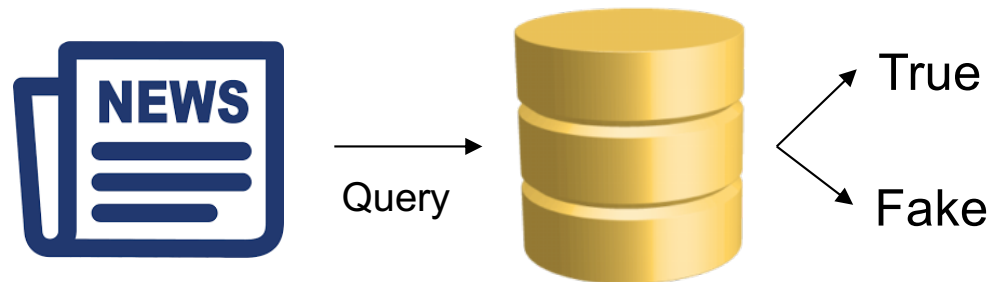
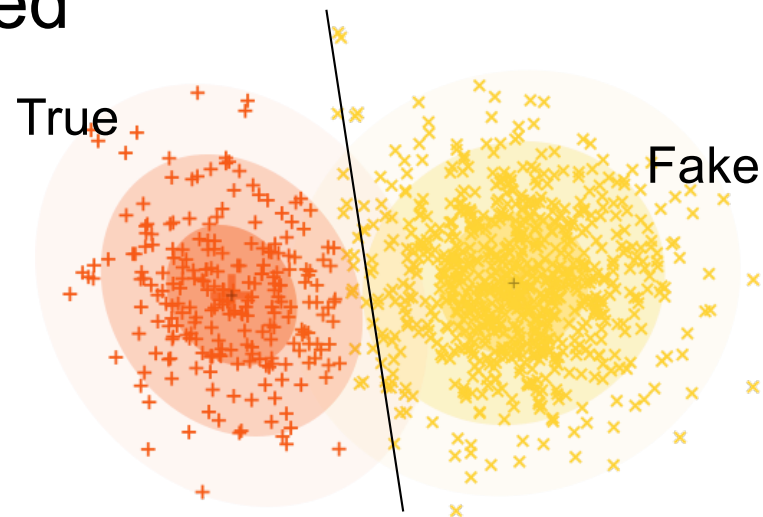
4. Virality



Source: Vosoughi et al., Science 2018

To Detect False Information

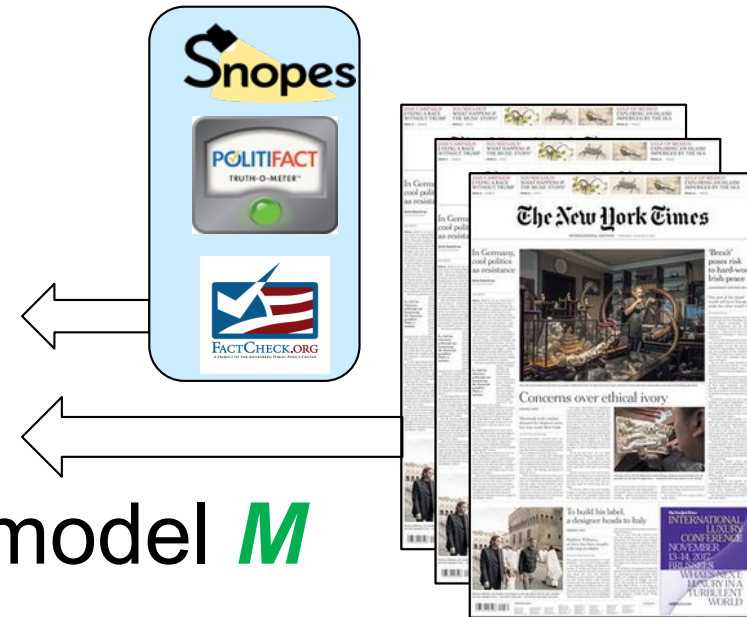
- Human Based
 - Manual fact-checking
 - Crowdsourcing based
- Machine Based
 - AI approach
 - DB approach



AI: Machine Learning Approach

In Training

- Learning
 - **P**: Features from “fake” news
 - **N**: Features from “true” news
- Feed (**P**, **N**) to ML to build a model **M**



In Deployment


- Feed a news story **A** to **M**
- **M** determines if **A** is fake or true news story

LAB: Fake-O-Meter

- In your smartphone browser, go to **Kahoot.it**
- Enter Game PIN, and Nickname to play

A shark photographed on the Highway outside of Naples, Florida during Hurricane Irma in 2017

13



Skip

0 Answers

▲ Fake

◆ Real

Educational Fake News Games

- <http://factitious.augamestudio.com/>
- <https://www.fakeittomakeitgame.com/>
- <https://playfakenews.com/>
- <https://hoaxy.iuni.iu.edu/fake-news-game/>
- <http://fakenews.game/>
- <https://boardgamegeek.com/boardgame/235085/fake-news-or-not>



Fake It
To Make It



https://www.fakeittomakeitgame.com/

LAB: Play Game (30 minutes)

Welcome to Fake It To Make It, a social-impact game about fake news. You can read more about this game [on this page](#). However, I suggest that you play first!

