DIGITAL PLATFORMS IN CRISIS:

A DECADE IN THE MAKING

APRIL 2018

digitalcitizens alliance

TABLE OF CONTENTS

INTRODUCTION

Digital platforms such as Facebook, Google, and Twitter are in crisis. A crisis of their own making. It's a crisis of trust, fueled by revelations that Russians used them to try to manipulate the 2016 election, that users' personal information was misused, that terrorists used them to spread Jihadi videos, and that they have become a "Dark Web Lite" marketplace for illicit goods and services.

It's not productive or beneficial for digital platforms to fall from grace with users. Companies such as Facebook, Google, and Twitter are not only a critical component of the modern economy but part of the fabric of society. Therefore, our society and economy are healthier when these platforms are trusted. But when a business model is ready-made for criminals and bad actors, it is inevitable that ultimately we would arrive at this moment.

There's an adage in the entertainment world that it takes half a lifetime to become an overnight success. For digital platforms, this crisis of trust is a decade or more in the making and is rooted in their unwillingness to monitor or take responsibility for the content that appears on their sites, no matter how harmful.

Platforms such as Facebook, Google, and Twitter seem to base that unwillingness on legal and business grounds.

From a legal standpoint, once they take responsibility for some, the platforms contend they would have to take responsibility for all. And they have a powerful protection to fall back on: Section 230 of the 1996 Communications Decency Act insulates digital platforms from responsibility for the content that appears on their sites—if they didn't participate in its creation.

From a business standpoint the content on their sites—even if objectionable and harmful—makes companies billions of dollars a year in revenues. In fact, taking action against this content could be contrary to their very business model.

INVESTIGATIONS *AND CALLS FOR* **REGULATION** *ARE THREATENING TO CONSUME DIGITAL PLATFORMS.*

On March 29, 2018, a memo surfaced in which a top Facebook executive said the company shouldn't let negative consequences get in the way of its mission. "Maybe it costs a life by exposing someone to bullies. Maybe someone dies in a terrorist attack coordinated on our tools. And still we connect people," wrote Andrew "Boz" Bosworth, Facebook Vice President.

That short-sighted thinking—based on legal exposure, growth and profits—has led us to a challenging place for both society and the platforms:

- Federal, state, and congressional investigations into whether Facebook compromised the personal information of up to 87 million of its users when it allowed data analytics company Cambridge Analytica to gain access and exploit the information.

- A multi-billion-dollar advertiser boycott against Google after revelations that the company allowed Jihad, hate speech, and other objectionable content on YouTube.

- A federal investigation into whether Russians used digital platforms, in particular Facebook, to spread disinformation in an effort to influence the 2016 election.

- After years of being Washington political darlings, the early lobbying efforts of Google and Facebook against anti-sex trafficking legislation were rejected by Congress—and now, for the first time, these digital platforms are being held responsible for the content they host on their sites.

But the potential damage goes well beyond what policymakers and advertisers think. It goes to the heart and soul of digital platforms: their users. While these companies have profited, they have alienated their users and significantly harmed trust in the platforms. According to a new Digital Citizens survey conducted March 24-March 30, faith in the platforms seems at an all-time low:

- Seventy-one percent report that over the last year their trust in the platforms has dropped.

- A majority of Americans (51 percent) said that platforms such as Facebook, Google, and Twitter are not responsible companies "because they put making profits most of the time ahead of trying to do the right thing." Only 20 percent said that they are "responsible companies because they try to do the right thing most of the time even if that gets in the way of it making profits."

- Just over half (50 percent) now believe that these digital platforms should be regulated. Only 1 in 4 (25 percent) said they should not be regulated.

- By a 65 percent-21 percent margin Americans said that companies such as Google should take a more active role in monitoring and taking down inappropriate content on their own instead of relying on users to flag it.

- Fifty-four percent said that companies such as Facebook, Google, and Twitter brought their recent problems on themselves by not doing a good enough job policing their content. That's compared to 26 percent who said the problems were outside their control.

