



\ Insights Lab \

# The Good Censor

*How can Google reassure the world that it protects users from harmful content while still supporting free speech?*

CULTURAL CONTEXT REPORT - MARCH 2018

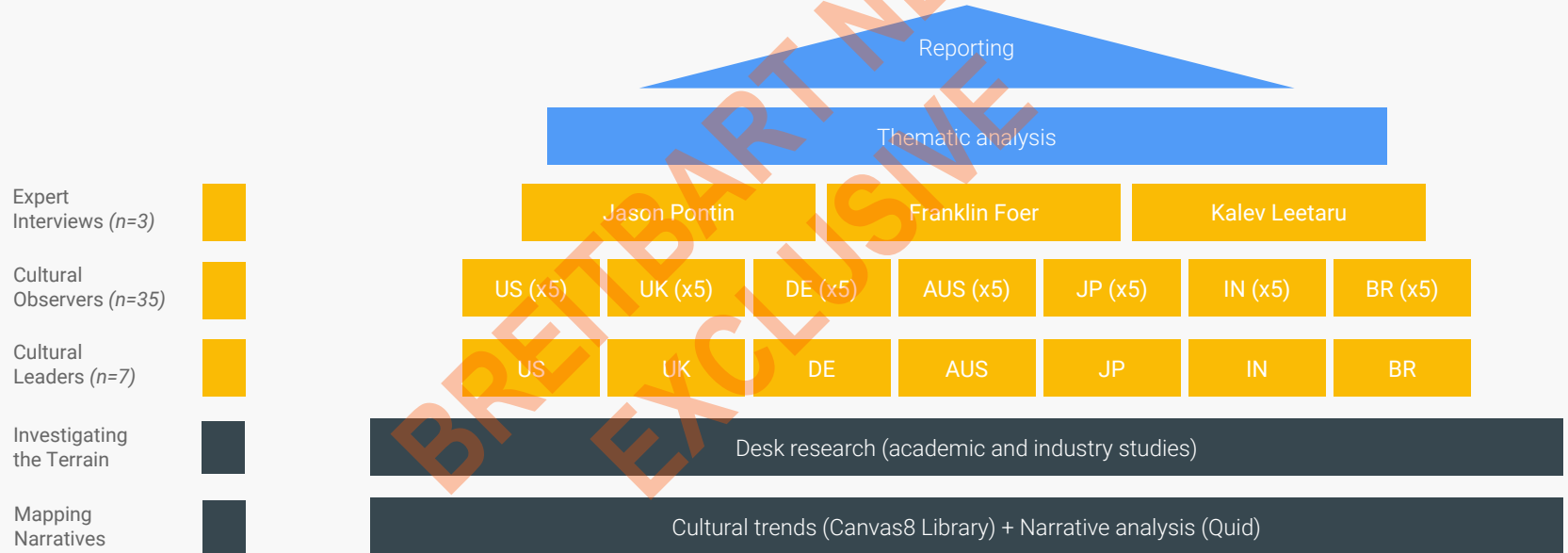
# Can Google protect free-speech *and* police harmful content?

From elections and political propaganda, trolls and gendered bigotry, to hate speech and religious extremism, **debates about who can and should be heard on the internet rage like never before.** As governments struggle to apply existing legislation to the Wild West online, **users are asking if the openness of the internet should be celebrated after all.**

Bots and troll farms lash out at free thought and controversial opinion, while faceless users attack each other without empathy. **Free speech becomes a social, economic and political weapon.** Automated technologies lack the sophistication to adjudicate effectively. In response, people think twice before airing their thoughts aloud, while critique is buried under avalanches of automated rebuttals, vitriolic attacks and nonsensical rhetoric. **As the tech firms struggle to deal with the issues, the public and governments grow increasingly impatient.**

Yet, amongst all this negativity, seeds of political harmony, gender and racial equality, and tolerance are sown on the internet. **Is it possible to have an open and inclusive internet while simultaneously limiting political oppression and despotism, hate, violence and harassment? Who should be responsible for censoring 'unwanted' conversation, anyway? Governments? Users? Google?**

This report is the result of several layers of research



We worked with cultural leaders and local observers to deepen our understanding...