If you need to stop in the middle of your game, it should be okay, as long as you have local storage enabled. (You probably do.) Just come back to this main menu, where you will find your automatically saved games.

While this game does work in mobile browsers, you will probably be happier playing on something wider, like a computer or a tablet, instead of a phone.

Your existing games:

Don

Last played July 31, 2018

Total funds: \$50.00



Play

Delete

Jim

Last played July 31, 2018

Total funds: \$50.00

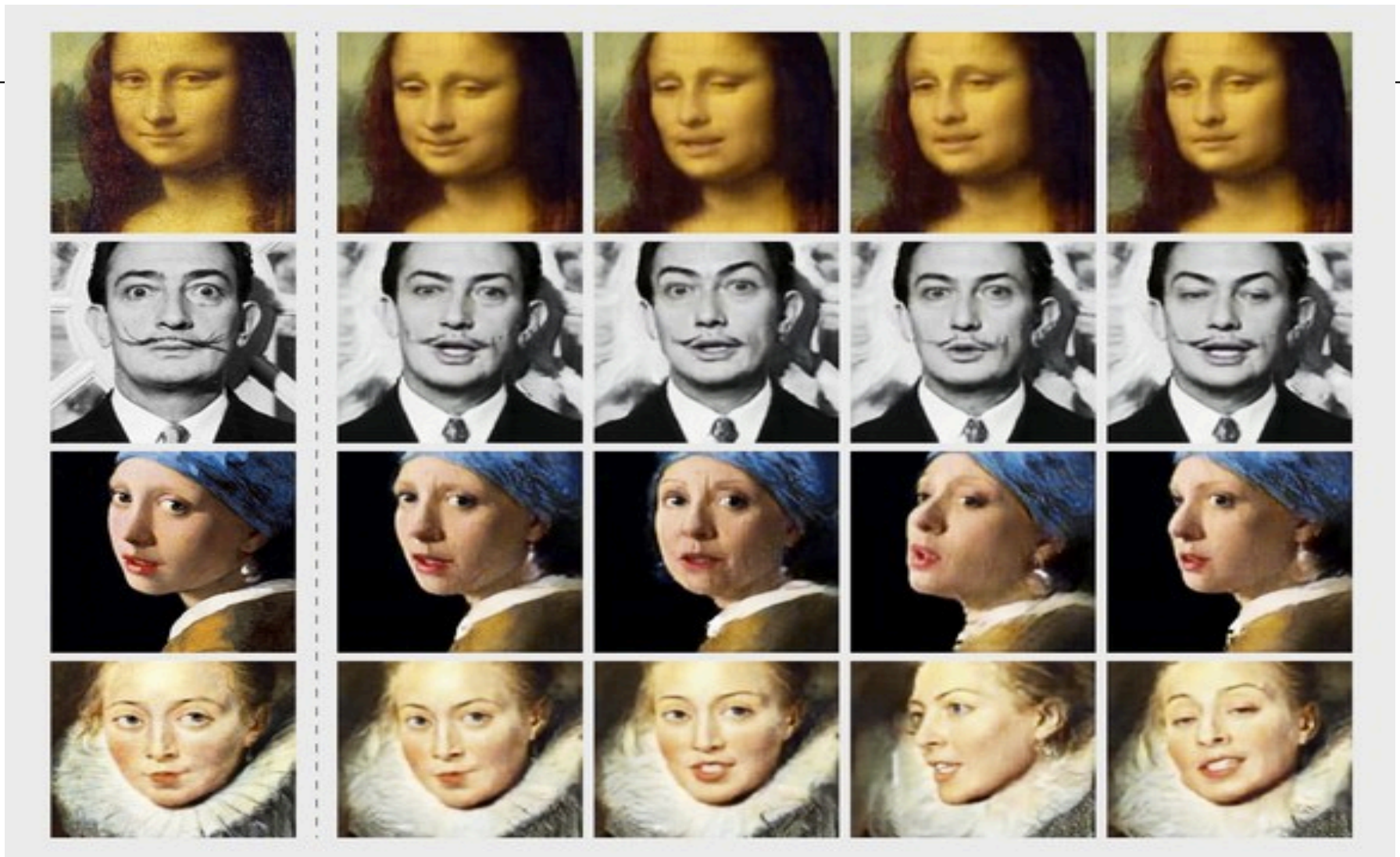


Play

Delete



Start A New Game



4. DEEPPFAKE

New Challenge: “Deepfakes”

