



The Shifting Landscape of Global Internet Censorship

The Harvard community has made this article openly available. [Please share](#) how this access benefits you. Your story matters

Citation	Clark, Justin, Robert Faris, Ryan Morrison-Westphal, Helmi Noman, Casey Tilton, Jonathan Zittrain. 2017. The Shifting Landscape of Global Internet Censorship. Berkman Klein Center for Internet & Society Research Publication.
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:33084425
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

June 2017

The Shifting Landscape of Global Internet Censorship

*An Uptake in Communications Encryption Is Tempered by
Increasing Pressure on Major Platform Providers; Governments
Expand Content Restriction Tactics*

Jonathan Zittrain

Robert Faris

Helmi Noman

Justin Clark

Casey Tilton

Ryan Morrison-Westphal

AUTHORS

Jonathan Zittrain is the George Bemis Professor of International Law and Professor of Computer Science at Harvard University, where he co-founded the Berkman Klein Center for Internet & Society. He has studied ethics and governance of autonomous systems, and he performed the first large-scale testing of Internet filtering beginning in 1999.

Robert Faris is the Research Director at the Berkman Klein Center for Internet & Society at Harvard University. His recent research has been focused on developing and applying methods for studying the networked public sphere.

Helmi Noman is a Research Affiliate of the Berkman Klein Center. His research focuses on Internet censorship in the Middle East and North Africa; exploring the impact of information and communication technologies on the Arab information societies; how the use of the Internet defies the social and political structures; and the potential systemic changes cyberspace can bring to real space in the Arab region.

Justin Clark is a Software Developer at the Berkman Klein Center for Internet & Society at Harvard University. Most recently, he has been adapting, designing and crafting systems for mapping the contours of information control on the Internet.

Casey Tilton is a Project Coordinator at the Berkman Klein Center for Internet & Society. He researches Internet censorship practices around the world and how networked communication technologies affect society.

Ryan Morrison-Westphal is a Senior Web Developer at the Berkman Klein Center for Internet & Society at Harvard University. He utilizes, creates, and promotes open source software, the Open Web, and strong web standards.

ACKNOWLEDGEMENTS

The authors would like to thank the following people for their helpful contributions: Priscilla Guo and Jeanette Si provided research assistance, David Talbot provided valuable comments and suggestions, and local and regional experts in 20+ countries were instrumental in the creation of country-specific URL testing lists. We are deeply grateful for the support provided by the research team at ICLab in giving access to vantage points for data collection.



INTERNET MONITOR is a research project to evaluate, describe, and summarize the means, mechanisms, and extent of Internet content controls and Internet activity around the world.

thenetmonitor.org



INTERNET MONITOR is a project of
the Berkman Klein Center for Internet & Society
at Harvard University
cyber.harvard.edu

23 Everett Street • Second Floor • Cambridge, Massachusetts 02138

+1 617.495.7547 • +1 617.495.7641 (fax) • <http://cyber.harvard.edu> • hello@cyber.harvard.edu

The Shifting Landscape of Global Internet Censorship

An Uptake in Communications Encryption Is Tempered by Increasing Pressure on Major Platform Providers; Governments Expand Content Restriction Tactics

By Jonathan Zittrain, Robert Faris, Helmi Noman, Justin Clark, Casey Tilton, and Ryan Morrison-Westphal

Executive Summary

Since 1999, the Berkman Klein Center has been part of independent academic efforts to enumerate the shaping and filtering of Internet content, and in particular that found on the web, by national governments. In research conducted over a decade and a half, we carried out tests for Internet filtering and documented content restriction practices in 75 countries. Today we release an update, based on research methods designed to evolve along with the web itself, that shows increasing sophistication by those who place content online, by those who provide content hosting platforms, and by those who wish to stop their citizens from finding and experiencing it. The citizens themselves, from a technical and skills standpoint, remain largely as they were in 1999: they use tools that can at different points secure and surveil their activities, and they can be readily dissuaded from seeking content if roadblocks are encountered. In one major difference from 1999, users are increasingly turning to a handful of aggregating platforms to learn about the world and to engage in dialogue with one another. How those platforms choose to structure their services, in turn, results in an outsized impact in the ongoing tug of war between those who wish to get to a particular destination online and those who wish to prevent them from doing so.

Governments around the world have been using technical, legal, and extralegal strategies to regulate online content for more than two decades. Over the last couple of years, a confluence of technological, behavioral, and market forces have ushered in a new reality in which the playing field has been fundamentally altered. The default implementation of encrypted connections by major social media and content hosting platforms along with messaging applications has effectively downgraded the filtering apparatuses used by states that filter the Internet by counting on "deep packet inspection" or URL analysis to intercept unwanted connections as users attempt to forge them. In those cases, state authorities can no longer selectively block individual accounts, web pages, and stories. For example, governments can generally no longer selectively block a specific article on the *New York Times* or Wikipedia, or a particular account on Twitter or Facebook, without blocking those sites and services in their entirety.