A MAJORITY OF AMERICANS NOW SAY THAT DIGITAL PLATFORMS ARE **NOT RESPONSIBLE** *COMPANIES AND SHOULD BE* **REGULATED.**

Up until now, digital platforms either have not appreciated the gravity of their challenges or had the willingness to own up to what needs to be done. In fact, for years they have resisted nearly every effort to take these issues more seriously. Perhaps that was because they were media darlings and the cash was flowing (Facebook's advertising skyrocketed 500 percent to $39.9 billion in four years).

But now that they face the threat of regulation, advertiser backlash and—for them, the worst news—users signing off, perhaps they will make this a priority. One thing we know about technology companies: they have some of the most brilliant minds that can achieve anything. When they want to.

DIGITAL DISTRUST: HOW DID WE GET HERE?

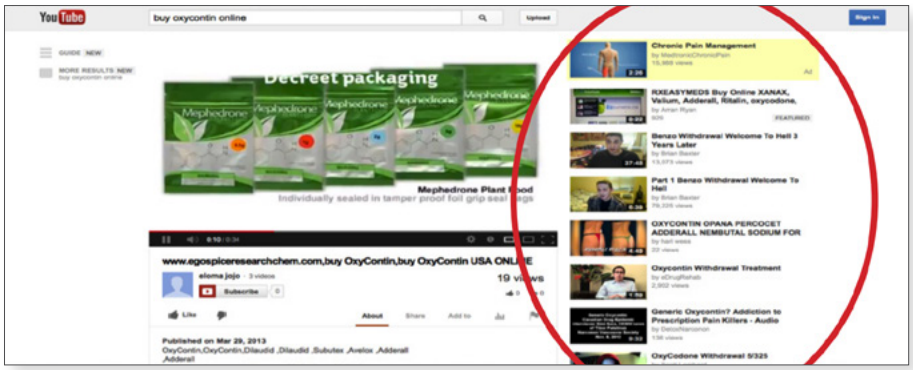
GOOGLING INDIFFERENCE TO HARM

We got here because when digital platforms such as Facebook and Google let it be known that they would be lax about policing their sites, it sent a signal to criminals and other bad actors. It's no different from a street corner. If the authorities ignore a drug dealer or someone peddling stolen merchandise, the criminal feels confident to operate freely.

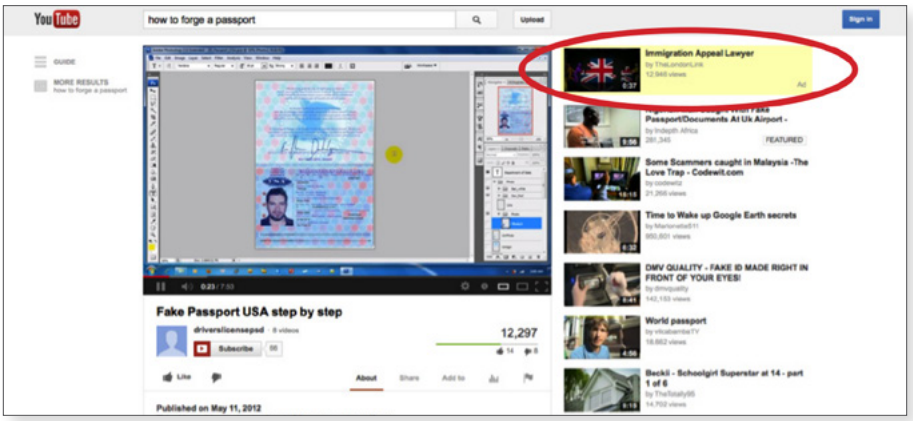
Criminals and bad actors seized upon Facebook, Google, and YouTube because they knew that the risks and costs of doing business are both low. Digital Citizens has seen it up close and issued warnings. Google shrugged and ignored the warnings.

Five years ago, Digital Citizens raised alarms about the proliferation of illegal and objectionable content on Google platforms such as YouTube. DCA cited examples of the platform being used to sell and promote illegal narcotics, prescription drugs without a valid prescription, knock-off merchandise, and fake IDs, including driver's licenses and passports.