1. AI method (GAN) generated artifacts
2. Manipulated artifacts hard to distinguish
 - Not “Shallowfakes”

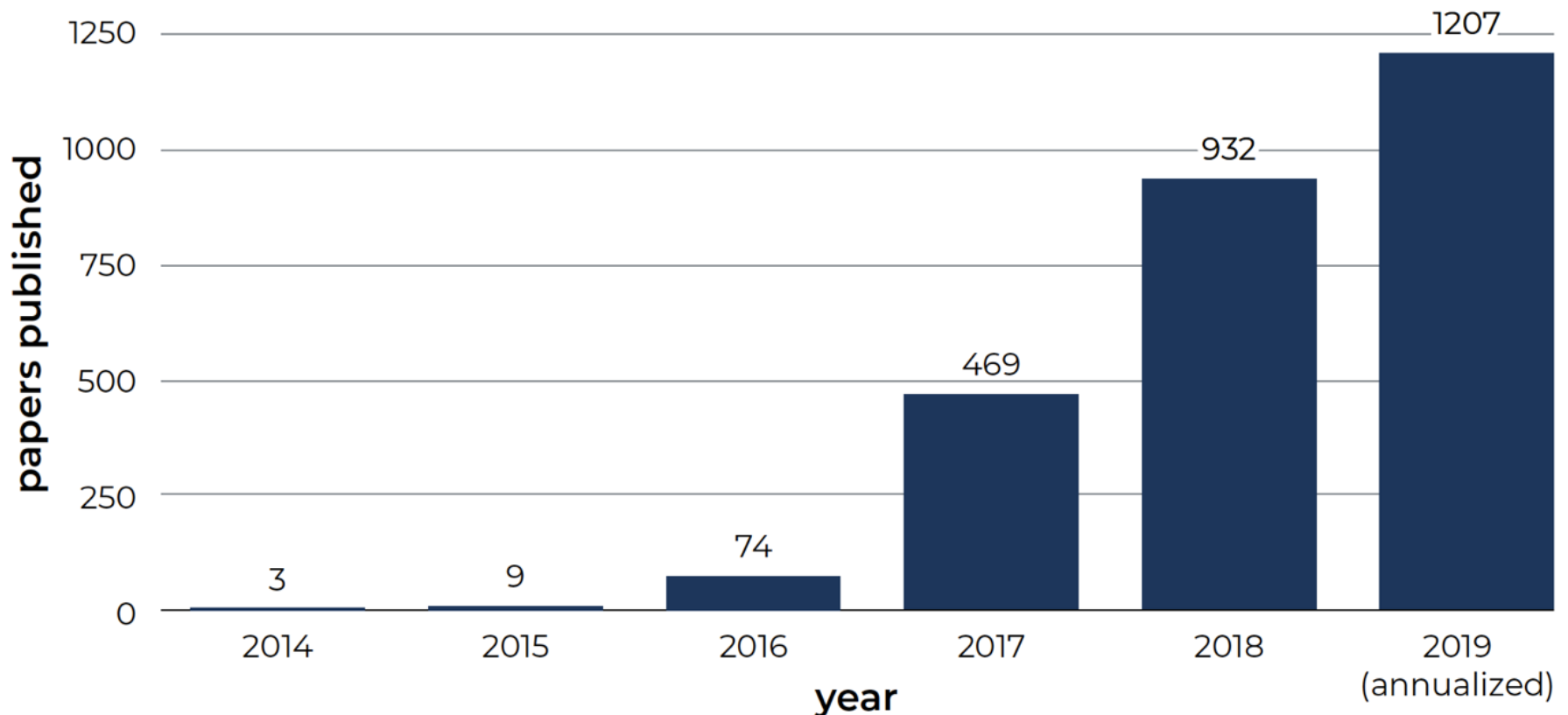


- Explosive effect ← When used in social media together with:
 - False information, Social bots, Clickbaits

Landscape of “Deepfakes”

- 14,678 deepfake videos [DeepTrace, 2019]
 - 96% are pornographic videos

Academic papers mentioning GANS in title/abstract published per year on the arXiv³



Eg, Deepfaked Text #1

Examples

Link found between autism and vaccines

Select an example or build an article below

Inputs

Domain

cnn.com

Generate

Give your article a domain source to write against

Date

October 10, 2019

Generate

Give your article a date to write against

Authors

Dongwon Lee

Generate

Give your article one or more authors to write against

Headline

Link found between autism and vaccines

Generate

Give your article a headline to write against

Article

Fake Article



cnn.com

Link found between autism and vaccines

October 10, 2019 - Dongwon Lee

Written by Dongwon Lee, CNN

Scientists say they have found a link between autism and vaccinations, citing increasing evidence over the past several years that vaccines are linked to cases of the condition.

Researchers at Stony Brook University in New York presented new findings on Wednesday linking autism with vaccines to the American Psychiatric Association's International Neuropsychiatric Association meeting.

"Many parents are particularly interested in this question because it has often been argued that vaccines cause autism," says the study published by the journal PNAS.

Autism is a mental disability that affects cognitive, language and social development. It is a leading disorder among children and is usually diagnosed in the first three years of life.

Signs that a child has autism include repetitive behaviors, delays in speech and language development, lack of affection and social interaction and

**Grover by
U. Washington**

Eg, Deepfaked Text #2

GPT-2 by OpenAI

Human

SYSTEM PROMPT (HUMAN-WRITTEN)

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

MODEL COMPLETION (MACHINE-WRITTEN, 10 TRIES)

The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science.