## Local cultural leaders

### **Dr Peter Chen**

Academic, Australia

### **Bia Granja**

Entrepreneur, Brazil

### **Joana Breidenbach**

Anthropologist, Germany

### **Nobuyuki Hayashi**

Journalist, Japan

### **Grant McCracken**

Anthropologist, USA

### **Nikhil Pahwa**

Entrepreneur, India

### **Richard Watson**

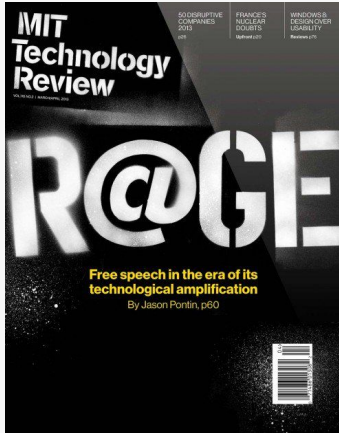
Futurist, UK

### **21x Micro Observers**

(3x per market)

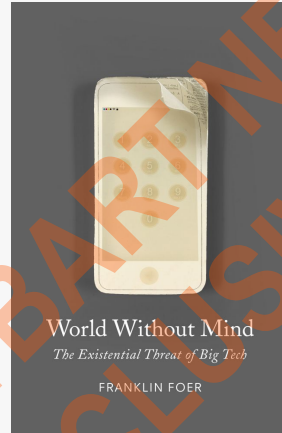


We've worked with some leading thinkers in this space



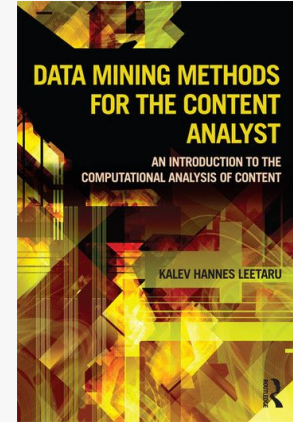
### Jason Pontin

Journalist and former editor  
in chief of MIT Technology  
Review



### Franklin Foer

Author of *World Without  
Mind* and former editor of  
The New Republic



### Dr. Kalev Leetaru

Senior Fellow at the George  
Washington University  
Center for Cyber &  
Homeland Security

1. The importance of free speech
2. Bad behaviour and censorship online
3. Fighting back against bad behaviour online
4. Tech firms - free speech or censorship?
5. How do people want Google to respond to this predicament?

With free speech, individuals can hold 'the powerful' to account...

BREITBART NEWS  
EXCLUSIVE

The freedom to speak holds the key to our two most valuable possessions...

**Personal liberty**

And the ability to express  
yourself freely

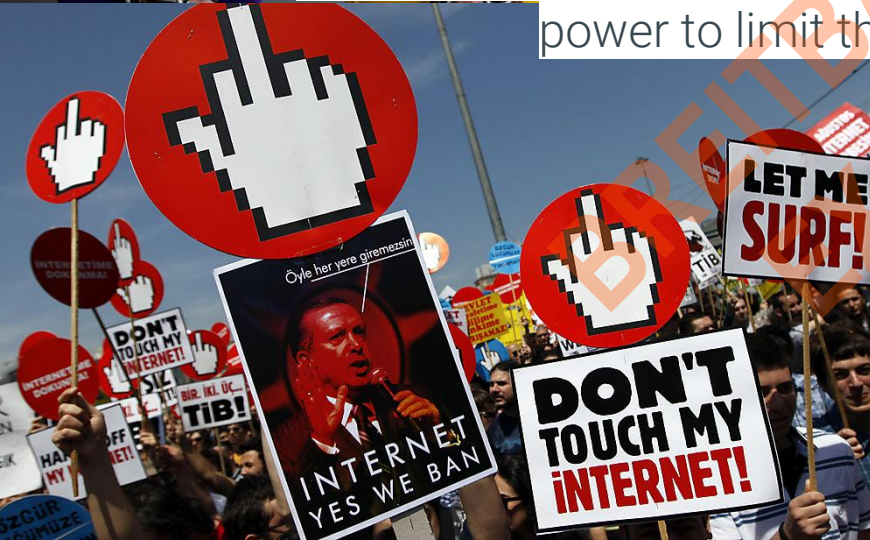


**Collective wellbeing**

And the prevention of harm



... but censorship can give governments – and companies – the power to limit the freedom of individuals



Underpinned by the First Amendment and belief in an “equality of status in the field of ideas” - the US is especially committed to free speech<sup>1</sup>

