

---

# Field Guide for Validating Premium Ad Inventory

---



**Dr. Augustine Fou**  
Anti-Fraud Forensic  
Scientist & Founder MSCG

marketing.science  
consulting group, inc.



**Andy Hibbler**  
VP, Brand Advertising &  
Partnerships

whitepages™



**Rami Essaid**  
CEO & Co-Founder

 **distil**  
networks

---

# Agenda

---

- Terminology and Standards
- The State of Digital Ad Fraud
  - Landscape and Trends
  - Display Fraud and Search Fraud
  - Top Issues for Advertisers and Publishers
  - Best Practices and Anti-Fraud Platforms
- Distil Networks Approach
- Whitepages Premium Ad Inventory Case Study
- Q&A

---

# Terminology and Standards

---

## Landscape of Compliance and Standards Organizations

- MRC ([www.mediaratingcouncil.org](http://www.mediaratingcouncil.org)) - viewability standards adopted and aligned with Quality Assurance Guideline
- IAB ([www.iab.net](http://www.iab.net)) - Anti-Fraud Working Group, authored the Ad Fraud Taxonomy
- TAG ([www.tagtoday.net](http://www.tagtoday.net)) - Formed from IAB, 4As, and ANA to focus exclusively on trustworthy supply chain, addressing fraud, piracy, and Non-Human Traffic (NHT)



# Bad Bots in the Taxonomy

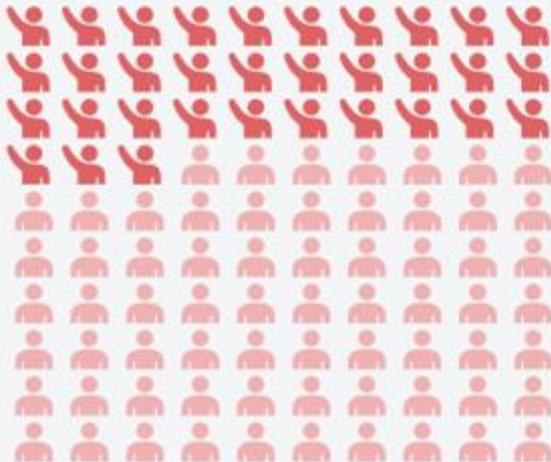
<b>Non-Human Traffic</b>	
<b>Type</b>	<b>Notes</b>
Hijacked session/device	fully automated browser
Crawler masquerading	posing as searchbot or human
Datacenter traffic	hosting companies, ISPs, etc.

<b>Non-Traditional Sources of Traffic on Ads</b>	
<b>Type</b>	<b>Notes</b>
Adware	additional HTML calls without user action
Proxy traffic	anonymous or known
Non-browser user agent	fake mobile app, invalid UA, etc.
Browser pre-rendering	loading ad before page

# How would you define your organization within the Ad industry?

## BUYERS

Purchasing Ad Inventory



33% Buyers

## SUPPLIERS

Representing Ad Inventory



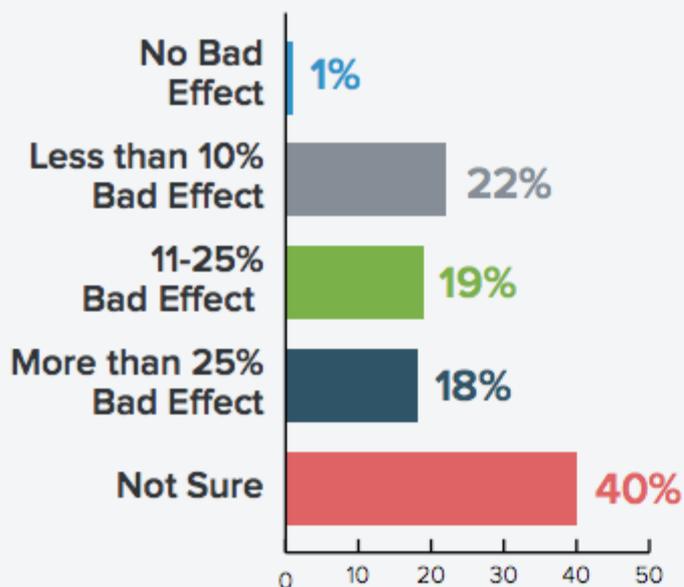
67% Suppliers

Brought to you by:

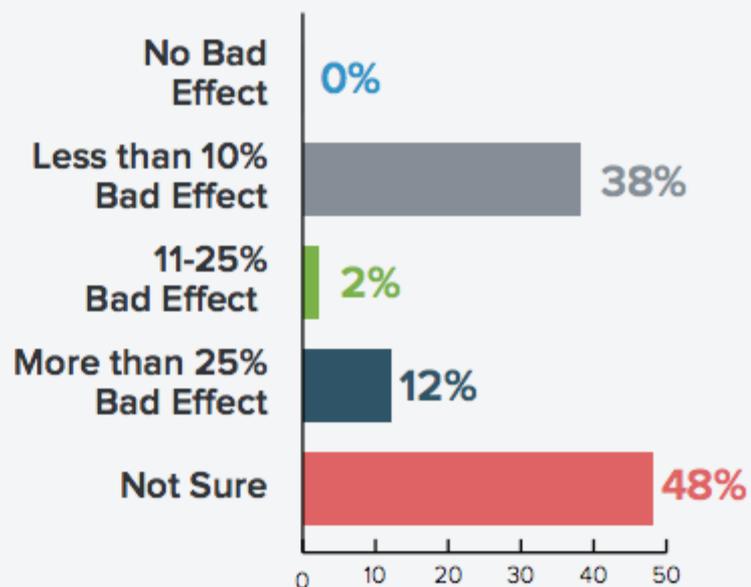


## How much does bot traffic negatively affect your web properties or the performance of online campaigns

### BUYERS



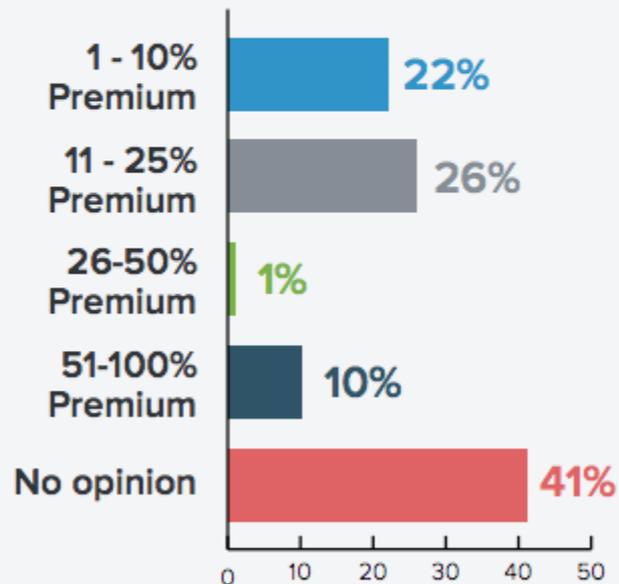
### SUPPLIERS



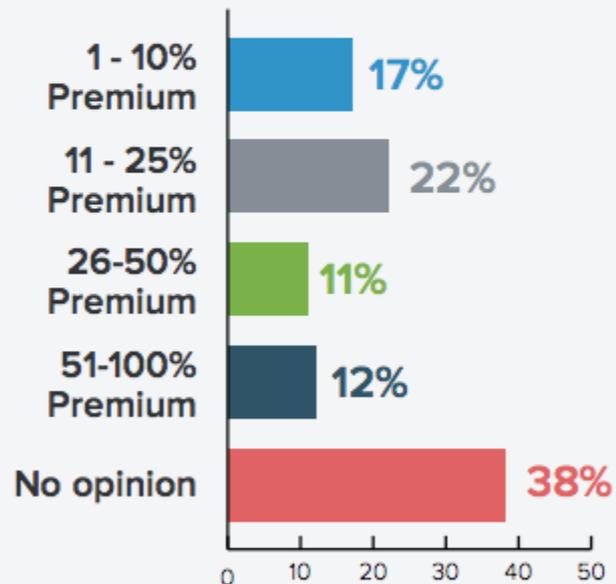
Brought to you by: **iab.** 

## How much of a premium rate do you believe “certified human” traffic is worth?

### BUYERS



### SUPPLIERS



Brought to you by: **iab.**  **distil**  
networks

---

---

# Live Survey

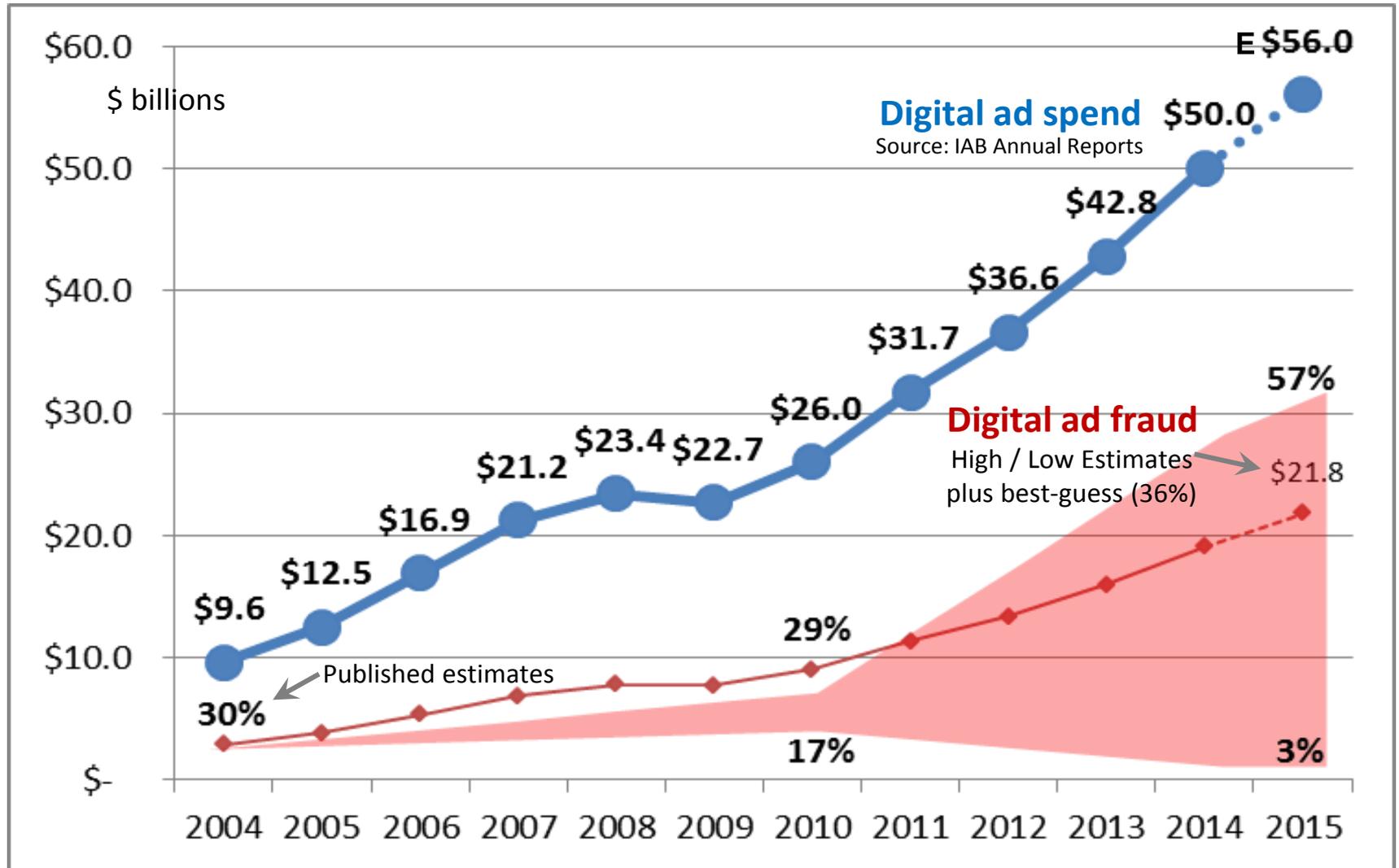
---

# State of Digital Ad Fraud

---

## **Landscape and Trends**

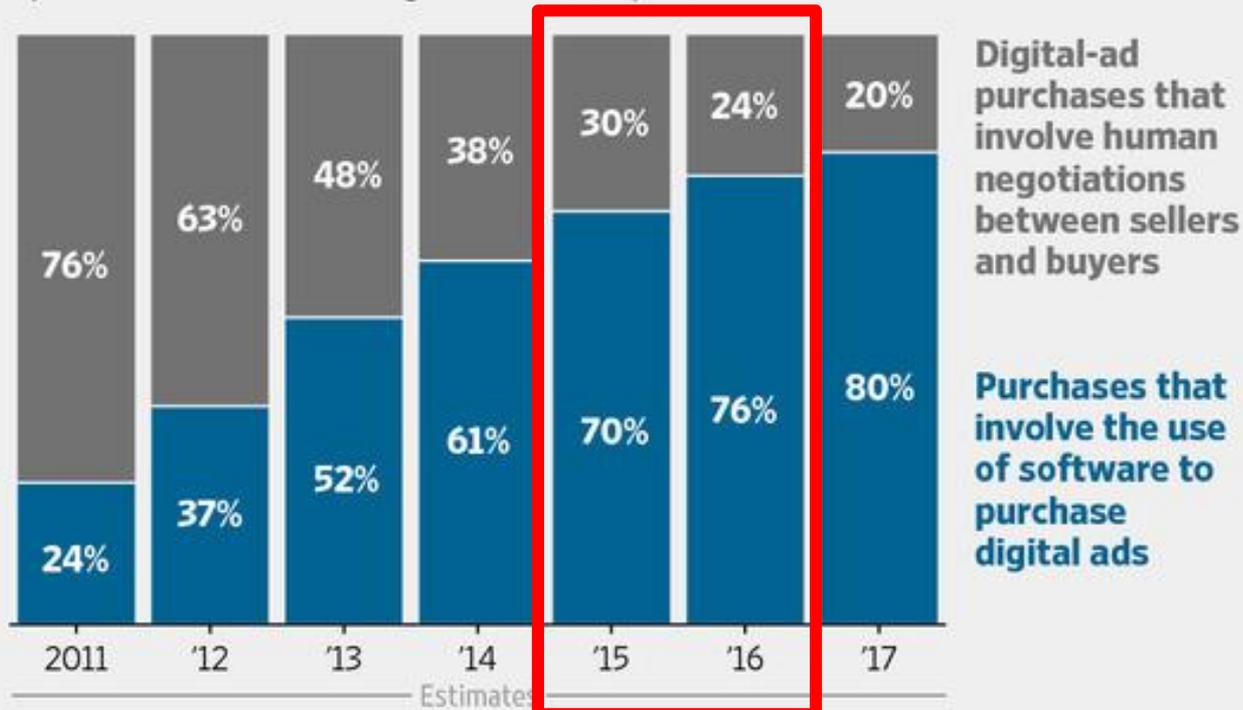
# Fraud goes up as digital ad spend continues upward



# Motive & Opportunity – More programmatic, more fraud

## Unseen Hand

More than half of online display-ad buys last year by value were made using automated systems.