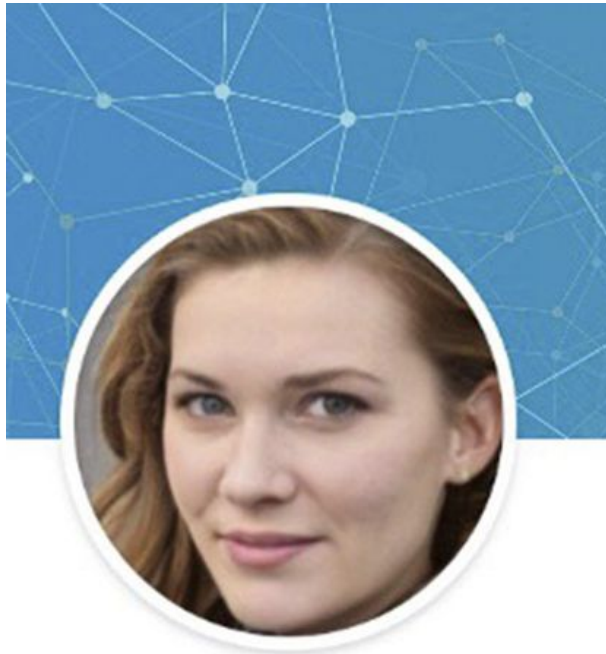
Machine

Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.

Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.

Pérez and the others then ventured further into the valley. "By the time we reached the top of one peak, the water looked blue.