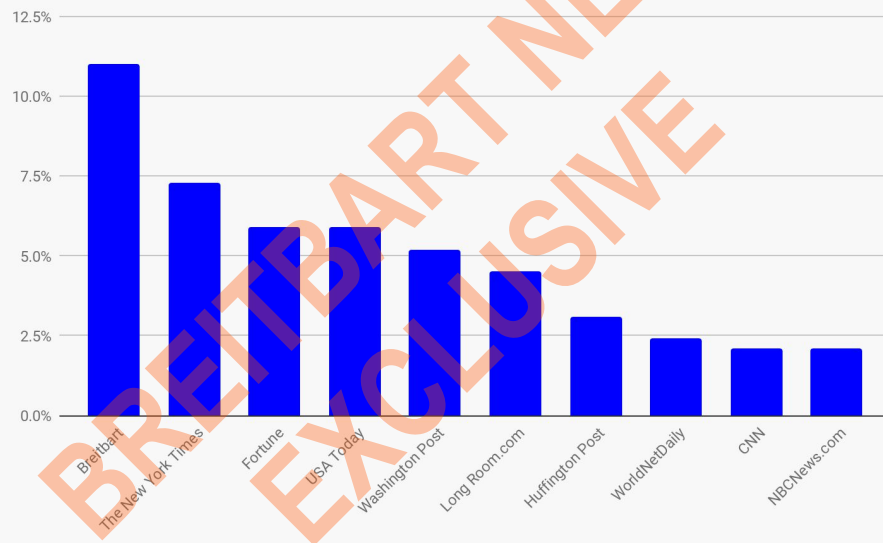
“

The reason that we only [ban] speech when it's intended to and likely to cause imminent violence is because as long as there's time enough to deliberate and to discuss— [there's faith that] **the best remedy to evil counsels is good ones**, that **counter-speech is more appropriate than suppression**, and that **reason will ultimately prevail**... it's the essence of our constitutional system.

Jeffrey Rosen, 2016

[The Deciders: The Future of Free Speech in a Digital World, Harvard Kennedy School](#)<sup>2</sup>

And conversations about the importance of free speech are alive on both sides of the political spectrum



Source: Quid with Canvas8 analysis. Media analysis of conversations around "free speech" (Google, Apple, Facebook, Amazon), Sep to Dec 2018. Focus on US & UK, mainstream media.

The internet was also founded on utopian principles of free speech...

“

*Governments of the Industrial World, you weary giants of flesh and steel, I come from Cyberspace, the new home of Mind. On behalf of the future, I ask you of the past to leave us alone. You are not welcome among us. You have no sovereignty where we gather.*

John Perry Barlow, 1996

A Declaration of the Independence of Cyberspace

This free speech ideal was instilled in the DNA of the Silicon Valley startups that now control the majority of our online conversations...



"[Google's] atmosphere of creativity and challenge... has helped us **provide unbiased, accurate and free access to information** for those who rely on us around the world."

Larry Page and Sergey Brin  
2004 Founders' IPO Letter<sup>1</sup>



"[Facebook is a tool to create] a **more honest and transparent dialogue around government**. [The result will be] better solutions to some of the biggest problems of our time."

Mark Zuckerberg  
2012 manifesto for investors<sup>2</sup>



"[Twitter is] the **free speech wing of the free speech party**"

CEO Dick Costolo  
2017<sup>3</sup>

An important US Federal statute from 1996 supports this position of neutrality

Under section 230 of the Communications Decency Act, **tech firms have legal immunity from the majority of the content posted on their platforms** (unlike 'traditional' media publications).

This protection has empowered YouTube, Facebook, Twitter and Reddit to **create spaces for free speech without the fear of legal action** or its financial consequences.

"It's hard to say what the global internet would look like if Section 230 had never become the law of the land.

**Would YouTube have even been possible?"**

April Glaser, [Slate](#)



And the internet has  
certain, unique qualities  
that have supported  
these ambitions further...

Communication is **fast and frictionless**

**Anonymous** conflict is possible

Everyone has a **voice**

We meet **like-minded** people

**Scale** is unprecedented

This commitment to free, uncensored conversation has had positive outcomes...

The Arab Spring was the the high point of this positivity - a visceral example of the power of digitalised free speech.



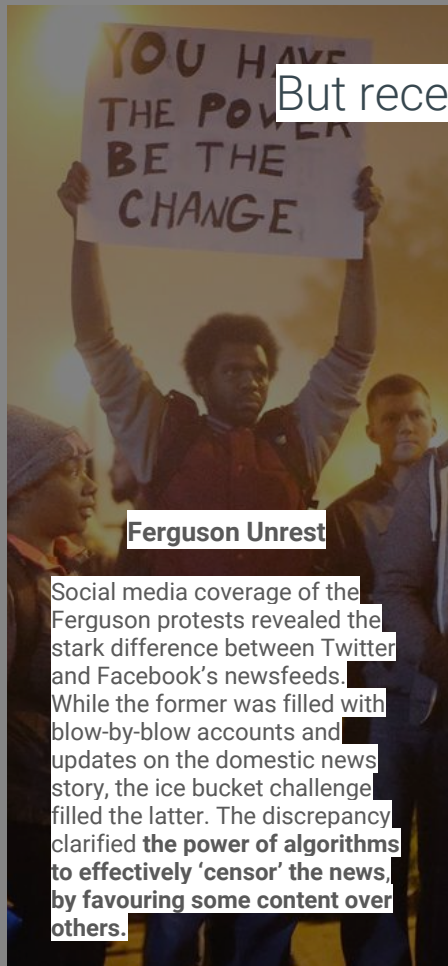
Free speech flourished online as governments struggled to contain it