When confronted with this dilemma, some countries have chosen not to block platforms that host accounts and content that were previously subject to blocking. For example, Wikipedia is available in Iran in its entirety, and all of Twitter is accessible in Saudi Arabia. Both of these platforms were subject to selective filtering in the past. Other countries have opted to block platforms entirely. The

2017 blocking of Wikipedia in Turkey is one example in which the Information and Communications Technologies Authority explicitly mentions the use of HTTPS as the basis for the blocking decision.¹ The 2017 blocking of Medium in Egypt is another example. Some have also sought to encourage home-grown platforms more amenable to user monitoring and censorship by the platform operators, and where possible to insist upon data and service localization by foreign platforms so as to enlist them in a similar effort. For example, Iran has launched its own version of YouTube² and Turkey is building a domestic search engine and email service.³ China has been particularly successful in restricting content to home-grown platforms.

Still, where local governments face all-or-nothing filtering decisions because service providers are uncooperative, the stakes are now higher, the political peril of filtering decisions greater, and the potential impact of collateral damage larger.

In similar measure, the issues of sovereignty and jurisdiction have only grown more complicated, and the clout of large Internet platforms has never been greater. Governments have always been constrained in their ability to go after content hosted in foreign countries and to pursue authors who write anonymously or reside overseas. The inability to selectively block individual accounts removes a key tool for governments intent on blocking authors who are particularly problematic for them. As major U.S.- and Europe-based platforms have adopted policies requiring formal processes for removal of content, access by users around the world has, on balance, increased. However, the ability of citizens around the world to fully participate in digital life, and to take advantage of the economic, social, political, and cultural opportunities it affords, remains vulnerable to the actions of regulators and the evolution of service providers' policies, to include various forms of "zoning" whereby content can be blocked by demand to citizens of one state while remaining open to others.

The most striking manifestation of the all-or-nothing choice sometimes left to governments that cannot influence individual platform providers is the rapidly growing incidence of Internet shutdowns around the world in response to social and political events. To our knowledge, this tactic was first used by Nepal in February 2005 when all international Internet connections were cut when martial law was declared. The rise in the use of this blunt tool represents a deep and abiding anxiety of governments over online social activism, and a political willingness to rein this in when threatened, without regard for the damage it might inflict on commerce, let alone on the development of a vibrant digital society.

In this report, we analyze the current state of global Internet content restrictions with an emphasis on state-sponsored filtering through technical means—one of the principal strategies for restricting content online. Our analysis is based on empirical data collected in 45 countries along with a review of rigorously researched secondary sources.

Filtering practices continue unabated for a large majority of countries that are known to employ technical filters to block Internet content. While many governments point to the reinforcement of social and cultural norms as a basis for conducting filtering, political motives are evident in the patterns we document in this study. A small number of the countries we studied engage solely in blocking socially sensitive content such as pornography or gambling. A much larger number of

¹ <https://twitter.com/BTKbasin/status/858628447278694401>

² <https://phys.org/news/2012-12-iran-version-youtube-web.html>

³ <https://turkeyblocks.org/2017/01/06/turkey-building-domestic-search-engine-and-email/>



countries that engage in Internet filtering block both social and political content. For a large majority of the countries that have invested in the technical and administrative infrastructure for blocking Internet content, the bar for blocking political content is evidently no higher than that for blocking social content.

Observing the Internet filtering practices in countries such as China, Russia, Iran, Turkey, and Egypt, it is apparent that filtering practices closely follow the political contours of the respective governments. Repressive states are more likely to block political content, and when they block political content, they most often target political views that are critical of the government. Those states that have less respect for freedom of expression and human rights tend to take an aggressive approach to blocking online content and use all available means to discourage political activity online. Because of the scale, global reach, and rapid production of online content, aggressive Internet filtering regimes typically delegate to third parties—usually private companies that specialize in selling filtering technologies—decisions over which websites to block. For countries that value freedom of expression, due process, and rule of law, carrying out such large-scale blocking is a step too far. Those countries that filter the Internet are those that are able to make such freedom of speech decisions by administrative fiat.

The increase in Internet filtering around geopolitical conflicts is well demonstrated in a review of blocking related to regional conflicts in the Middle East. We find that several countries in the region target content related to the ongoing conflict in Yemen, and that filtering practices, which historically have focused on non-state sources, now include a primary focus on the information and news produced by rival governments.

Introduction and Background

This report presents a global review of state-sponsored Internet censorship. We draw upon an empirical study of Internet filtering in 45 countries carried out over the past year, and we describe the most important trends in the political, social, and technological spheres that shape the current state of global Internet freedom. We begin by summarizing the approaches and tools that governments have at their disposal to address unwanted content online.

State censors continue to employ a variety of tactics to restrict content online drawing from a well developed menu of content control techniques. One option is for governments to go after unwanted content at the source. If the content is hosted on domestic servers, this might be carried out by court order or simply a phone call to the content provider. For content hosted on foreign servers, governments typically have less leverage and must either convince private companies to take down content or enlist foreign governments to act on their behalf.

Another way to deal with content hosted outside a government's jurisdiction is to filter traffic: a primary focus of this study, which we present in the next section. This has been a principal tool for addressing unwanted content for more than two decades along with direct interventions to remove content at the source.⁴ Alternative versions of filtering include throttling traffic or the extreme step

⁴ Jonathan Zittrain, "Internet Points of Control," *Boston College Law Review* 44 (3) 2003.

of shutting down the Internet entirely.