Just as troubling, Google sold ads on YouTube videos promoting things like drugs, prostitution, and forged documents, it was effectively an advertising partner with bad actors because when YouTube users click on those ads, Google's business model is to split the ad revenue with those video producers. Here are just a couple of examples, from the thousands found, of inappropriate content that Digital Citizens discovered in 2013 on YouTube:



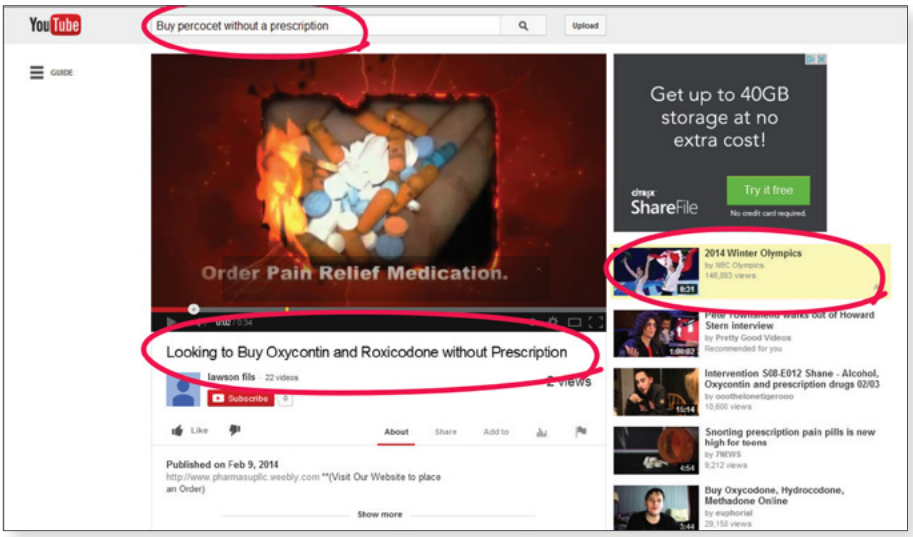
Digital Citizens researchers found countless YouTube videos for sites marketing the sale of prescription pain medicines to individuals without the proper prescription. This video in this screenshot is from a site marketing the sale of the opioid pain medication, OxyContin, with a paid advertisement for chronic pain management.



This screenshot shows an ad for an "Immigration Appeal Lawyer" right next to a video marketing the creation and sale of fake U.S. Passports.

Google's response was telling, stating that its "review teams respond to videos flagged for our attention around the clock, removing any content that violates our policies. We take user safety seriously and have Community Guidelines that prohibit any content encouraging dangerous, illegal activities. This includes content promoting the sale of drugs."

By pointing out that it responds to content "flagged for our attention," Google was acknowledging that it wouldn't monitor itself. Google took down the videos that Digital Citizens highlighted, but because the search and advertising giant treated it as a PR problem, AND NOT an Internet safety or user trust issue, within months the inappropriate content was back.



Here's a YouTube video promoting the sale of OxyContin and Roxicodone without a prescription. Note the 2014 Winter Olympics ad on the right.

Once again, Google removed the incriminating content, and declared how many videos it takes down a day. Months later, a DCA review found stolen credit cards for sale on YouTube. Along with ABC News, Digital Citizens contacted credit card thieves who demonstrated how they can be used and even how criminals can make their own cards using fake names.

Digital Citizens' researchers found this Target advertisement running alongside a video pushing stolen credit cards, social security numbers, and bank logins on YouTube. When we clicked on the ad, we went directly to Target's website. At that time, the company was spending billions to regain trust on the heels of its massive December, 2013 data breach.

These two screenshots are among dozens of examples that Digital Citizens captured that are available at http://www.digitalcitizensalliance.org/get-informed/digital-citizens-investigative-reports/.