“

*On the global scale, **the internet and the social platforms have been a wonderful boon for free speech.** The internet has given platforms to billion of people to express themselves and has made it **almost impossible for governments** – even in highly controlled nations like China – **to control people's speech effectively.***

Jason Pontin

But recent global events have undermined this utopian narrative



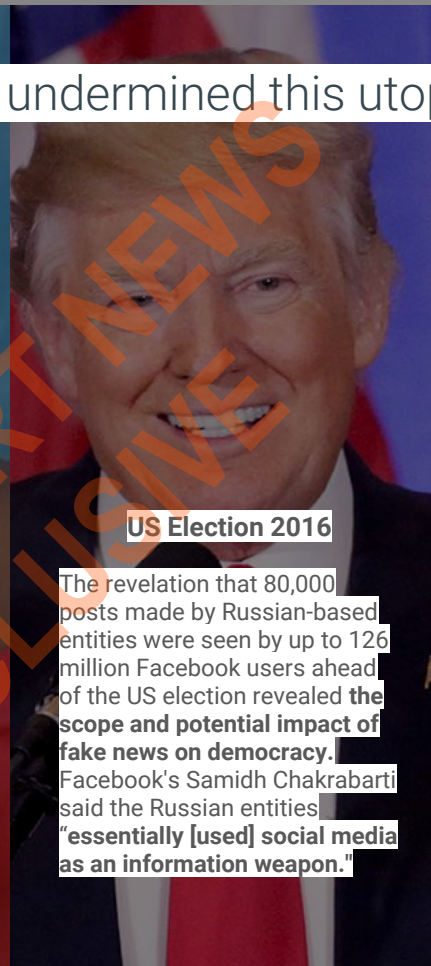
Ferguson Unrest

Social media coverage of the Ferguson protests revealed the stark difference between Twitter and Facebook's newsfeeds. While the former was filled with blow-by-blow accounts and updates on the domestic news story, the ice bucket challenge filled the latter. The discrepancy clarified the power of algorithms to effectively 'censor' the news, by favouring some content over others.



Leslie Jones vs Trolls

Actor Leslie Jones was subjected to **persistent sexist and racist trolling on Twitter**. After Jones quit the platform there was extensive media coverage and public outcry, and alt-right ringleader Milo Yiannopoulos was thrown off. He responded: **"This is the end for Twitter. Anyone who cares about free speech has been sent a clear message: 'You're not welcome on Twitter.'"**



US Election 2016

The revelation that 80,000 posts made by Russian-based entities were seen by up to 126 million Facebook users ahead of the US election revealed **the scope and potential impact of fake news on democracy**. Facebook's Samidh Chakrabarti said the Russian entities **"essentially [used] social media as an information weapon."**



Kashmir Clashes (IN)

Facebook and Twitter were implicated in **governmental censorship of clashes between rebels and Indian authorities in Kashmir**. The platforms removed posts and suspended accounts about the events, including images of rebel Burhan Wani's funeral, highlighting the **platforms' complicity with government censorship** as they attempted to stay on the right side of global authorities.

But recent global events have undermined this utopian narrative

### Philando Castile

The aftermath of the shooting of Philando Castile was broadcast on Facebook Live. Where traditional media would have had time to consider how to broadcast such sensitive, violent and controversial content, live-streaming sidesteps this editorial process. The clip highlights the huge importance of context in moderation and the fluctuating line of appropriate censorship.

### The Rise of the Alt-Right

The rise of far-right political parties and institutions such as Britain First, Germany's AfD and Unite the Right opened people's eyes to how alt-right beliefs have been able to flourish on the internet. **Once controversial voices have been emboldened by like-minded individuals and are making their way offline, both on the streets and at the polls.**

### Queermuseu (BR)

The conservative *Free Brazil Movement* used social media platforms to rally against an art exhibition called Queermuseu, because it discussed homosexuality and paedophilia. The exhibit was eventually shut down, raising concerns about **freedom of expression in digital spaces and the censorship of online/offline spaces in Brazil.**

### Logan Paul

Hugely popular Youtuber, Logan Paul drew criticism for an insensitive clip of him seeing a suicide victim in Aokigahara forest, Japan. Youtube responded by removing Paul from its premium advertising program and reforming its ad restrictions. **The controversy raised the question of how much censorship we should demand from Youtube and whether it is putting profit before people.**

While revelations and exposés increased calls for change

### Extremist Content

**Major brands**, including the UK government, Marks & Spencer and McDonald's, **boycotted YouTube** after it was revealed that their ads were appearing on controversial clips and extremist content. Google responded with promises of an **overhaul to advertising policies**, including more control and transparency for advertisers, but fell short of promising to rid the platform of such content entirely, as the latter would place them squarely in the realm of 'curator and censor'.