Source: Magna Global Note: Figures don't add up to 100% because of rounding. The Wall Street Journal

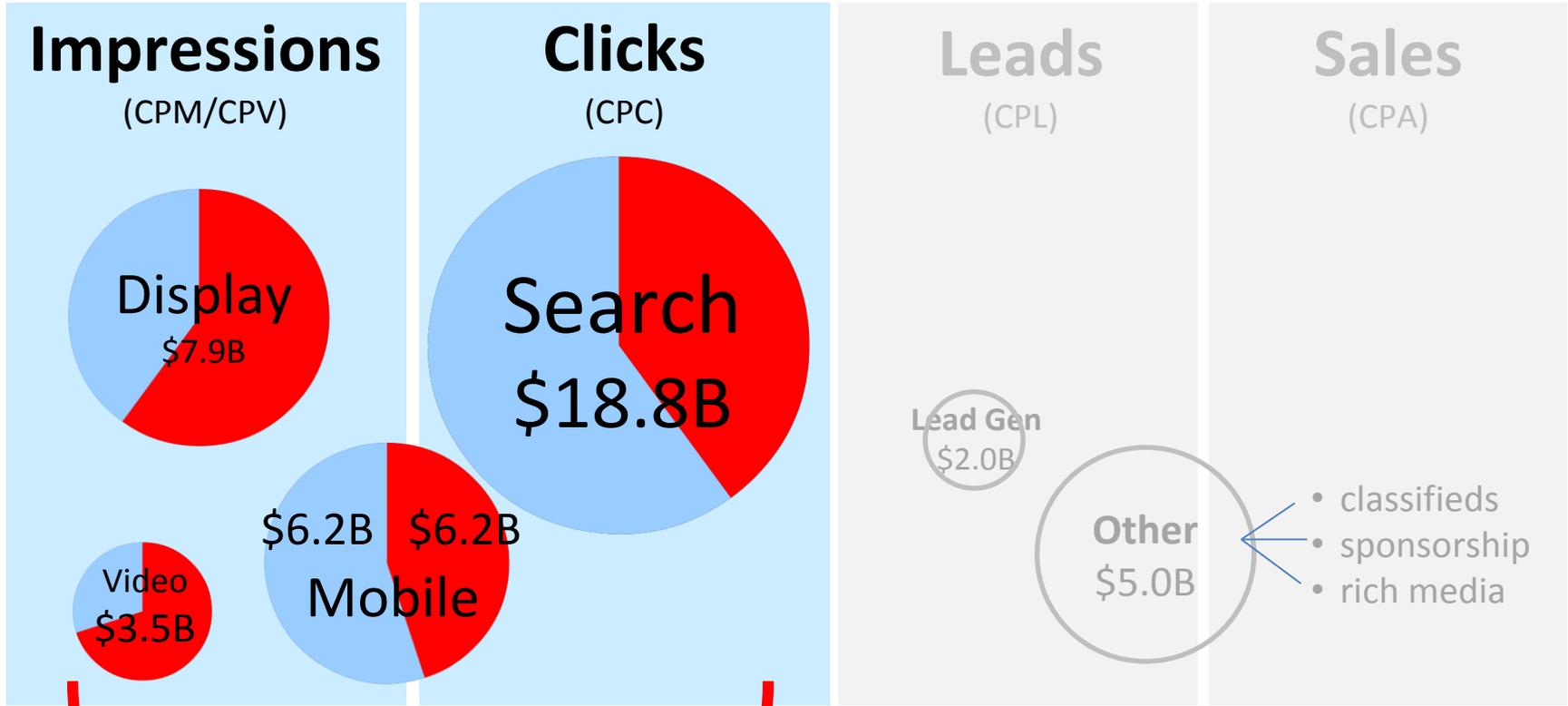
As more digital ads are placed entirely programmatically, the opportunity for fraud continues to increase.

Bad guys also fully automate their digital ad fraud operations, using programmatic tools.

# Impressions and clicks remain the biggest targets

**\$42.5B**

**\$7B**



**86% digital spend**

 estimated fraud  
 not at risk

- classifieds
- sponsorship
- rich media

# Industry associations are leading the charge

September 30, 2014 - IAB, 4A's, and ANA announce cross-industry organization to fight ad fraud, malware, and piracy.

*"It is imperative that the ecosystem addresses online fraud and improves media measurement and transparency."*

**Bob Liodice**

President and CEO of the ANA

*"We've invested [heavily] in data and data technology and [have] been very much at the forefront of fraud verification,"*

**John Montgomery**

GroupM

*"Together we can continue to rebuild the trust that is necessary for the interactive marketing industry to thrive"*

**Nancy Hill**

President and CEO, 4A's

*"We must create an all-inclusive program that identifies qualified participants, and commits them to good actions, guaranteed by continual monitoring and sanctions for non-compliance"*

**Randall Rothenberg**

IAB

---

# State of Digital Ad Fraud

---

## Display Fraud and Search Fraud

# Two main types of fraud and how each is generated

## Impression (CPM) Fraud

(includes mobile display, video ads)

1. Put up fake websites and load tons of ads on the pages

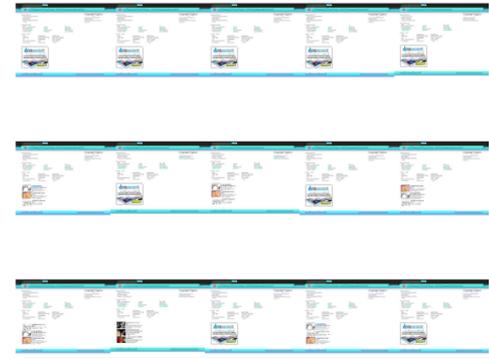


2. **Use bots** to repeatedly load pages to generate fake ad impressions (launder the origins of the ads to avoid detection)

## Search Click (CPC) Fraud

(includes mobile search ads)

1. Put up fake websites and participate in search networks



- 2a. **Use bots** to type keywords to cause search ads to load
- 2b. **Use bots** to click on the ad to generate the CPC revenue

# Bad bots are the primary tool for stealing ad dollars

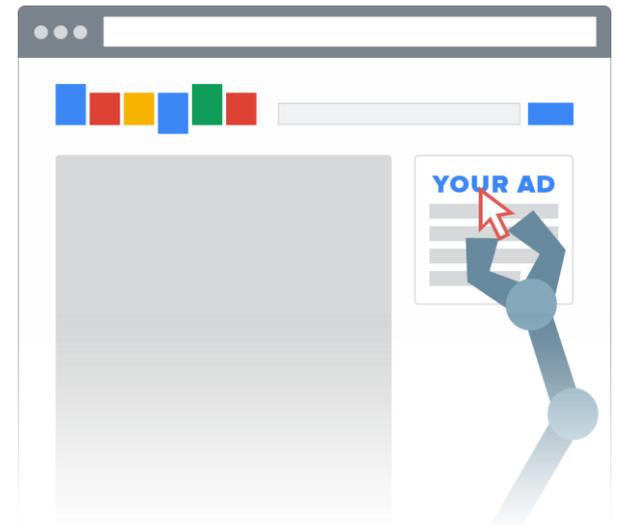
---

## Generate fake ad impressions and clicks

- Bots repeatedly load pages or ads to make impressions
- Bots can also “type” search keywords and click on search ads to create fake clicks

## Retargeting fraud

- Bots come to advertisers site to collect cookie
- Bots visit other sites owned by bad guys
- Retargeting follows them around to show ads
- When ad impression is served on bad guys sites, they earn the CPM revenue

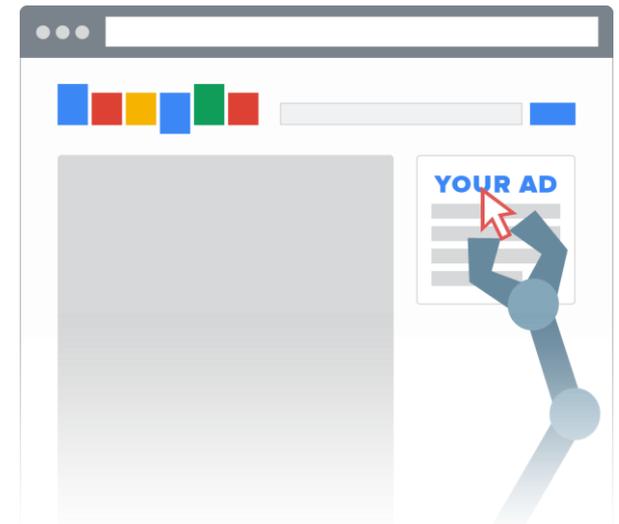


# Bad bots are the primary tool for stealing ad dollars

---

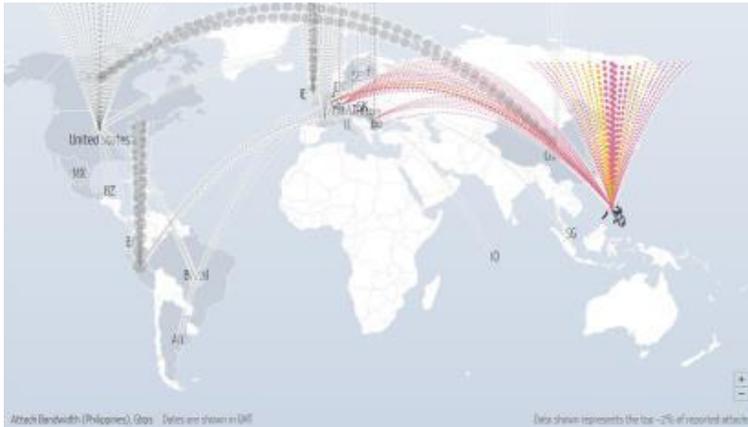
## Diverting ad dollars

- **CTR manipulation:** bots go to premium site to generate pageviews without clicks (lowers the sitewide CTR ); optimizers divert dollars to sites with higher sitewide CTR (bad guy sites)
- **Fake viewability:** bad guys stack all ads behind each other so they are all viewable; optimizers divert dollars to sites with higher viewability (bad guy sites)