Eg, Deepfaked Image



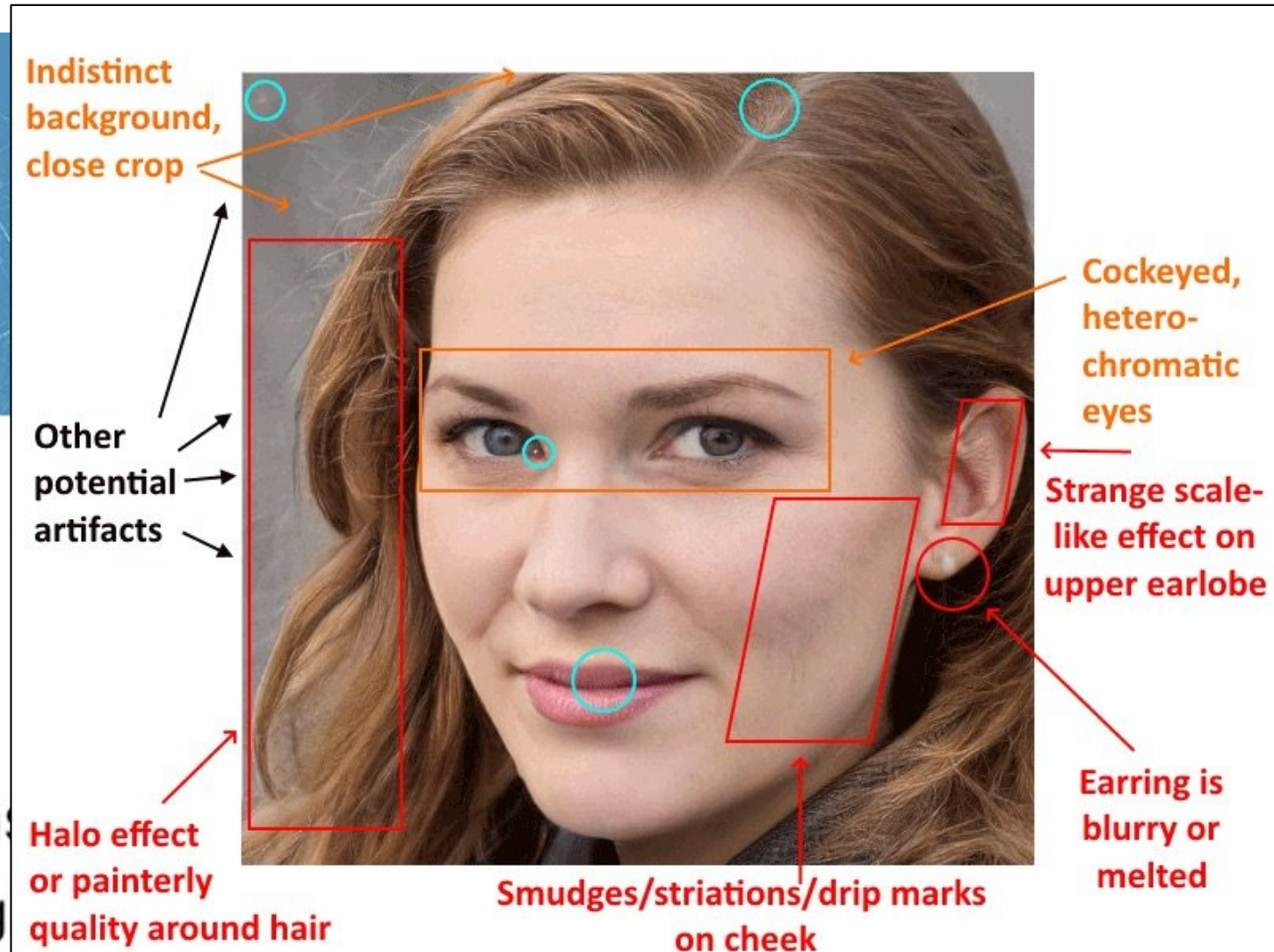
Katie Jones

Russia and Eurasia

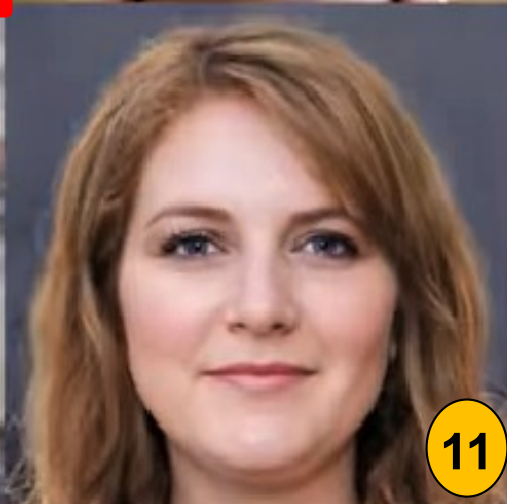
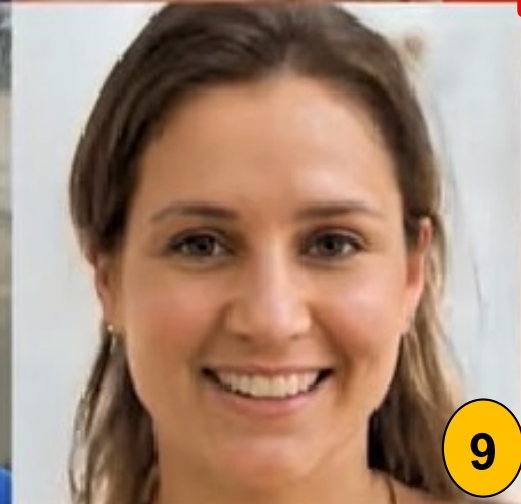
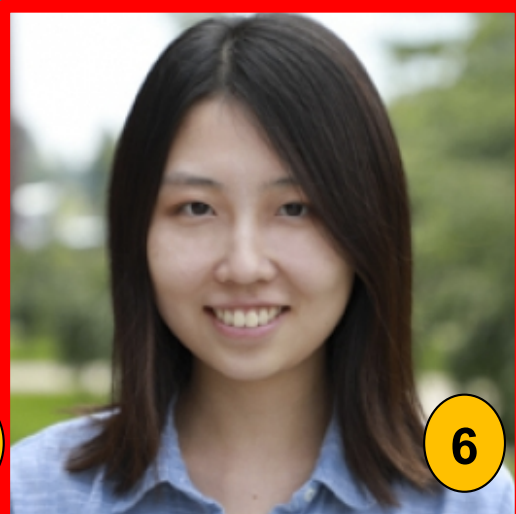
Center for Strategy

University of Michigan College of Literature, Science...

Washington · 49 connections

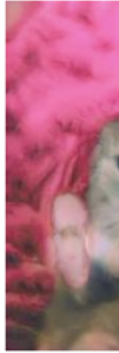


<http://thispersondoesnotexist.com>



These Cats Do Not Exist

Learn More: [General](#)