### Peppa Parodies

Articles by the New York Times and writer James Bridle called attention to **troubling and inappropriate video content on YouTube**, which is not only accessible to children but often **targeted at them using popular kids' characters** like Peppa Pig, Frozen's Elsa and Spiderman, and tags to game the platform's algorithms. The resulting outcry led to **promises of reform from YouTube** and revealed the **shortcomings of relying on algorithmic filtering**.

As the “we’re not responsible for what happens on our platforms” defence crumbles, users and advertisers are demanding action...

“ We're through the first early utopian period of social media and free speech, we're through the middle period, where there was excitement about the benefits of these platforms, and now **we're into a third era where we've become more jaded about their functionality**. Now **we're looking to the networks themselves to better manage their own utility**, and there is **conversation about governmental obligations** on these networks as well.

– Jason Pontin

1. The importance of free speech
2. Bad behaviour and censorship online
3. Fighting back against bad behaviour online
4. Tech firms - free speech or censorship?
5. How do people want Google to respond to this predicament?

What's driving this furore around free speech and censorship online?

BREITBART NEWS  
EXCLUSIVE

“

The early utopian period of the internet has collapsed under the weight of bad behaviour....

Jason Pontin

Users, Governments  
& Tech firms  
are all behaving badly...

BREITBART NEWS  
EXCLUSIVE

# How are users behaving badly?

## Hate speech

Although people have long been racist, sexist and hateful in many other ways, they weren't empowered by the internet to recklessly express their views with abandon. From film-stars to activists, viciousness is aimed at a diverse range of users.

**2.6 million**

tweets contained anti-Semitic speech during the US presidential election<sup>1</sup>

## Reprisals and intimidation

Online and offline worlds are blurring as more people are physically assaulted for speaking on the net. In 2017, a Christian cartoonist was murdered in Jordan for mocking Islamist militants' vision of heaven, while a journalist was killed in Myanmar for using FaceBook to post about corruption.<sup>2</sup>

**8 countries**

witnessed murderous reprisals for online speech in 2017<sup>3</sup>

## Trolling

By provoking arguments and flaming disruption, trolls threaten valuable debate and infuriate users. The problem has become so rampant that several websites have even resorted to removing comments entirely.

**26%**

of American users are victims of internet trolling<sup>4</sup>

# How are users behaving badly?

## Cyber harassment

From petty name-calling to more threatening behavior, harassment is an unwelcome component of life online for all too many users. With sustained stalking and one-off incidents defining the spectrum, some experiences are easier to escape than others.

**40%**  
of internet users  
have been harassed online<sup>1</sup>

## Cyber racism

Supremacy, destiny and nationalism. Otherness, separation and hostility. Cyber racism exists in many guises, but it most often describes a "range of white supremacist movements in Europe and North America" and "the new horizons the Internet and digital media have opened" for them.<sup>2</sup>

"We're stepping off the internet in a big way. We have been spreading our memes. We have been organising on the internet. And now we're coming out."

Robert Ray, Daily Stormer (neo-Nazi website), at the Charlottesville protests<sup>3</sup>

## Venting

When they're angry, people vent their frustrations. But whereas people used to tell friends and family about bad experiences, the internet now provides a limitless audience for our gripes. As more and more people vent, online conversations fill with anger and nastiness.

"Just pick a random, poor, innocent idiot on the internet and just attack them. Go after them. And find others to join you, who are also angry at that moment."

Kalev Leetaru

## Why are users behaving badly?

“

**Human beings en masse don't behave very well.** They particularly don't behave very well if there aren't clear rules, and especially if speech is unaccountable, consequence-free, and in many cases anonymous. What happened on these networks is that, **in the absence of rules and consequences, everyone has behaved maximally badly.**

– Jason Pontin

A man with dark hair, wearing a dark blue suit, a light blue and white striped shirt, and an orange tie, is speaking. He has a serious expression and is gesturing with both hands. The background is a blurred red curtain. A large, diagonal, semi-transparent orange watermark reading "BREXIT NEWS EXCLUSIVE" is overlaid across the center of the image.

And the same reasons why the internet is great for free speech  
mean it's also primed for bad behaviour...

## On the net...

Communication is fast and frictionless

Anonymous conflict is possible

Everyone has a voice

We meet like-minded people

Scale is unprecedented

Why are users  
behaving badly?

which means...

Relentless, 24/7 online conversations encourage people to **dive-in with their opinion** before it's too late, even if they're misinformed. And because we think with our emotional brain before our rational one, instant responses **amplify emotion-led discourse** not thoughtful debate.