# Different kinds of bots generate fake impressions/clicks

## Botnets (from datacenters)



## Malware (on PCs)



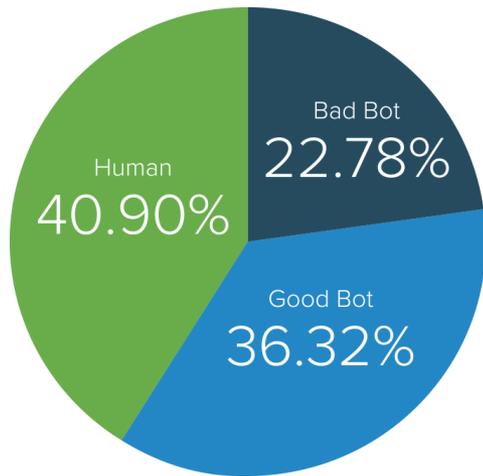
## Javascript (on webpages)



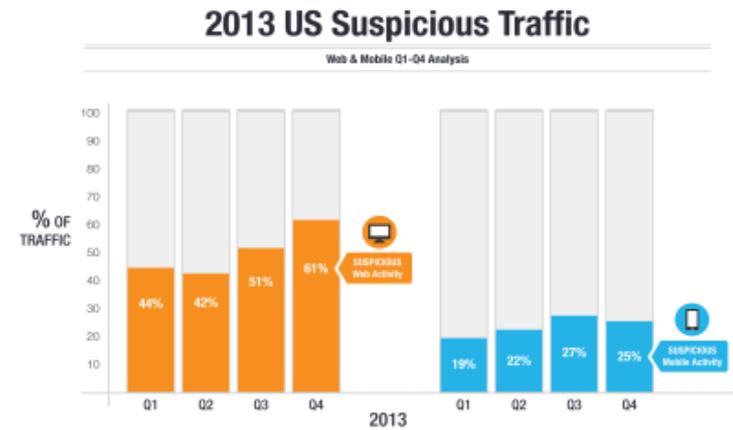
## Toolbars (in-browser)



# Multiple sources confirm levels of bot activity and traffic



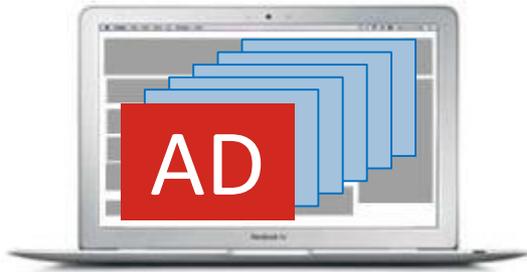
Source: [Distil Networks 2015 Bad Bot Report](#)



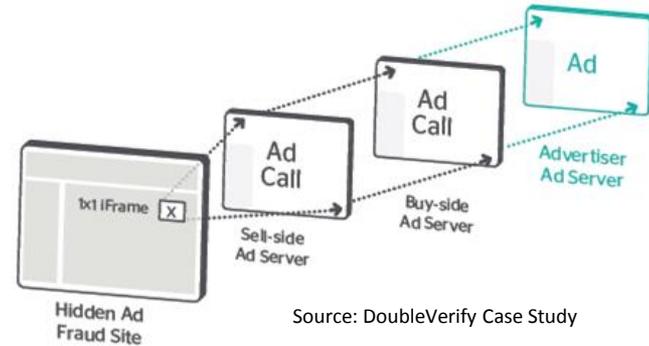
Source: [Solve Media 2013](#)

# Techniques to generate more impressions

## Ad stacking (dozens in one call)

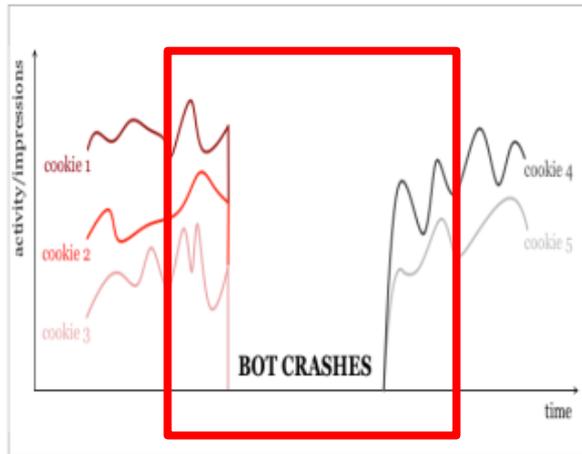


## Pixel Stuffing (hiding more ads)



Source: DoubleVerify Case Study

## Sourced Traffic (more pageviews)



## Ad Injection (replace ads)

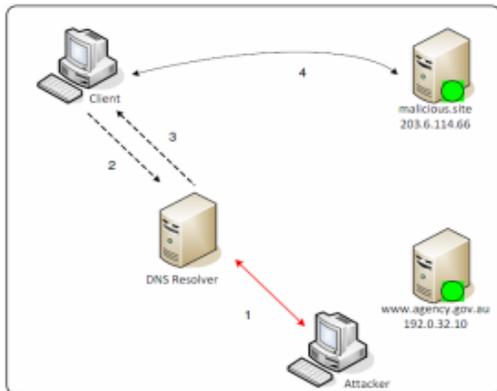


Image Source: [BenEdelman.org](http://BenEdelman.org)

The ANA/WhiteOps study also found rampant injection fraud, including one publisher whose site was hit with 500,000 injected ads every day for the duration of the two-month study.

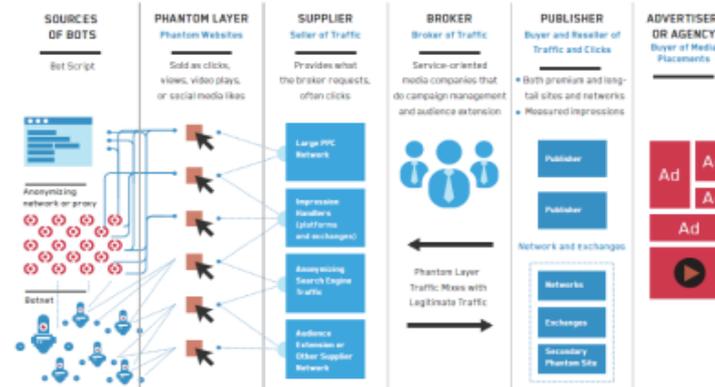
# Techniques to hide fraudulent activity

## Domain / IP Spoofing



Make it appear that the impression came from legit site

## Impression Laundering



Source: [Whiteops/ANA Study Dec 2014](#)

## Stacked Redirects

1. <http://viewsite.me/landing.php?p=QlpoOTFBWSZTWXI%2BUww>
2. <http://viewsite.me/landing.php?s=e8>
3. <http://74.117.113.21/cuout.php?cu=>
4. <http://74.117.113.21/cuout2.php?cu=>
5. <http://2402208.r.msn.com/?id=8vN2vmOGqqOFJPwUMROPyconcentrated-regular-bleach%252F%3>
6. <http://s0.srtk.net/www/delivery/rd.phpconcentrated-regular-bleach%2F%3>