Google's approach seemed influenced by an alarm that would have served, for most companies, as a wake-up call: a sting operation organized by a Rhode Island prosecutor found that the company had illegally helped overseas pharmacies illegal promote the sale of prescription drugs in the United States.

An investigation using a federal prisoner posing as the operator of illegal online pharmacies showed that Google was not only aware that Canadian pharmacies that advertised on its site were providing painkillers such as Oxycontin—one of the drugs fueling the opioid crisis—without a prescription but aided them in developing advertising.

State prosecutors hinted the knowledge of the scheme was among the top executive echelons of Google, and in 2011 the company settled by agreeing to pay $500 million.

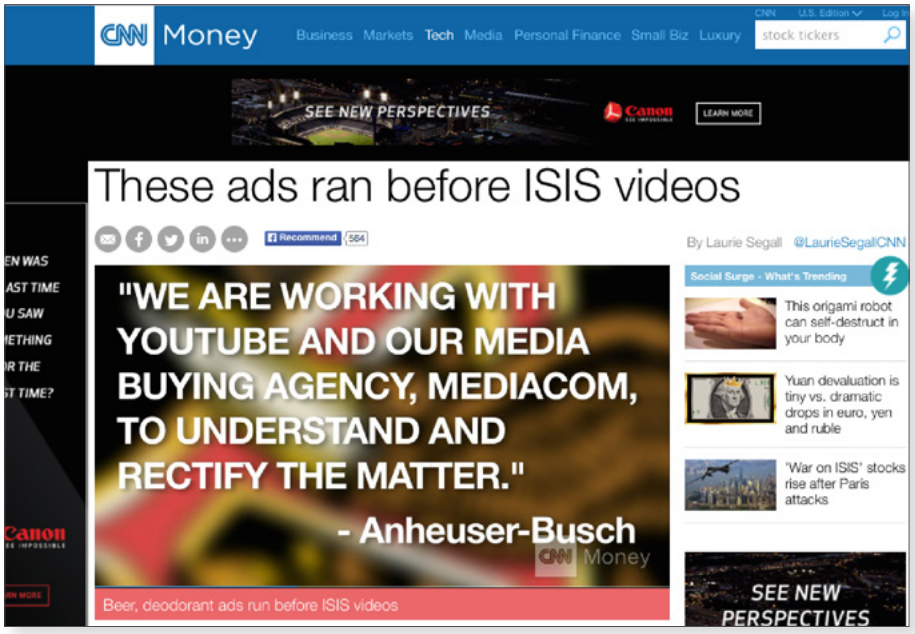
As Google took a hands-off approach to monitoring its content, it also began to weaken its policies that protected its users' privacy. When the company first published its privacy policy in 1999, it was less than one page and stated that Google would not disclose identifiable information to any third party without receiving the customer's permission and that beyond the initial search and result click, Google did not track a user and the user's data.

But by 2010, Google was apologizing for inappropriately collecting computer passwords and emails and downloading personal information from wireless networks as part of its Street View project. By 2012, Google was tracking users and user data across YouTube, Gmail and its search engine and combining that data to create user profiles. Opting out of this tracking was not an option, the company said.

While some change in settings was inevitable and necessary as new location-based services are offered, Google hasn't been very transparent about it. For example, in 2017 Google acknowledged that its Android phones collected cell tower information that enabled the company to track individuals' locations and movements even when their devices are off.

For a timeline of Google's privacy policy changes, please visit: http://www.digitalcitizensalliance.org/google-privacy.

In the end, Google did move to clean up inappropriate and objectionable content in 2017, but not when warned by Internet safety groups or policymakers—only when it faced a billion-dollar backlash from advertisers who said they would no longer advertise on the platforms.



Anheuser-Busch's response to a CNN report of advertisements for its products running on YouTube ahead of ISIS videos.

A year after Google ignored Digital Citizens' warnings that Jihad and other hate speech videos were not only proliferating on the platform but compromising mainstream brands—companies such as AT&T, Verizon, Pepsi, Walmart, Dish, Budweiser, Starbucks, and General Motors—pulled their advertising. Only then, faced with the loss of profits, did Google vow to aggressively police "hateful, offensive and derogatory content."