The **social norms** that hold society together and keep people from hurting one another offline, **shift faster online**. It's more tempting to be nasty and aggressive when there are **no warning signals or hurdles** to slow people down. And because the internet removes physical communication barriers, **users are detached from the effects of their actions**

## On the net...

Communication is fast and frictionless

Conflict can be anonymous

Everyone has a voice

We meet like-minded people

Scale is unprecedented

Why are users  
behaving badly?

which means...

When they can't be seen or found, **people are more likely to cheat, lie and attack each other**. Anonymity isn't the full story - group dynamics, online cultures, and even the time of day can encourage bad behaviour - but the opportunity to **behave badly without fear of repercussion** does bring out the worst in people. **This is especially true when we aren't forced to empathize**

Offline, we avoid confrontation because it can more easily lead to physical harm - but with online anonymity, people don't worry so much. When we think nobody can see us, **we're keener to transgress moral norms**. And the ability to have multiple identities enables people to **say one thing and do another**.

## On the net...

Communication is fast and frictionless

Anonymous conflict is possible

Everyone has a voice

We meet like-minded people

Scale is unprecedented

Why are users  
behaving badly?

which means...

The **'little guys and girls' can now be heard** - emerging talent, revolutionaries, whistleblowers and campaigners. But **'everyone else' can shout loudly too** - including terrorists, racists, misogynists and oppressors. And because "everything looks like the New York Times" on the net, it's **harder to separate fact from fiction**, legitimacy from illegitimacy, novelty from history, and positivity from destructivity.

When consumers/producers feel like they 'own' their media platforms, their **experiences of free speech and censorship feel more personal** too. They increasingly value their ability to speak freely, but also feel personally assaulted when confronted through their own channels, **lashing out more violently** when their voice and opinions are threatened.

## On the net...

Communication is fast and frictionless

Anonymous conflict is possible

Everyone has a voice

**We meet like-minded people**

Scale is unprecedented

Why are **users**  
behaving badly?

which means...

The internet has **united political activists, dissidents and like-minded communities** of all shapes and sizes, including the oppressed minorities. On the flip-side, **minority groups once pushed underground** by public opinion of their abhorrent views **have discovered a safer space** in which to communicate, organise and reach-out to new sympathizers.

Because the **internet helps people to bunker down**, surrounded by similar opinions, mindsets and behaviours, **opportunities for learning and life-changing experiences are threatened**. These closed filter bubbles and echo chambers make **positive and transformative political debate less likely**, not more.

## On the net...

Communication is fast and frictionless

Anonymous conflict is possible

Everyone has a voice

We meet like-minded people

Scale is unprecedented

Why are users  
behaving badly?

which means...

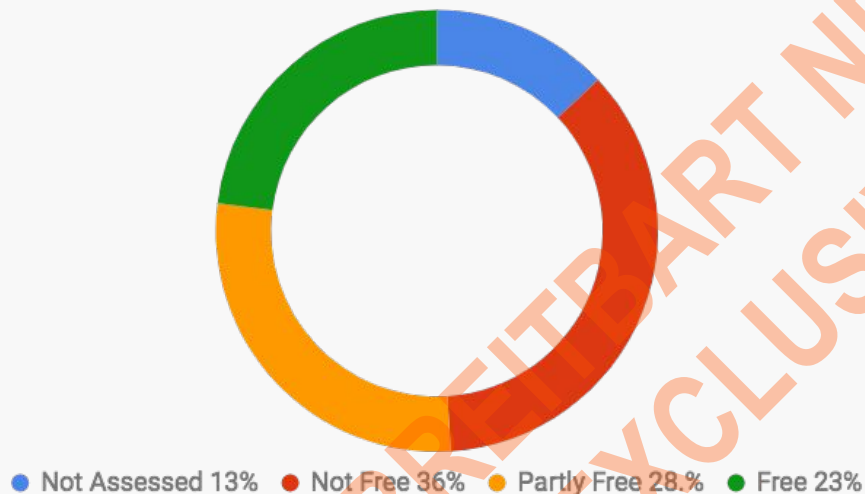
Across the supranational platforms of the net, **local stories become global events. People unite across borders and time zones.** But this global explosion has created a land grab for power.

**Regional laws lose their significance and influence.** Borderless filters aren't relevant everywhere - who decides what is or isn't censored? Jokes and critique don't always translate well. **Crummy politicians jump on the confusion** to expand their influence. **Users' bad behaviour falls between the cracks.**

If that's how users are behaving badly...

What about governments?