**Known blackhat technique** to hide real referrer and replace with faked referrer.

Example how-to:  
<http://www.blackhatworld.com/blackhat-seo/cloaking-content-generators/36830-cloaking-redirect-referer.html>

## Passing fake variables

[http://www.olay.com/skin-care-products/OlayPro-X?utm\\_source=msn&utm\\_medium=cpc&utm\\_campaign=Olay\\_Search\\_Desktop\\_Category+Interest+Product.Phrase&utm\\_term=eye%20cream&utm\\_content=TZsrSzFz\\_eye%20cream\\_p\\_2990456911](http://www.olay.com/skin-care-products/OlayPro-X?utm_source=msn&utm_medium=cpc&utm_campaign=Olay_Search_Desktop_Category+Interest+Product.Phrase&utm_term=eye%20cream&utm_content=TZsrSzFz_eye%20cream_p_2990456911)

Easily trick analytics platforms to think the impression came from legit source

---

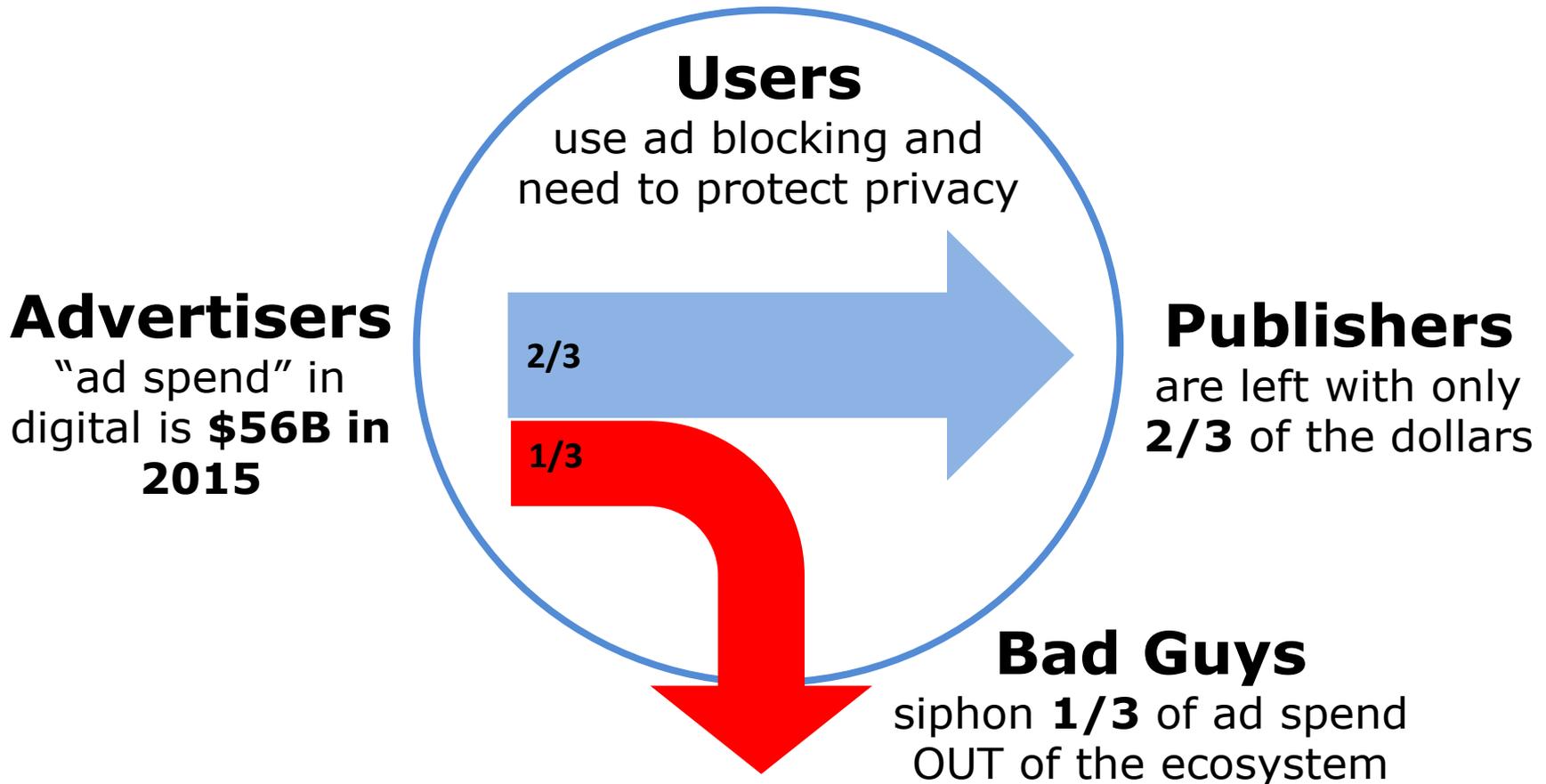
# State of Digital Ad Fraud

---

## Impact on Advertisers and Publishers

# Fraud siphons 1/3 of dollars out of ad ecosystem

- Ad dollars are being **siphoned OUT of the ecosystem** into the pockets of the bad guys
- Advertisers have lower ROI due to fraud (fake impressions, non-humans)
- Publishers have lower revenues (ad dollars stolen by bad guys)



---

# Advertisers and Publishers are Harmed Differently

---

## Advertisers

1. Advertisers attempt to increase revenues by increasing digital reach and frequency.
2. But more ad impressions doesn't mean more ads seen by people
3. Bots causing webpages and ad impressions to load directly cost advertisers -- in the form of wasted ad dollars that never have a chance to convert into paying customers

## Publishers

1. Publishers attempt to increase revenues by increasing the number of ads served or the CPM they charge
2. Bots coming to pages not only drain web serving resources but they also have the potential to damage legitimate publishers' reputation when selling ads
3. Bots also steal valuable content and collect cookies, both of which help bad guys divert traffic and ad revenue to fraudulent sites they own, resulting in direct loss of ad revenue for publishers.

---

# State of Digital Ad Fraud

---

## **Best Practices and Anti-Fraud Platforms**

---

# Recommendations for reducing fraud

---

## **Focus less on impressions served or traffic, more on observable actions**

If we focus on actions that can be observed rather than tonnage of ads served, and incentivize our agencies differently, we avoid the temptation of buying lower cost (lower quality) ad inventory

## **Move beyond just viewability; tackle NHT/bot fraud**

Even if viewability was solved, if the impression was caused to load by a bot, it is still fraud and money wasted for the advertiser

## **Allocate budget toward mainstream, premium publishers**

Allocate ad spend away from long-tail sites in ad exchanges that have the most fraud; the more direct way to reduce bot fraud is to buy mainstream sites

## **Demand full transparency and independently verify**

If they can't show you where each ad impression was served, don't buy it; put in place your own checks and balances

---

# Comparison matrix of anti-fraud solutions

	<b>Description</b>	<b>Mitigation</b>
<b>On-page</b>	Typically javascript code installed on websites (like analytics). Detects parameters about the user and computing environment	blacklist domains that send traffic that is mostly bots
<b>In-ad exchange</b>	Real-time decisioning in the bid-stream using rules and blacklists (e.g. if a site is on a blacklist, don't serve the ad)	eject users and reject sites in real-time based on blacklists
<b>Challenge based</b>	Using captchas and other challenges which ask humans to verify themselves as humans	identify and whitelist humans for use in RTB decisioning
<b>Reverse Proxy</b>	Inspection of all page requests to prevent scraping, content theft, and fraud BEFORE the ad call and pixel fires. Provides real time protection without relying on IP address	blocks all traffic from bad bots, rejects individual browsers, builds real-time whitelists and blacklists