Large hosting platforms have enacted a form of selective filtering on behalf of states in which the platforms block access to certain content when the user request comes from specific countries. This geolocational filtering allows companies to abide by requests from governments to block content that is illegal in that country—for example, a video that insults the Thai king—without removing the content globally. YouTube has selectively filtered content by location since 2007, and Twitter activated a geolocated content-blocking policy in 2012.

A different strategy for restricting online content is to discourage the publication of unwanted content by identifying and going after the authors. This may entail legal and extralegal threats and actions, and it goes hand in hand with surveillance needed to locate and identify authors. The ultimate goal of this strategy—self-censorship—not only is a particularly effective tactic but is very difficult to monitor and document. A more extreme and increasingly common measure is to engage in technical disruption against the content host's server. This might be a DDoS attack or targeted hacking operation directed at a particular individual or website. A key trend over the past several years is the development and commercialization of targeted surveillance tools that are available to governments; these tools frequently operate by surreptitiously installing software on a target's mobile phone or computer.

Finally, governments engage in debates and campaigns to shape media narratives. State actors increasingly join the online public space, participate in discussions, and try to influence discourse. Campaigns representing state interests sometimes hide behind entities posing as non-state actors. While the core strategies for restricting Internet content have changed little over the past two decades, there have been consequential changes in the economic, political, social, and technological context. Internet content providers are increasingly migrating their content to social media. Citizens around the world have adopted encrypted mobile messaging apps like WhatsApp and Viber that allow users to spread information quickly and securely.

Many websites and social media platforms have moved from HTTP to secure HTTPS connections in recent years, which has prevented censors "in the middle" from seeing exactly which pages a user visits. This, in turn, has made blocking specific pages on a domain impossible using standard filtering techniques. There are a number of possible explanations for the increasing use of HTTPS, including new protocols and new features of existing protocols that significantly lower the technical resources required for negotiating HTTPS connections. Efforts such as Let's Encrypt,⁵ a certificate authority that offers free certificates, and well-documented migrations of highly visible websites have also pushed and popularized the deployment of HTTPS and reduced the financial and knowledge barriers to acquiring and deploying HTTPS certificates.⁶

⁵<https://letsencrypt.org/>

⁶ "Secure browsing by default," Facebook Engineering, <https://www.facebook.com/notes/facebook-engineering/secure-browsing-by-default/10151590414803920>



These trends have created additional challenges for government censors. Now more than ever, governments have an all-or-nothing choice when it comes to censorship. Instead of targeting individual webpages or social media accounts, government censors must choose between allowing all content on a social media platform or a messaging app and blocking all content to conceal the information they deem detrimental.

It is unclear whether this development will over time result in lesser or greater access to information. While the adoption of HTTPS by key highly visited websites such as Wikipedia results in greater accessibility in some countries, other government censors choose to block these platforms entirely. China and Iran, for example, block both Twitter and Facebook, forcing users in those countries onto other platforms, except for the small proportion of users who use VPNs or other circumvention tools to get around technical filters.

In a striking departure from the controversial but common practice of selective filtering, a growing number of countries have resorted to shutting down Internet connectivity altogether for periods of time. In India alone, 20 incidents of Internet shutdowns were recorded in the first six months of 2017.⁷ Countries that shut down the Internet in all or parts of the country in 2016 include Bahrain, Ethiopia, Pakistan, Iraq, Turkey, and Malaysia.⁸

State Internet censorship practices are increasingly intertwined with intraregional political dynamics. Politics inform censorship policies, which are used to advance political causes and hinder those of state rivals. We observe an increase in state actors that ban content originating from other state actors driven by political tensions in many parts of the world. Previously, censorship centered around state actors targeting content from non-state actors. As bilateral and regional geopolitical conflicts continue, more states are implementing censorship policies to block access to conflict-related content originating from state political adversaries.

A related trend is that regional political alliances are now shaping Internet censorship policies. This is most visible in the Middle East, where geopolitical conflicts produce alliances that translate into bloc-centered shared Internet censorship. Most recently, a Saudi-allied bloc of countries have begun to block the same websites originating from Qatar as Saudi Arabia does. Also, countries with the same position on the armed conflict in Yemen, the Muslim Brotherhood, or Hezbollah ban the same content.

State censors have extended the reasons and rationales for Internet censorship. The fight against terrorism has been a frequently invoked justification for expanding political censorship, and states have targeted political speech they find offensive. Recently, state censors have started blocking content they label "fake news." For example, Egypt blocked 21 websites in May 2017 for spreading terrorism and fake news.⁹

⁷ <https://www.hrw.org/news/2017/06/15/india-20-internet-shutdowns-2017>

⁸ "Latest News on #KeepItOn," *Access Now*, <https://www.accessnow.org/keepiton-news/>

⁹ <http://www.reuters.com/article/us-egypt-censorship-idUSKBN18K307>

Global Internet Filtering

For this report, we conducted an empirical study of Internet filtering in 45 countries and found evidence of filtering in 26 of them. The methodology follows the approach developed by the OpenNet Initiative (ONI), which we founded with researchers from the Universities of Toronto and Cambridge, running from 2004 to 2014.¹⁰ Two sets of URL testing lists were developed to gather the data that were used to make determinations about the level of Internet filtering across different content categories for each of the 45 countries.