According to Freedom House, only a quarter of the world's internet users reside in countries where the internet is 'Free from Censorship' <sup>1</sup>



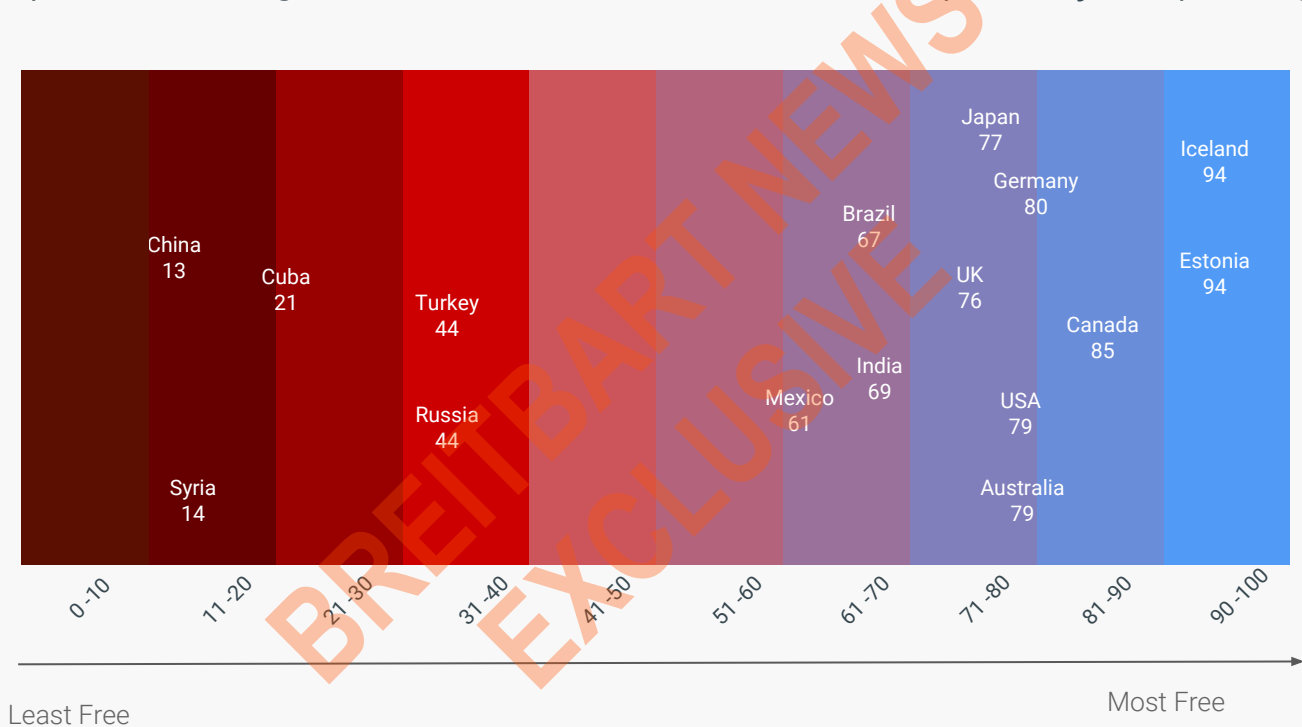
Percentage of total global internet users, by 'freedom of net' status

**'Free from Censorship' means there are...**

No major obstacles to access  
No onerous restrictions on content  
No serious violations of user rights in the form of unchecked surveillance  
No unjust repercussions for legitimate speech.<sup>2</sup>

Freedom on the Net 2017

The spectrum of global internet freedoms isn't especially surprising, but...



Countries are ranked on a 100-point scale based on three broad categories: obstacles to access, limits on content and violations of user rights. The higher the score, the more restrictive a country's internet controls.



Global internet freedoms have gone downhill for the past seven years

## Political interference is on the increase....

Online manipulation and disinformation  
influenced elections in more than 18 countries in  
2017, including the US<sup>1</sup>

Despite having a more vibrant and  
diverse online environment than most,  
disinformation and hyperpartisan content  
are having a bigger impact.

BRETTBART NEWS  
EXCLUSIVE

## And there are worrying signs of new government encroachments

Eg.1

When a Twitter user objected to Trump's immigration policy in January 2017, U.S. Customs and Border Protection agents responded by asking the firm to reveal their identity, before backing off when Twitter fought them in court.