---

# Distil Networks Approach

---

## Distil Networks Bot Detection and Mitigation

---

# Dealing with Unsolicited NHT

---

## Single Pixel with Ad

Tech: Pixel for Tracking

Single js injection

Predictable .js every time

Probabilistic

Reactive, post ad call

UA & IP = fingerprint

IP/Network intelligence

vs

## Reverse Proxy

→

Tech: Security & Performance

→

Multiple injections

→

Randomized, Spoofproof

→

Technical tests on browser

→

Proactive, pre-page load

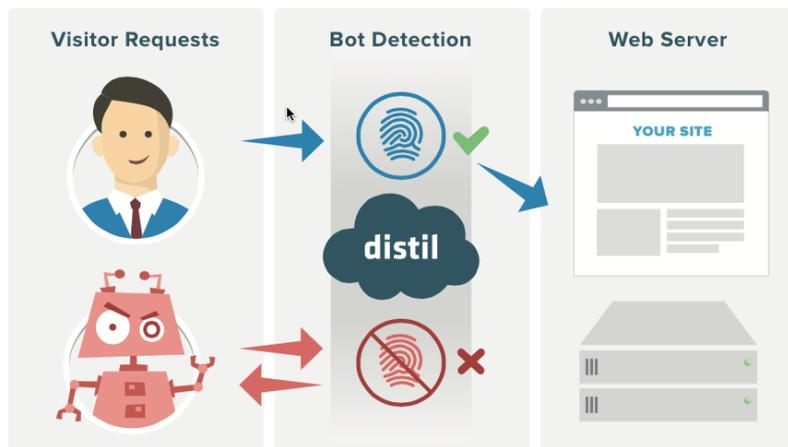
→

40+ parameters for fingerprint

→

Domain specific modeling

# Distil's Accurate Protection



## Inline Fingerprinting

Fingerprints stick to the bot even if it attempts to reconnect from random IP addresses or hide behind an anonymous proxy.

## Known Violators Database

Real-time updates from the world's largest Known Violators Database, which is based on the collective intelligence of all Distil-protected sites.

## Browser Validation

The first solution to disallow browser spoofing by validating each incoming request as self-reported and detects all known browser automation tools.

## Behavioral Modeling and Machine Learning

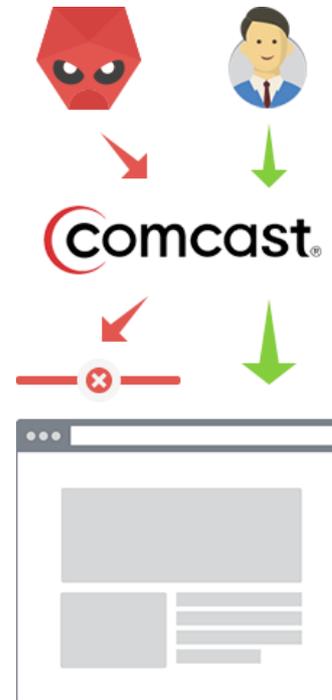
Machine-learning algorithms pinpoint behavioral anomalies specific to your site's unique traffic patterns.

# No Impact on Legit Users

Without **Distil**



With **Distil**



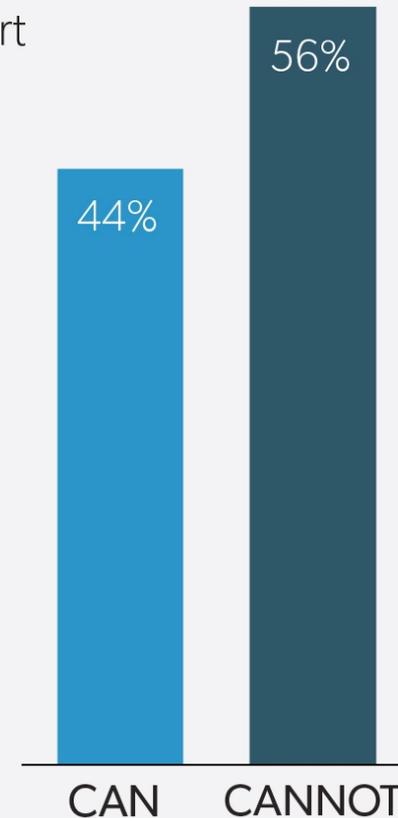
# DIGIDAY & Distil Networks Survey

So if confronted by advertisers with a report of bot-driven clicks and a refund request...

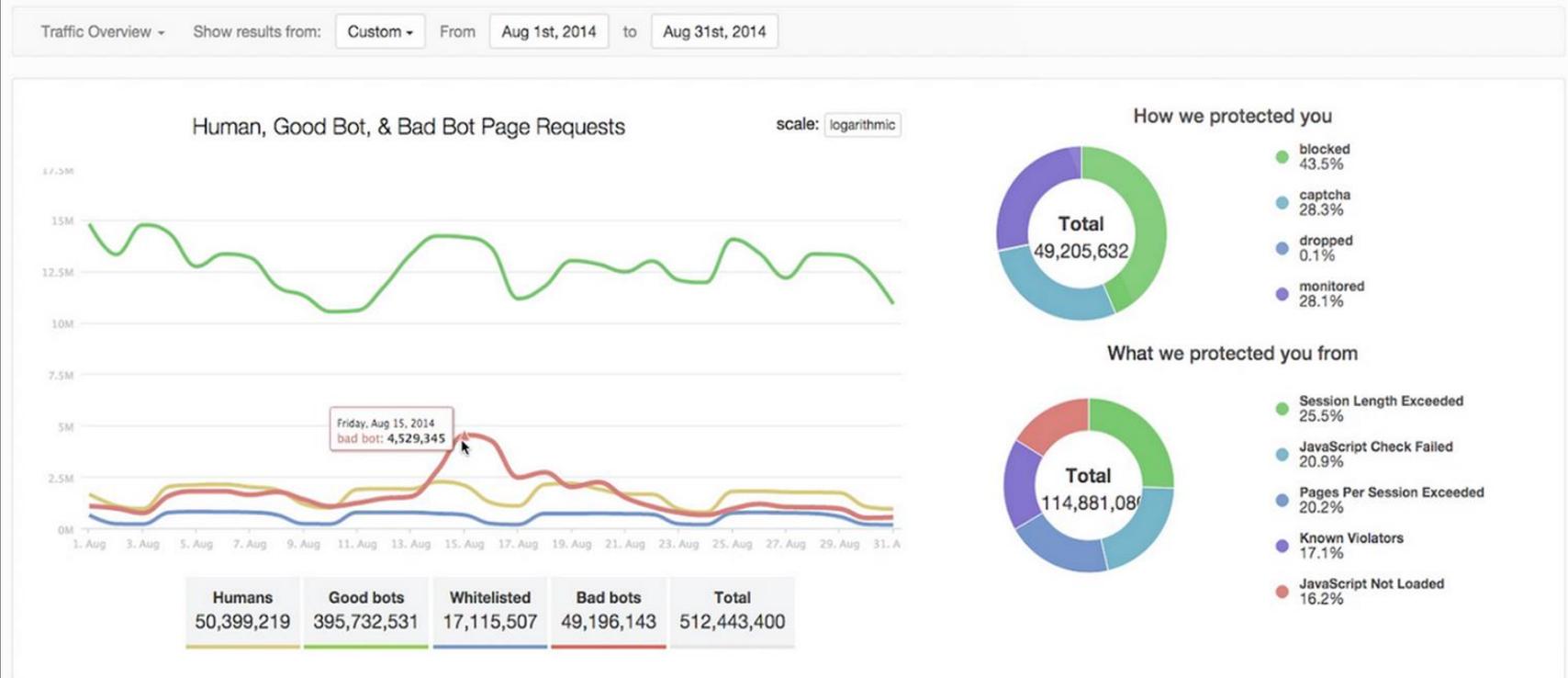
---

...how many publishers are able to quickly **investigate the issue** and **present their own findings** based on proprietary data?

This has led to **56 percent** of publishers losing ad dollars because of either **real** or **perceived digital ad fraud**.