ENTIRE GUEST SUITE

Ponta Mer Media Aireoper

Athens



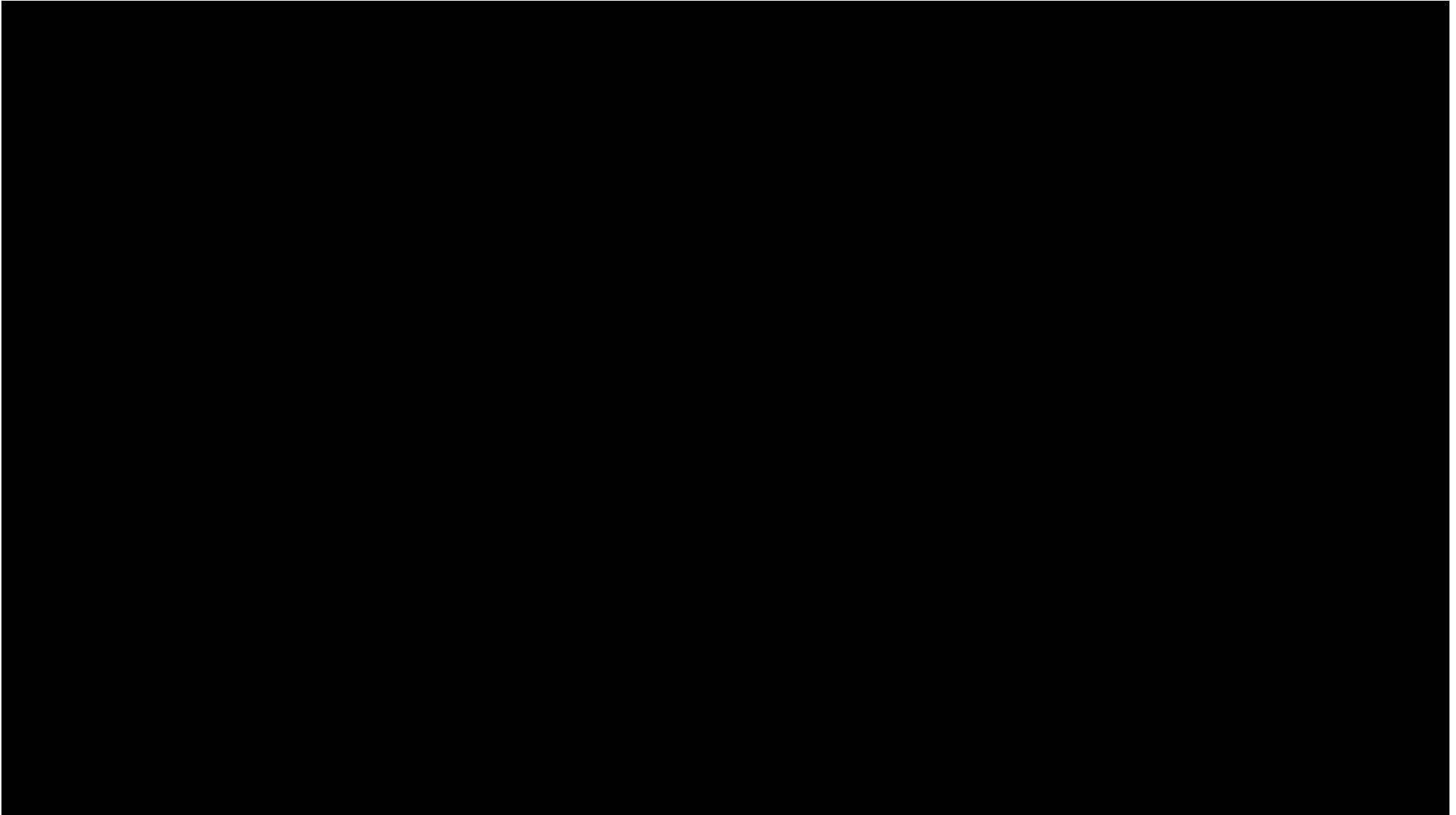
Alice

7 guests 3 bedrooms 4 beds 2 baths

My place is 2 mins frou shopping malls and restaurants open 360 HM have a double cable television. We will offer all the necessary noise and make your stay enjoyable with high speed internet available in the kitchen but also will be fully resolved to a comfortable room while day. Guests have access to the balc. Tons of green pubs and all the restaurants are 3 trams ride away. Steeped in a bathroom, WC, French balco front and back porch to ensure guests to sit around the corner for entry or dinner. There is

<https://thisrentaldoesnotexist.com/>

Eg, Deepfaked Video #1



Eg, Deepfaked Video #2



Eg, Deepfaked Video #3

Text-based Editing of Talking-head Video

Ohad Fried^{*}, Ayush Tewari[^], Michael Zollhöfer^{*}, Adam Finkelstein[†], Eli Shechtman[‡],
Dan B Goldman, Kyle Genova[†], Zeyu Jin[‡], Christian Theobalt[^], Maneesh Agrawala^{*}

^{*} Stanford University

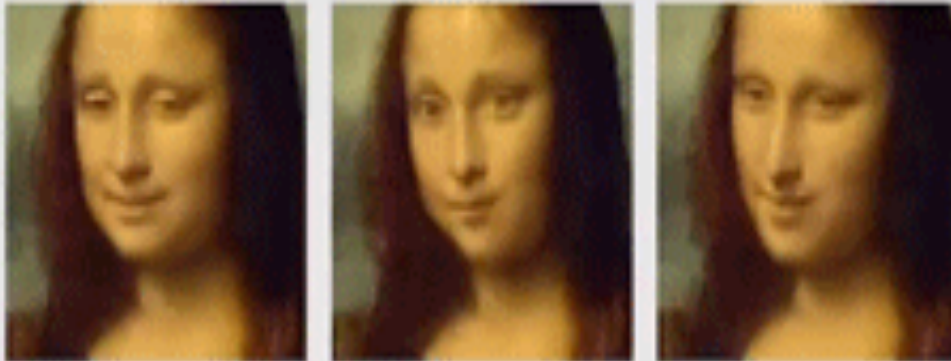
[^] Max Planck Institute for Informatics