Eg.2

In August 2017, the Department of Justice contacted DreamHost (a hosting company) to demand the names of 1.3 million users who had visited [#DisruptJ20](#) - an anti-Trump protest website.<sup>1</sup>



# How are governments behaving badly?

## Bots

Bots account for more web traffic than humans.<sup>1</sup> Governments employ impersonators, scraps, spammers and hackers to manipulate conversations, quell dissent and discredit information. From Washington to Moscow, bots are deployed by governments against foreign adversaries and domestic opponents.<sup>2</sup>

**28.9%**

of all web traffic is estimated to come from 'bad bots'<sup>1</sup>

## Troll farms

With shadowy secrecy, governments employ armies to manipulate online discussions in their favour.<sup>3</sup> This fabricated support silences opponents and critics at home and abroad. And the propaganda makes it more difficult to know which opinions are 'real' and who is really supportive of those in power.<sup>4</sup>

**448 million  
comments**

posted by users employed by the Chinese government to impersonate ordinary citizens<sup>5</sup>

## Restrictions and cyber attacks

Governments are increasingly restricting mobile internet services for political gain or security reasons.<sup>6</sup> And users were stopped from live streaming anti government protests in more than nine countries in 2017. To limited free speech further, states are also restricting encryption and virtual private networks (VPNs).<sup>7</sup>

**34**

countries witnessed cyber attacks against government critics in 2017<sup>8</sup>

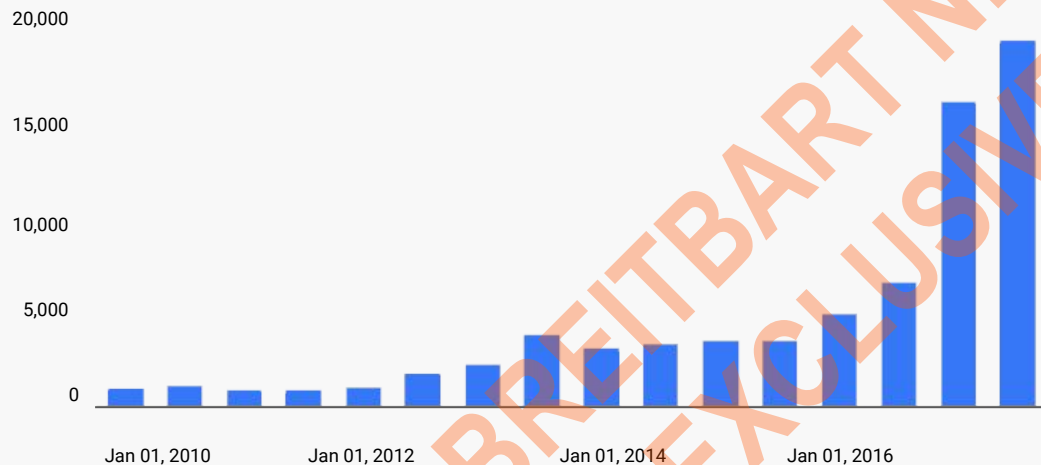
“

The use of paid commentators and political bots to spread government propaganda was pioneered by China and Russia but has now gone global. The effects of these rapidly spreading techniques on democracy and civic activism are potentially devastating.

Michael J. Abramowitz, president of Freedom House<sup>1</sup>

BREITBART NEWS  
EXCLUSIVE

Governments also trying to tighten their grip on political discourse by asking Google to censor more and more content



Total number of content removal requests made to Google by courts and governments worldwide

**50.6%** of these requests relate to YouTube and **19.8%** to Search.

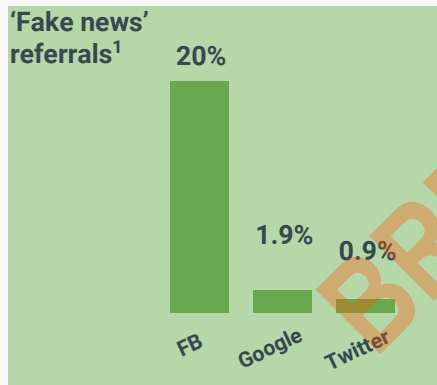
What's the role of the tech firms in all of this?

BREITBART NEWS  
EXCLUSIVE

# How are tech firms behaving badly?

## Incubating fake news

Untrustworthy sources and misinformation have thrived on tech platforms. Dubious distributors have capitalised on a lack of sense-checking and algorithms that reward sensationalist content. And rational debate is damaged when authoritative voices and 'have a go' commentators receive equal weighting.



## Ineffective automation

With 400 hours of video uploaded to YouTube<sup>3</sup> and 340,000 tweets<sup>4</sup> sent every minute, it isn't surprising that platforms outsource moderation duties to AI and automation. But even the most sophisticated tech can censor legitimate and legal videos in error, while erroneous content can elude the safeguards.

**"We got it wrong..."**

Our system sometimes make mistakes in understanding context and nuances

YouTube Creator Blog<sup>2</sup>

## Commercialized conversation

Shares, likes and clickbait headlines - monetized online conversations aren't great news for rational debate. And when tech firms have an eye on their shareholders<sup>5</sup> as well as their free-speech and censorship values, the priorities can get a little muddled.

**"For a business, free speech can only be a meaningful value if it doesn't really cost anything"<sup>6</sup>**

Sarah Jeong, VICE Motherboard