# Understand Your Traffic



Click Fraud - Show results from: Last 14 days - Group by Agency - Help

1

Total Bot Clicks vs. Human Ad Clicks

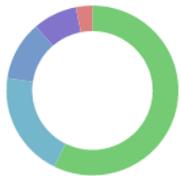


Human Ad Clicks  
22,857 clicks

Bad Bot Ad Clicks  
3,870 clicks

**14.48%**  
TOTAL BOTS

Top 5 Agencies by Bot Clicks



**Fraud Ads Inc.** 13,183 clicks

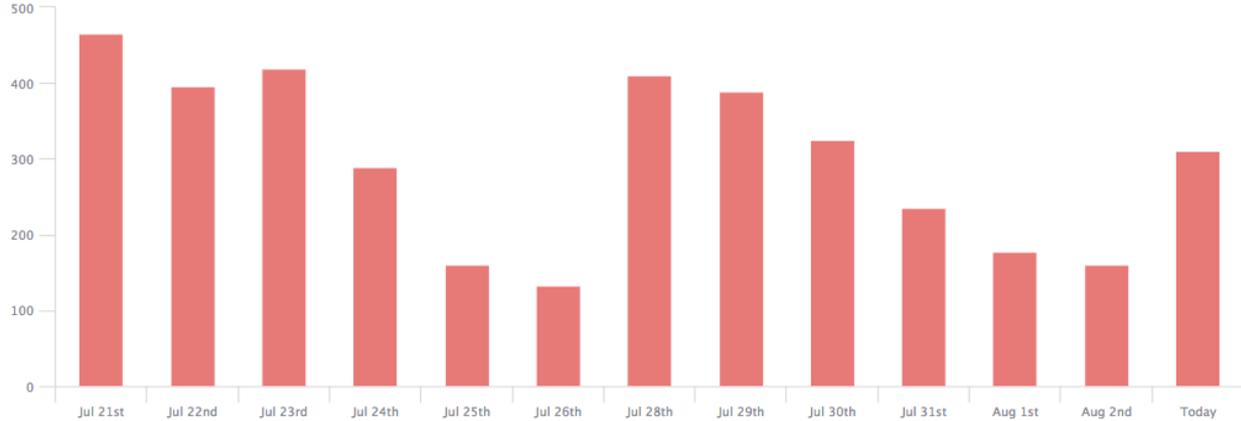
Google - Black+Friday%7CPri.. 4,593 clicks

Google - Black+Friday%7CSec... 2,586 clicks

Bing 1,913 clicks

Google - Google 712 clicks

Daily Click Fraud Activity



Fraud Ads Inc.

Agency	Total Bot Clicks	Total Ad Clicks	Bot Ad Clicks vs Human Ad Clicks
Fraud Ads Inc.	13,183	13,183	
Google - Black+Friday Primary	653	4,593	
Google - Black+Friday Secondary	392	2,586	
Bing	274	1,913	
Google - Google	100	712	
Bing - Black+Friday Primary	95	660	
Google - AP+75	61	372	
Google - OTT	60	373	

# Serve CAPTCHAs to Bots

[Back to dashboard](#)

[Account Management](#)



Domains

Reports

Settings

CAPTCHA Requests

Show results from:

Last 7 days

1

### Total Traffic vs. CAPTCHA Served

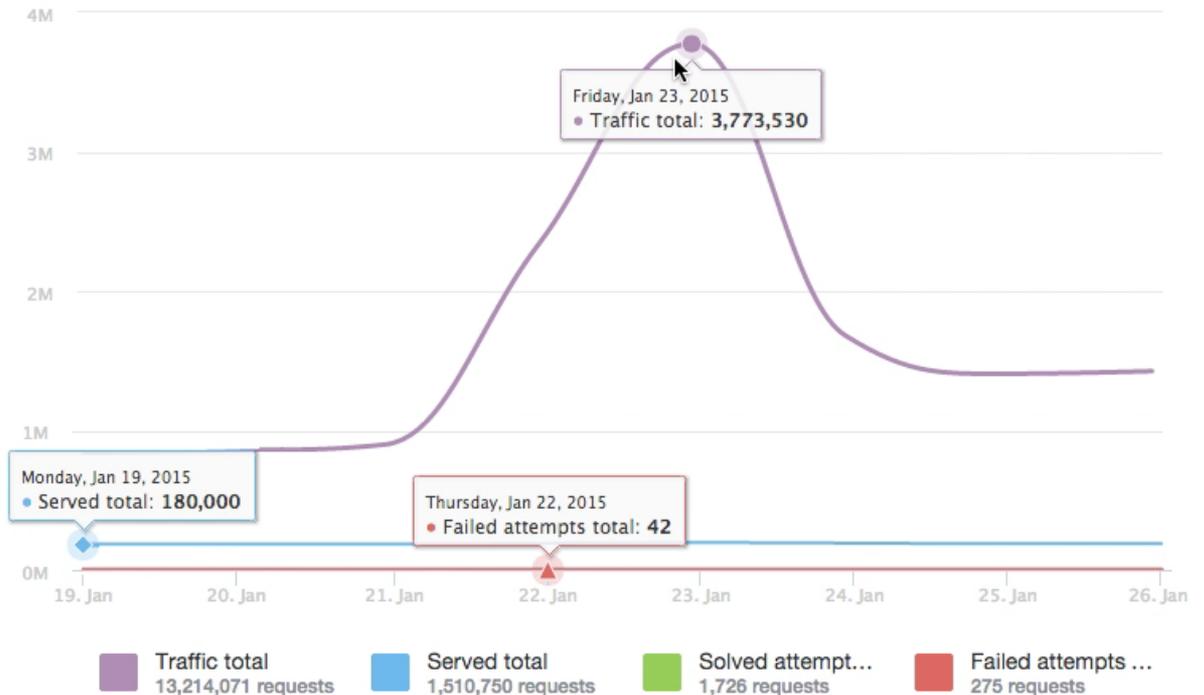


- Total Traffic  
13,214,071 requests
- CAPTCHA Served  
1,510,750 requests

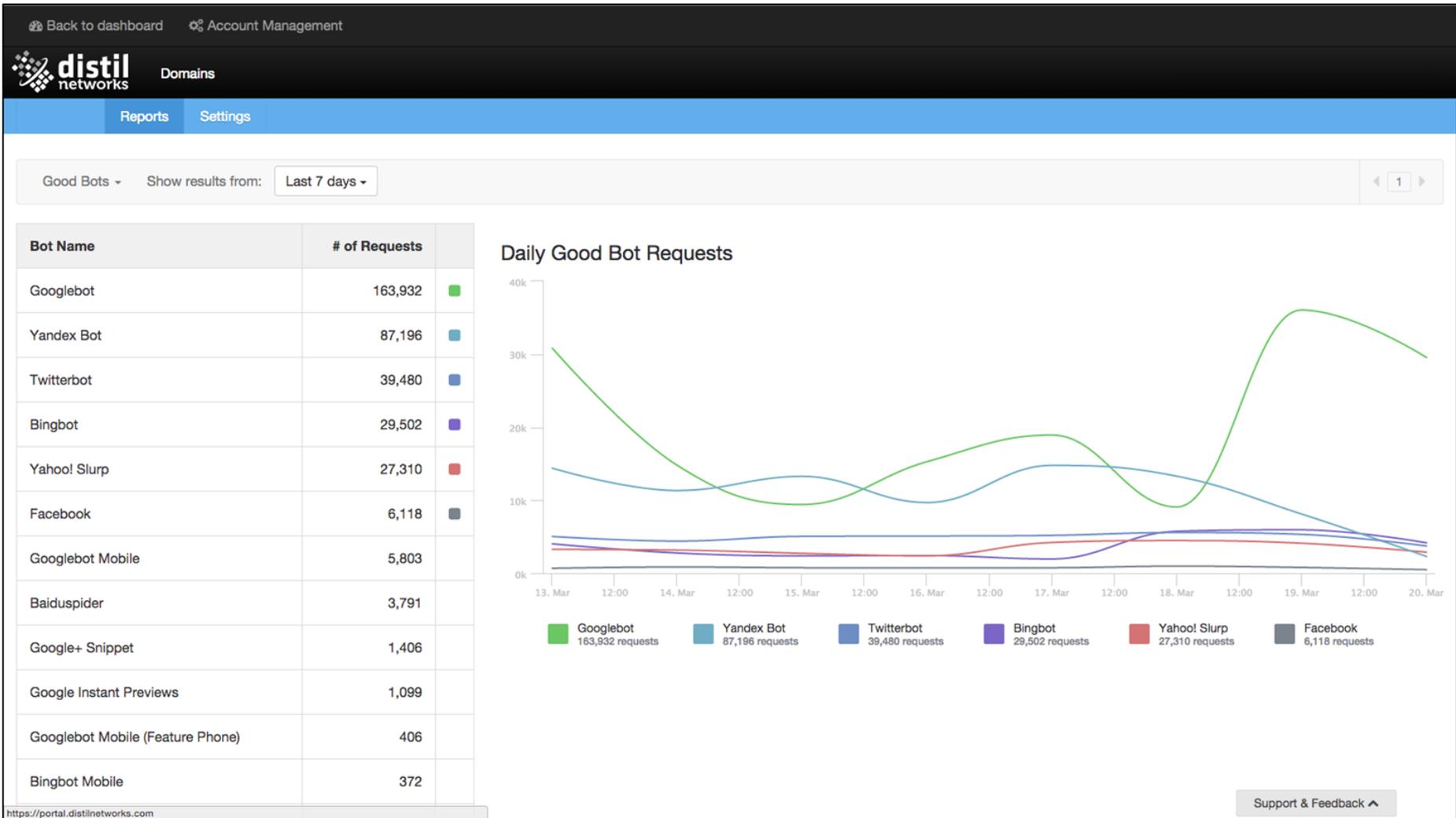
### Attempts Solved vs Failed



### Daily CAPTCHA Requests



# Good Bot Report



# Choose Your Bot Response

## Manage your threat responses

### **block** ▾ requests from **known violators**

Known Signatures protects you from Distil's curated list of known content thieves and malicious bots. This is a global IP and bot signature ACL that is maintained based on information gathered from all of our customers.

### **block** ▾ requests from **identities**

Enabling protection based on Identities will protect you from a Distil curated list of known user agents that primarily hog resources and cause unwanted stress on your webserver. In addition, this will also validate that specific request headers and request behaviors fall in line with specifications of the reported user agent.

### **block** ▾ requests from **aggregator user agents**

Monitor User Agents from accessing your website will use a Distil curated list to filter out content from RSS feed reader and Wordpress plugins. If you want your content accessed via these user agents, please leave this setting on "Monitor".

Monitor

Block

Captcha

Drop

### **block** ▾ requests from **automated browsers**

Automated Browsers are tools that often run a modified or light version of a web browser, capable of accessing your website and viewing the raw HTML, though unable to actually fully render your website or load and process external resources like JavaScript.

## Monitor

Monitor to inspect requests and record the traffic to Distil and/or your own server logs

## Block

Set to Block to serve the client an unblock verification form

## CAPTCHA

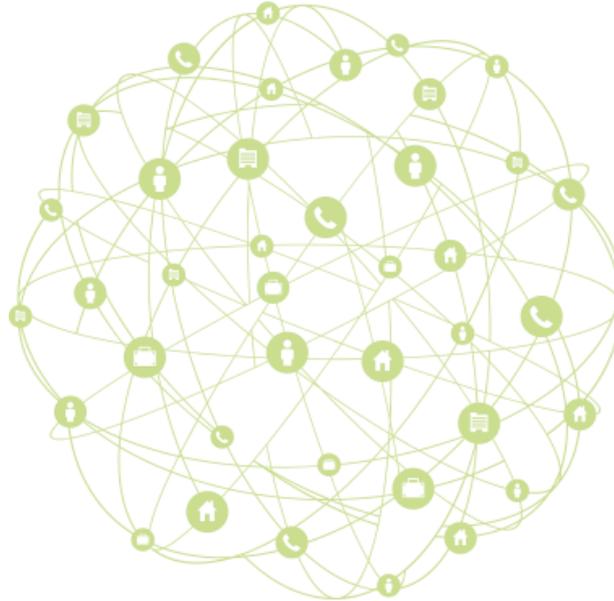
Serve a hardened CAPTCHA to test the client for verification

## Drop

Drop them to present them with an access denied page



# Digital identity starts here



Connect  
with people  
& businesses



Identify  
calls &  
texts

**PRO**

Detect &  
prevent  
fraud



# Whitepages is committed to eliminating fraud

---



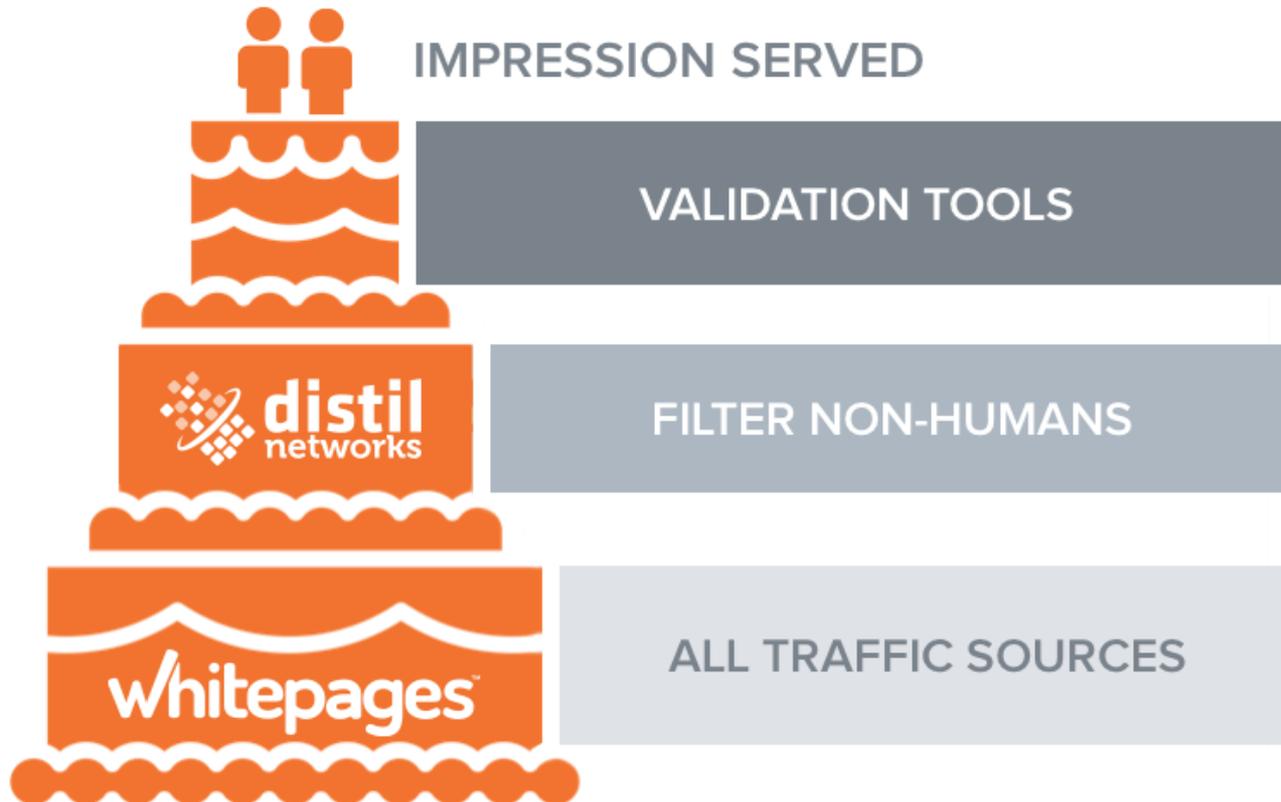
---

# Bots don't buy cars

---



# Our recipe for delivering honest impressions



---

# Whitepages Results Partnering with Distil Networks

---

- Stopped “comment spam” and prevented scrapers from stealing data
- Reduced time spent detecting bad bots and eliminated false positives
- Allowed servers to run and load impressions at optimal speeds
- Created a reliable and safe environment for brands (<5% IVT)

---

# Q&A

---



**Dr. Augustine Fou**  
Anti-Fraud Forensic Scientist  
& Founder MSCG

marketing.science  
consulting group, inc.



**Andy Hibbler**  
VP, Brand Advertising &  
Partnerships

whitepages<sup>®</sup>



**Rami Essaid**  
CEO & Co-Founder

 **distil**  
networks