Country-specific testing lists were compiled by topic area experts to cover a range of subjects and content relevant to the political and social landscape of each country, including content in local languages and content categories that have been targeted by Internet censors in the past. A separate global testing list was used in tests across all of the countries. The global list consists of internationally relevant websites and covers major global media organizations, freedom of expression and human rights websites, and social content such as escorts, dating, sex education, and pornography. It also includes prominent Internet tools and websites including social media platforms, free email, anonymizers, free web hosting, search engines, and hacking and translation websites.^{11 12}

In addition to the country-by-country testing and analysis, we devised a research approach using the political backdrop of the MENA (Middle East and North Africa) region to determine if state censors block access to websites that originate from or are affiliated with other states in the region for reasons related to ongoing geopolitical conflicts and intraregional adversaries. The approach consists of two primary methods: we compiled special country-specific lists of websites promoting news aligned with the views of each of the governments and tested each list from within each of the countries.

For the survey of 45 countries, we collected data from vantage points within the network of each country to be tested that supported the analysis to determine which of the URLs on the testing lists were subject to filtering.¹³ We used automated tools to make a preliminary determination for each URL/country test pair based on a combination of weighted metrics about whether a website was intentionally blocked in that country. The software factored in such metrics as the presence of a block page, evidence of DNS tampering, and various connection errors.¹⁴

¹⁰See Ronald Deibert, John Palfrey, Rafal Rohozinski, Jonathan Zittrain, eds., *Access Denied: The Practice and Policy of Global Internet Filtering* (Cambridge: MIT Press), 2008. Also available at <http://access.opennet.net>

¹¹ The global and local test lists are available on our Github: <https://github.com/berkmancenter/test-lists>

¹² A list of content categories and their definitions can be found here: <https://thenetmonitor.org/internet-content-categories>

¹³ In this study, the in-country vantage points were provided by ICLab (<https://iclab.org/>) and Dyn (<https://dyn.com/>). The selection of the 45 countries was determined in part by the availability of vantage points. We plan to expand coverage to additional countries for future projects.

¹⁴ Further details on our filtering detection methods can be found in the Appendix.

To summarize the results of our work, we have evaluated each country to reflect the extent of filtering in each of four thematic areas: politics, social content, conflict/security, and Internet tools. We have divided the degree of filtering into three levels:

- Pervasive filtering is defined as blocking that spans a number of categories while blocking access to a large portion of related content.
- Substantial filtering is either a medium level of filtering carried out over a few categories or a low level of filtering carried out across many categories.
- Selective filtering is either narrowly targeted filtering that blocks a small number of specific websites across a few categories or filtering that targets a single category or issue.

Key Results

	Political Content	Social Content	Conflict/Security	Internet Tools
Bahrain	●●	●●●	●●	●●
China	●●●	●●	●●●	●●●
Egypt	●●	--	●	●
Hungary	--	●	--	--
India	●●	●	●●	●●
Indonesia	●●●	●●●	●	●●●
Iran	●●●	●●●	●●●	●●●
Kazakhstan	●●	●	--	●
Kuwait	●●	●●	●●	●●
Lebanon	--	●	--	--
Malaysia	●●	●●	--	●
Oman	●	●●●	●	●●
Pakistan	●	●●●	●	●●●
Palestinian Territory – West Bank	●	--	--	--
Qatar	●	●●●	●	●●●
Russia	●●	●●●	●●●	●●
Saudi Arabia	●●●	●●●	●●●	●●
Singapore	●	●	--	--
South Korea	●●	●●	●●●	●
Sudan	--	●●	--	●
Syria	●●●	--	●●	●●
Thailand	●●	●●	--	●
Turkey	●●	●●	●●●	●
United Arab Emirates	●●●	●●●	●●●	●●●
Uzbekistan	●●	●●●	●●●	●●
Yemen	●●●	●●●	●●●	●●●

●●● Pervasive filtering; ●● Substantial filtering; ● Selective filtering; -- No evidence of filtering



The table above shows the patterns of censorship across the 26 countries in which we found evidence of filtering. The diversity of filtering practices is evident, from the heavy interventions by China, Saudi Arabia, and Iran to the more limited interventions by Singapore, Hungary, and Lebanon.

Political Content

The political theme consists of content categories such as news and media, human rights, religion, freedom of expression, and environmental controversy. Filtering directed at political opposition to the ruling government is a common type of blocking in many countries across the world. In fact, every country in which we found evidence of technical filtering blocks political content to some degree except for Hungary and Lebanon. The list of countries that engage in substantial political blocking includes Bahrain, Egypt, India, Kazakhstan, Kuwait, Russia, South Korea, Thailand, and Turkey, and the countries that pervasively filter political content are China, India, Indonesia, Pakistan, Saudi Arabia, Uzbekistan, and Yemen.