Digital Citizens' researchers found this RAM truck advertisement running alongside this ISIS related content on YouTube, which was viewed over 33,000 times.

FACEBOOK UNLIKED

If anything, Facebook has faced even more challenges with user privacy but typically sidestepped much criticism because the company was more open about its troubles. As an engineering-driven company initially, Facebook didn't have a good antenna for user issues.

A decade ago, Facebook rolled out Beacon, an ambitious advertising platform that allowed other companies to track purchases by users and then, without consent, alert the users' friends of what they had bought. Facing a backlash, the company backtracked and allowed its users to opt out. Again, this time in 2011, Facebook settled with the FTC over claims it gave third-party apps access to users' personal data without their permission—a case not altogether different from the current controversy surrounding Cambridge Analytica's access to user data.

Facebook now faces a follow-up FTC investigation into whether the Cambridge Analytica data dump violated the 2011 agreement to not share user information without permission. If found in violation, Facebook could be hit with a significant fine, new restrictions, and additional scrutiny and monitoring. As investigators dig into the case they will learn more how about Cambridge Analytica proposed using sensitive data from Facebook to target Google search ads.

While Cambridge Analytica may be the straw that breaks the camel's back, Facebook's trust problems started when its hands-off approach to content was exploited by Russian and partisan interests trying to game the 2016 election.

Well before the latest controversy, Facebook had to know that it was in potential trouble with its users. According to a Digital Citizens survey in early 2017, 64 percent of Americans said their trust in digital platforms had dropped in the last year. The same amount said that that the Fake News issue had made them less likely to trust the Internet as a source of information.

To its credit, Facebook committed to hiring thousands of monitors that would police content. But a year later, Facebook is heading into a tense 2018 U.S. election with many of the same issues that plagued it in 2016: a proliferation of divisive and misleading content and uncertainty whether users are real—or bots or sleepers to distract and disrupt.

TWITTER'S CHARACTER

To quantify the impact digital platforms have on politics and society, all one has to do is look at the Tweeter-in-Chief in the White House.

While it has not received the same level of scrutiny as Facebook, Twitter is perhaps the most blatantly utilized digital platform to spread disinformation.

Within hours of the Parkland school shooting, thousands of bots—many of them connected to Russia—were activated to flood Twitter with propaganda to stir up emotions. According to Botcheck.me, a website that tracks 1,500 political propaganda bots, in the aftermath of the shooting those bots began tweeting exclusively about that event. The top hashtags included #Parkland, #guncontrol, and #guncontrolnow.

While it offers a different type of service than Google or Facebook, Twitter has one thing in common with them: criticism that it hasn't been transparent. Twitter initially downplayed the number of Russian-linked accounts it found (174)—only to update that to 3,000 months later, sparking criticism. Sen. Mark Warner labeled Twitter's Capitol Hill testimony as "inadequate on almost every level" and later said they were the least responsive of all the tech companies.

Even when it knows its users, Twitter is challenged to keep inappropriate content off its platform. In a 2013 article entitled, "Twitter: The New Face of Crime," *USA Today* demonstrated how "political extremists, criminals and gang members are advertising their wares, flaunting their exploits and recruiting new members in 140 characters or less."

The challenge for Twitter is the same as Facebook: in just over seven months, the U.S. election will take place and the companies haven't inspired confidence that they have a handle on false information and the integrity of their users.

CREATING SAFE DIGITAL NEIGHBORHOODS

We all want to live, work and play in a safe neighborhood. A core element of a safe neighborhood is trust—and the platforms are rapidly losing it. In the recent research survey 57 percent of Americans said that Facebook "is an unsafe neighborhood."

Facebook and other digital platforms have a tough road ahead to rebuild trust among their users, companies, and advertisers that rely on them, and policymakers who need to be convinced that the platforms will be part of the solution, not the problem.