[†] Princeton University

[‡] Adobe

Eg, Deepfaked Video #4

Living portraits



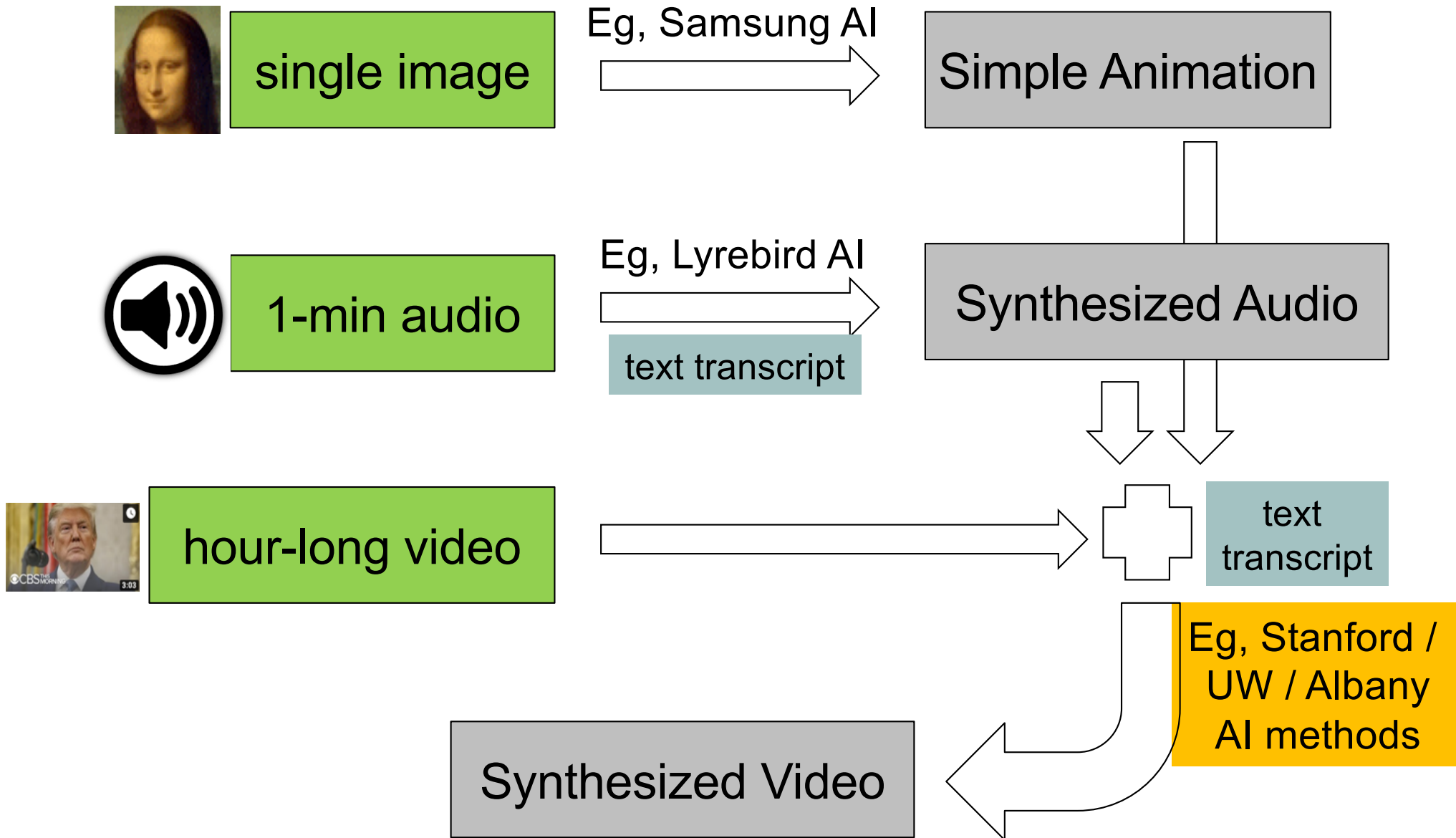
Eg, Deepfaked Video #4



Eg, Deepfaked Video #5

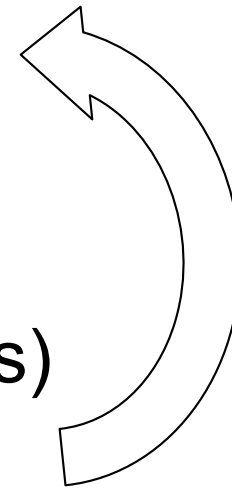


Potential Deepfake Scenario



If I were an Adversary ...

- **Human** adversary
 - Create a fake image/video
 - Write a fake news story
 - Plant it into social media (via bots)
- **Machine** adversary with **deepfake capability**
 - BEGIN
 - **Synthesize** a fake image/video
 - **Synthesize** a fake news story
 - Plant it into social media (via bots)
 - END



Repeat
Million
times

Implications

- No known instances in which deepfakes have actually been used in disinformation campaigns” – Deeptrace, 2019
- Documentation is no longer evidence
- “Implied **false** effect”
- “Reality apathy” – Aviv Oyadya, 2019

The biggest threat of deepfakes isn't the deepfakes themselves

The mere idea of AI-synthesized media is already making people stop believing that real things are real.

by **Karen Hao**

Oct 10, 2019

“The Liar’s Dividend”
-- Robert Chesney
and Danielle Citron

MIT
Technology
Review

Arms Race against Deepfakes

Deepfake Detection Challenge

Deepfake Detection Challenge

Deepfake Detection Challenge invites people around the world to help detect deepfakes and tampered media. Identifying tampered deepfakes rapidly evolve, so we're working together to build better

Events Starting in October 2019

NEWS

Home | Video | World | UK | Business | Tech | Science | Stories | Er

Technology

Google makes deepfakes to fight deepfakes

f m t

Real Fake

浙江卫视 G20 楊瀾訪談錄 YAN LAN ONE ON ONE

这期间有着全球的金...