The Chinese government, for example, creators of one of the world's most sophisticated regimes of Internet filtering and information control, pervasively censors websites that contain criticism of the Communist Party or that report on its human rights record and policies toward Taiwan. For example, the censors block China Digital Times, a bilingual website that aggregates and provides analysis on politically sensitive topics.¹⁵ The website for the Uyghur Human Rights Project (UHRP) is also blocked in China. The site reports on the deteriorating human rights situation in East Turkestan, an area in central Asia under Chinese control.¹⁶

Religious filtering straddles the political and social themes. We categorize websites with religious content in the political theme because they intersect with political filtering. In some cases the censors block religious content that does not conform with state-sanctioned religious belief under the pretext of maintaining political stability. Results reveal an increase in the scope of faith-based filtering in the MENA region, especially around sectarianism. Saudi Arabia, the UAE, and Bahrain block Shi'ite content, and Iran blocks Sunni content. Saudi Arabia, Yemen, UAE, and other Muslim-majority countries including Iran, Indonesia, and Malaysia filter websites that promote Christianity among Muslims or provide critical review of Islam.^{17 18 19}

Overall, we detected evidence of blocking of religious content in China, India, Indonesia, Iran, Kazakhstan, Kuwait, Malaysia, Pakistan, Russia, Saudi Arabia, Singapore, the UAE, Uzbekistan, and Yemen.

¹⁵ <http://chinadigitaltimes.net/>

¹⁶ <http://uhrp.org/>

¹⁷ <http://islamreview.com/>

¹⁸ <http://www.light-of-life.com/>

¹⁹ <http://www.prophetofdoom.net/>



Social Content

Filtering of social content occurs in a majority of the countries for which we found evidence of Internet blocking. Social filtering is focused on topics that go against a country's accepted societal norms, including pornography, gambling, alcohol and drugs, LGBTQ content, and online dating. We found that Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, the UAE, and Yemen in the Middle East, as well as China, Indonesia, Kazakhstan, Malaysia, Pakistan, Russia, Singapore, South Korea, and Thailand, censor pornography, the most commonly targeted category for filtering.

Many of the same countries block gambling, the second most commonly blocked category, to various degrees. Unlike the many countries that block content across a wide variety of categories, Hungary and Lebanon block only gambling sites.

Conflict and Security

A perceived or stated threat to national security is a common rationale that governments use for blocking content. Countries including Egypt, India, Iran, Kuwait, Pakistan, Russia, Saudi Arabia, South Korea, Turkey, the UAE, Uzbekistan, and Yemen all substantially or pervasively filter websites that cover armed conflicts or that host content supporting insurgents. Some ban extremist and terrorist content.

The Iranian government pervasively blocks websites that relate to the conflict and security theme. For example, Iran blocks websites affiliated with ethnic groups within the country such as the Iranian Kurds, whom the government considers a security threat because they seek separation from Iran. It has also recently expanded blocking in the conflict category to include media websites originating from its regional rival, Saudi Arabia. Similarly, Saudi Arabia blocks key news websites originating from or affiliated with Iran.

South Korea, despite its status as a world leader in Internet speed and broadband adoption, pervasively blocks websites that support North Korea in the ongoing conflict between the two countries or that advocate for unification of North and South Korea.^{20 21}

Internet Tools

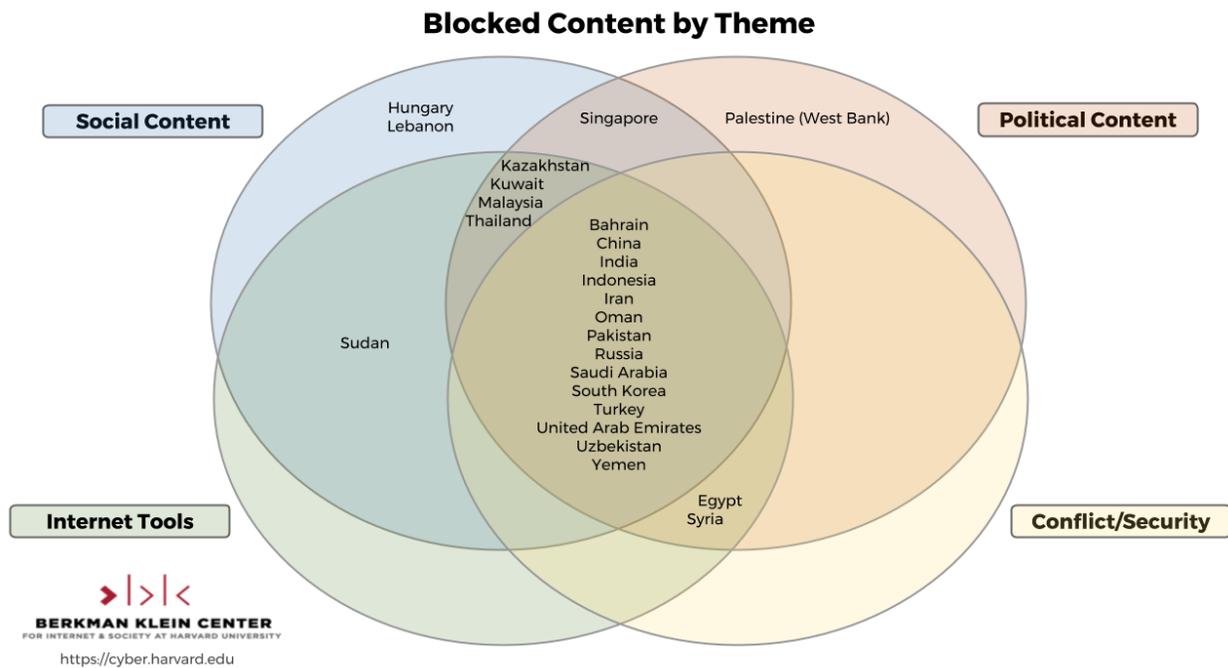
Our Internet tool theme is made up of categories such as anonymizers, censorship circumvention tools, social media platforms, and streaming and P2P file sharing sites. The websites counted in this theme were blocked in multiple countries around the world. For example, anonymizers/circumvention tools and social media platforms, the most frequently blocked categories in the Internet Tool theme, are filtered in more than 10 countries.