It starts with a simple act: an honest conversation with the American people about how these platforms have been used and misused. Full disclosure of any other incidents and, finally, a clear declaration of responsibility.

Government regulation is already being implemented in Europe with new requirements that platforms remove illegal and objectionable content within one hour of notification. It's now seriously being discussed in the United States as well. And for the first time since Digital Citizens began tracking perceptions of digital platforms, a majority of Americans believe that regulation is necessary.

At this point it's moved beyond digital platforms merely avoiding onerous regulation or managing investigations. Regaining trust is at stake. Since mid-March, the number of Americans who say that Facebook is an "irresponsible company" has jumped from 35 percent to 52 percent.

TO REGAIN TRUST, DIGITAL PLATFORMS HAVE TO COMMIT TO TAKING **RESPONSIBILITY** *FOR THE CONTENT THAT APPEARS ON THEIR SITES.*

Digital platforms have to reassure a skeptical user base that they can properly manage their personal information and police the content that appears on their sites. If not, #deletefacebook will grow as a movement and likely spread to other digital platforms.

To avoid that, the companies have to own up to the fact that the aggressively hands-off approach they took to policing content made it easy for criminals and other bad actors to exploit the platforms, which in turn has blurred the lines between mainstream sites and the "Dark Web."

To show that they are serious about regaining trust, Facebook, Google, and Twitter have to make real changes. This includes the hiring of a more diverse multi-cultural workforce dedicated to identifying inappropriate content and illegal activities and then removing them. Digital Citizens has long noted that Google's technology enables it to place relevant ads even on inappropriate content. Surely that algorithm could be deployed to flag suspicious content for inspection.

Second, there should be a cross-platform initiative to identify and ban bad actors. This is something Digital Citizens has long advocated for while acknowledging it will pose technical and legal challenges. This could include analyzing usage data that they already collect to highlight behavior that is anomalous and suggests illicit, unlawful or illegal conduct.

Platforms could create digital fingerprints of unlawful conduct that are shared across platforms to proactively block such conduct, as is done with child pornography. There is also the model used by casinos to identify cheats and share that information globally.

With this information, digital platforms would have the capability to make decisions whether to de-list or demote websites offering illicit goods and services, and the ability to stop the spread of illegal behavior that victimizes its users.

Given rising privacy concerns, digital platforms should collaborate to create uniform basic privacy settings that are easily understood by users. Internet users are generally at a loss at how their information is collected and disseminated.

Finally, digital platforms should commit to not using their dominant position to harm would-be competitors, because as these companies grow in size that is inevitably the next major concern for policymakers and regulators.

Over the last year, more often when you hear about Facebook, Google, or Twitter, it's in the context of Russian election meddling, privacy breaches, or illegal or illicit behavior. This behavior has had an impact on users' trust and is certainly not the reputation by which these digital platforms wish to be known. Hopefully these powerful companies understand that lawyers and lobbyists can do a lot to protect them, they can do nothing to regain users' trust.

THE SURVEY

The survey of 1,020 Americans included in the "Digital Platforms in Crisis" report was conducted by Survey Monkey from March 24, 2018–March 30, 2018 and has a margin of error of +/- 3 percent.

ABOUT DIGITAL CITIZENS

This report was created by the Digital Citizens Alliance, a nonprofit 501(c)(6) organization that is a consumer-oriented coalition focused on educating the public and policymakers on the threats that consumers face on the Internet and the importance for Internet stakeholders—individuals, government, and industry—to make the Web a safer place. While all Digital Citizens hold themselves personally responsible to do all they can to protect themselves and their families, we are also concerned that technologies, standards, and practices are in place that will help keep all of us safe as a community. The industry has a critical role to play in ensuring those safeguards are established and updated as needed to address the continually evolving challenges we face online. We have much work to do, but we can't do it effectively without understanding the problems we face. That is why the Digital Citizens Alliance investigates issues such as those detailed in this report. By sharing our findings with consumers, we hope all Digital Citizens will engage in discussions about these issues.