Defense One

NEWS THREATS POLITICS BUSINESS TECH IDEAS

DARPA Is Taking On the Deepfake Problem

Arms Race against Deepfakes

Support The Guardian

Available for everyone, funded by readers

Contribute →

Subscribe →

Sign in

The
Guardian

News Opinion Sport Culture Lifestyle



World ▶ Europe US Americas Asia Australia Middle East Africa Inequality Cities More



California

California makes 'deepfake' videos illegal, but law may be hard to enforce

H.R.3230 - Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2019

116th Congress (2019-2020) | [Get alerts](#)

BILL Hide Overview ✕

Sponsor: [Rep. Clarke, Yvette D. \[D-NY-9\]](#) (Introduced 06/12/2019)

Committees: House - Judiciary; Energy and Commerce; Homeland Security

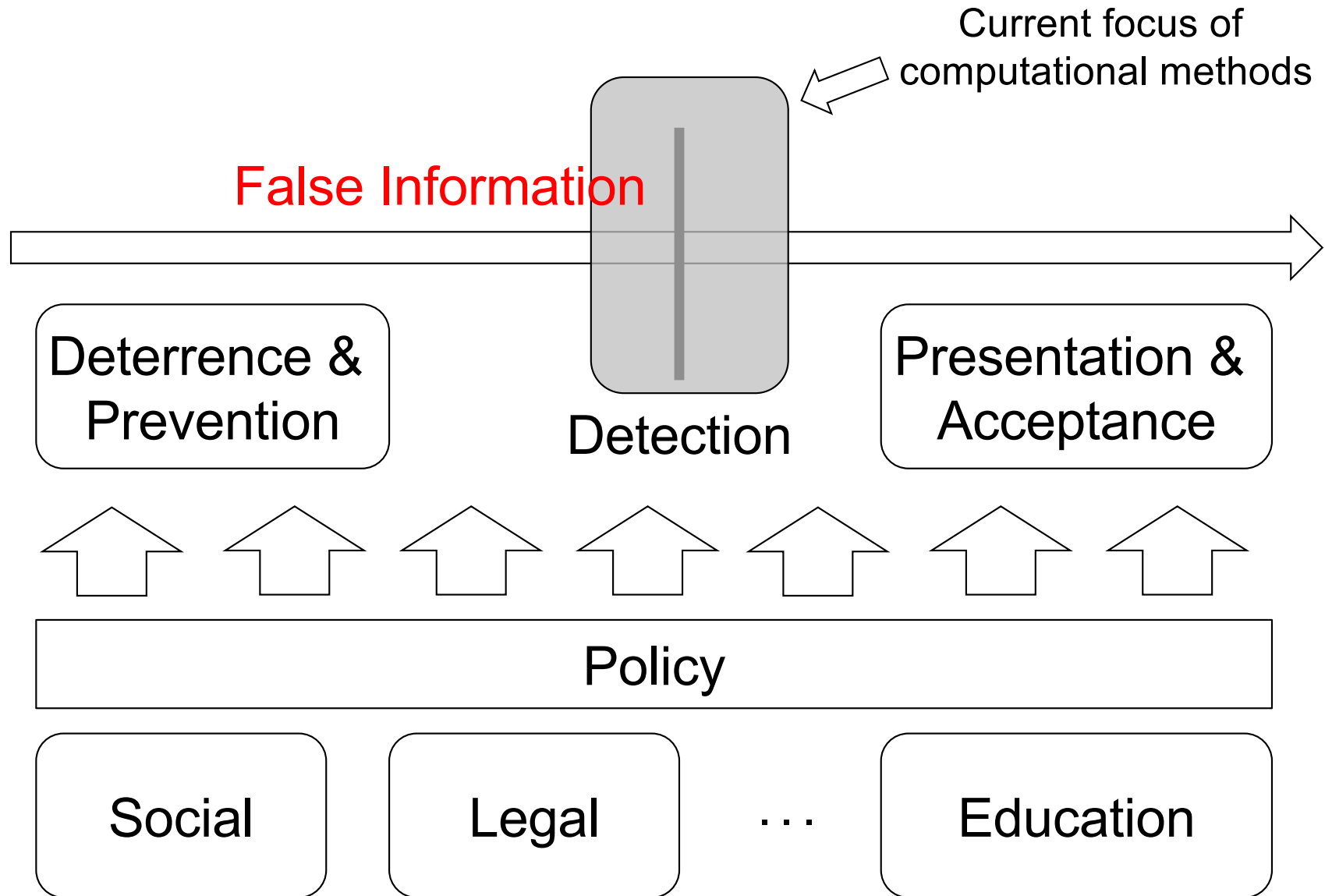
Latest Action: House - 06/28/2019 Referred to the Subcommittee on Crime, Terrorism, and Homeland Security. ([All Actions](#))

More Thoughts

- Technological solutions alone cannot solve the problem of **false information** (and **deepfakes**)
 1. Detection does not directly lead to removal
 - Eg, slowed-down Pelosi video
 2. Not fast enough
 - Eg, virality of false information
 3. Little help to real victims
 - Eg, deepfaked porn videos

Source:
Angela Chen, MIT
Technology
Review, 2019

More Thoughts



Development Timeline

- Plan to release V1 of materials: Dec 15, 2019

2-3 week long version

- Special topic plug-in to other related classes
 - CompSci undergraduates
- Recruit instructors to use part of materials in their classes in Spring and Summer 2020

Download of Materials

<https://tinyurl.com/fraud-informatics>