Filtering of content in this theme has increased in recent years. The governments of China, India, Indonesia, Kazakhstan, Russia, South Korea, Turkey, and Uzbekistan have all increased their levels of Internet tool filtering since prior rounds of filtering research. For example, filtering research by the ONI in 2010 found that India selectively filtered Internet tools, but today, the government filters a substantial number of anonymizers, cloud storage sites, and P2P torrent sites.

²⁰ <http://onekorea.org/>

²¹ <http://www.minjok.com/>

Across the different thematic areas, we find websites that are blocked from both the country-specific testing lists and the global list. A substantial portion of the websites that we identified as being blocked related to Internet tools and social topics are from the global list. A larger portion of political and conflict/security content is found on the country-specific testing lists.



The Venn diagram above shows that the majority of the 26 countries filter content across all four themes, although the depth of the filtering varies. A few countries selectively block content in only one or two themes. The Internet in Hungary and Lebanon, for example, remains unfiltered except for a small number of online gambling websites. The government of Singapore continues to selectively filter a small set of pornographic websites and a few websites that are critical of Islam. Egypt and Syria filter political and conflict-related content without blocking social content.

The overall trend suggests that once the administrative, technical, legal, and political obstacles to implementing a filtering regime have been overcome, there is a tendency to extend blocking to include political and social content as well as to the core tools and platforms that might be used to circumvent Internet filters. The evidence indicates that the slope is indeed slippery in almost every country that enacts Internet filtering. The incremental cost of adding additional topics and websites to blocking lists is apparently modest once the filtering system is in place. Several countries, including China, Egypt, Iran, Malaysia, and Turkey, block entire platforms to ban access to objectionable content, even if the targeted content represents a small proportion of the content hosted on the platform.

Key Trends

Internet filtering is increasing around the world in scope and depth. New countries have started filtering, and others have significantly expanded the scope of their filtering in the last several years. Since the last round of testing by the ONI in 2012, more countries have started to filter the Internet. These countries include Egypt, Russia, and Malaysia. This development is not particularly surprising, as each country has a history of media control practices that include harassment or imprisonment of online activists, journalists, and cyber-dissidents. The highly consequential all-or-nothing decisions over the blocking of major content hosting platforms have undoubtedly raised the political stakes of these filtering choices. This has had less of an impact on the manner in which countries address other filtering decisions. The evidence suggests that the decisions whether to block potentially thousands or millions of sites—depending on the subjects that are targeted—have not changed.

Geopolitical Internet Filtering

Internet filtering practices in the MENA region have expanded in recent years, especially with respect to the region's geopolitical tensions and conflicts between countries. Tensions that have led to increased filtering in the past year include the conflict between regional rivals Saudi Arabia and Iran, the armed conflicts in Syria and Yemen, and conflicts involving the armed Lebanese group Hezbollah and the transnational group the Muslim Brotherhood. Hezbollah and the Muslim Brotherhood have become increasingly entangled in regional politics—Hezbollah for its engagement in the military operations in Syria in support of the government of President Bashar al-Assad,²² and the Muslim Brotherhood after its member Mohammed Morsi won the office of president in Egypt in 2012 and later was ousted by a military-backed campaign.²³ Hezbollah is labeled a terrorist organization by most of the Arab states,²⁴ and the Muslim Brotherhood is banned in Egypt,²⁵ Saudi Arabia,²⁶ and the UAE.²⁷

Here are key results of testing access to the curated lists of URLs using in-country testers and vantage points in Saudi Arabia, Iran, Iraq, Yemen, the UAE, Qatar, Bahrain, Lebanon, and Turkey:

- Saudi Arabia blocks access to major websites affiliated with Iranian media, the Iranian ally Hezbollah, the Houthis in Yemen, Syrian websites editorially aligned with the Syrian government, and websites associated with the Muslim Brotherhood, which Saudi Arabia labels a terrorist group.
- Iran blocks media websites from Saudi Arabia, Bahrain, Kuwait, Yemen, and Lebanon, all of which are critical of Iran's foreign policy in the region.

²² Nadav Pollak, *The Transformation of Hezbollah by Its Involvement in Syria*,

²³ Beverley Milton-Edwards, *The Muslim Brotherhood: the Arab spring and its future face* (Routledge), 2016.

²⁴ <http://www.bbc.com/news/world-middle-east-35789303>

²⁵ <http://www.reuters.com/article/us-egypt-explosion-brotherhood-idUSBRE9BO08H20131225>

²⁶ <http://www.reuters.com/article/us-saudi-security-idUSBREA260SM20140307>

²⁷ <http://www.reuters.com/article/us-emirates-politics-brotherhood-idUSKCN0IZ0OM20141115>

- The UAE blocks media websites affiliated with Iran and Hezbollah as well as websites affiliated with the Muslim Brotherhood, which it bans as a terrorist organization.
- Bahrain blocks media websites affiliated with Iran and Hezbollah.
- Oman blocks websites affiliated with Hezbollah and the Muslim Brotherhood.
- Yemen's national ISP, which is under the control of the Houthis, blocks Saudi media websites. Saudi Arabia is leading a military coalition against the Houthis.

The expansion of geopolitical filtering contributes to the pressure on the state censors who face the conundrum of overblocking or underblocking as more content around the conflicts is hosted on centralized platforms including social media. A forthcoming study from the Berkman Klein Center around wartime censorship finds that warring parties in Yemen implement military-aligned Internet censorship policies to control flow of information. While they manage to block access to a large number of websites originating from their adversaries, citizens resort to centralized platforms that escape state filtering to frame their views on the war on their own terms.²⁸

New and Expanding Internet Censorship

Egypt

Starting in late 2015, Egypt has selectively blocked political websites that contain content critical of the government. In May 2017, the Egyptian authorities began to substantially filter political content, and as of June 2017, the lists of blocked URLs had grown to more than 100.²⁹ The government claims some of the websites are spreading false news or are affiliated with the Muslim Brotherhood, which the government bans as a terrorist group. Egypt also blocks official Qatari media websites and news dailies over allegations that Qatar supports the Egypt-based Muslim Brotherhood and terror activities. Additionally, Egypt recently blocked a website that protests the government's transfer of two islands from Egypt to Saudi Arabia.³⁰ In 2016, the Egyptian authorities attempted to block Signal, an encrypted communication application, before the tool operators managed to sidestep the blocking measures.³¹ Media reports suggest that the government implements Internet filtering at the primary Internet access point, thus bypassing the ISPs.³² The access point is thought to be the same hub used to enact most of the January 2011 Internet outage that received worldwide attention and

²⁸ Forthcoming Berkman Klein Center research paper.

²⁹ https://afteegypt.org/right_to_know-2/publicationsright_to_know-right_to_know-2/2017/06/04/13069-afteegypt.html?lang=en

³⁰ Hamza Hendawi, "Egypt committee approves deal on islands' transfer to Saudis," *ABC News*, <http://abcnews.go.com/International/wireStory/egyptian-committee-approves-transfer-islands-saudis-48003868>

³¹ Andy Greenberg, "Encryption App 'Signal' Fights Censorship with a Clever Workaround," *Wired*, <https://www.wired.com/2016/12/encryption-app-signal-fights-censorship-clever-workaround/>

³² <https://www.madamasr.com/en/2017/06/21/feature/politics/egyptian-government-bypasses-isps-to-block-access-to-websites-telecommunications-ministry-source/>



was carried out in response to the anti-government protests.³³

Palestinian Territories - West Bank

In June 2017, after a few years of no blocking, the Palestinian Authority ordered ISPs to block 12 news websites affiliated with the rival Islamist group Hamas, which controls the Gaza Strip; websites affiliated with dismissed Fatah leader Mohammed Dahlan; and 10 news websites that provide news and views on Palestinian politics. This political blocking is a result of political tensions between the Palestinian Authority headed by President Mahmoud Abbas, which controls the West Bank, and Hamas. Because of the political friction between the two sides, the ISPs in Gaza block different political content. Although connectivity to Palestinian territory is routed through Israel, there is no evidence that Israel conducts any "upstream" blocking, and the blocking of content that occurs appears to be carried out by Palestinian ISPs.

Russia

Our most recent round of testing in Russia has revealed extensive filtering of content in each content theme. The government blocks websites critical of the government and websites associated with militant or extremist organizations. It also pervasively blocks gambling, pornography, and alcohol and drug websites. The implementation of filtering in Russia follows upon a long history of content restrictions in Russia based on more subtle and difficult-to-document tactics, including extensive surveillance, state-sponsored information campaigns, and offensive cyberattacks.³⁴

Malaysia

Despite the Malaysian government's guarantee not to censor the Internet,³⁵ research by multiple organizations has revealed evidence of Internet blocking in Malaysia.³⁶ The government filters pornography and gambling websites substantially, and torrent sites selectively. The government also censors news outlets, medium.com, and other blogs that report on the Malaysian Prime Minister's alleged involvement in a billion-dollar misappropriation scandal in 2015. Additionally, the government blocks at least one website that is critical of Islam.

Regional Trends

Common Filtering Practices in the Commonwealth of Independent States (CIS)

Countries in the CIS in which we detected filtering—Kazakhstan, Russia, and Uzbekistan—have a high degree of overlap in the content categories they block or leave accessible. While any two countries chosen at random would have, on average, 57 percent agreement in their category-level

³³ <https://www.wired.com/2011/02/egypt-off-switch/>

³⁴ Ronald Deibert and Rafal Rohozinski. Control and Subversion in Russian Cyberspace in Ronald Deibert, John Palfrey, Rafal Rohozinski, Jonathan Zittrain, eds., *Access Controlled: The Shaping of Power, Rights, and Rule of Cyberspace* (Cambridge: MIT Press), 2010. Also available at <http://access.opennet.net>

³⁵ "MSC Malaysia Bill of Guarantees," Malaysia Digital Economy Corporation, 2017, <https://www.mdec.my/msc-malaysia/bill-of-guarantees>

³⁶ Maria Xynou et al., "The State of Internet Censorship in Malaysia," OONI, <https://ooni.torproject.org/post/malaysia-report/>



censorship decisions, these three countries have an average of 80 percent agreement. The overlap among Kazakhstan, Russia, and Uzbekistan is noteworthy because each country blocks websites in only 16 or fewer of the 40 categories that make up the four themes. Each country blocks websites related to gambling, political news, religion, various communication platforms, and pornography.

Russia and Uzbekistan both have intensely controlled online environments with substantial or pervasive filtering in each content theme. Although they block content in fewer categories than other countries do, the depth of filtering in the categories that they filter is high.

Increased Faith-Based Filtering

ISPs in the MENA region have increased the scope of faith-based filtering, especially around sectarianism. Saudi Arabia, the UAE, and Bahrain block Shiite content, and Iran blocks Sunni content. The increase in religious blocking intersects with political censorship to the extent that political opposition actors happen to be religious players or belong to a religious sect that is not officially sanctioned by the state. Also, the armed conflicts in the region have deep-seated sectarian dimensions, with each sect supporting one side of the conflict—hence the overlap with political filtering.

Increasing Adoption of HTTPS and Implications for Filtering

The adoption of HTTPS is markedly changing the scope and impact of Internet filtering. The logic and value of HTTPS extends far beyond questions of Internet filtering. Those domains and platforms that shift to HTTPS offer their visitors greater security and the ability to transfer information safely and out of the view of routine surveillance. There is considerable variation, though, in the use of security features across different websites. A recent study finds that security adoption varies across industry sectors and that it is weak within news and sports websites.³⁷ A side effect of adopting HTTPS is that it makes selective filtering impossible with standard filtering techniques.³⁸

Here we track the increasing adoption of HTTPS as a measure of this global trend. To measure the prevalence and usage of HTTPS technologies over time, we turned to the Common Crawl dataset of monthly web crawls.³⁹ This dataset consists of requests and responses to millions of websites and is therefore extremely useful for analyzing longitudinal trends on the web.

We limited our analysis to data from March 2015 to May 2017. For each month, we looked at the number of successful responses to requests for each of 2,046 domains on our global test list. This test list is composed of the top 1,000 most trafficked websites according to Alexa Internet, Inc., plus another approximately 1,000 websites that are deemed by experts to be of global interest. We split

³⁷ William J. Buchanan, Alan Woodward, and Scott Helme, "Cryptography across industry sectors," *Journal of Cyber Security Technology*, June 2017,

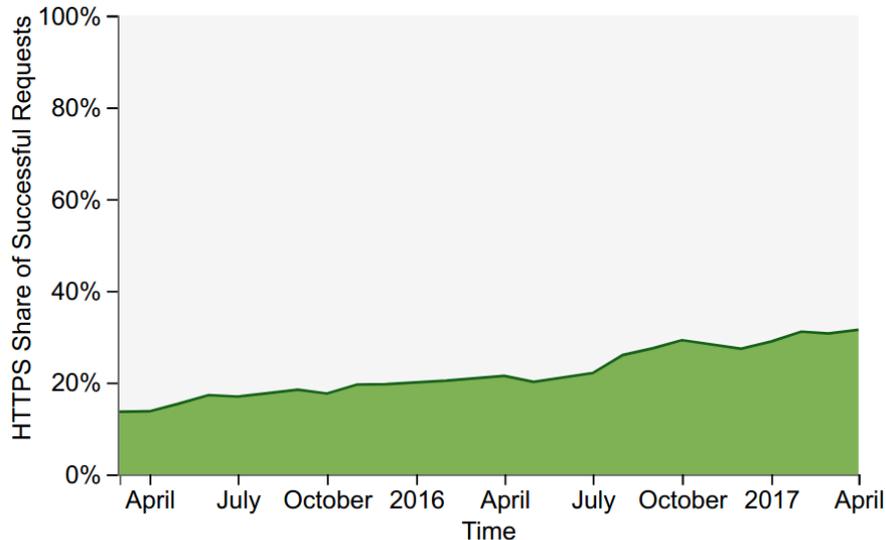
<http://www.tandfonline.com/doi/abs/10.1080/23742917.2017.1327221?journalCode=tsec20>

³⁸ In order to selectively block individual accounts or pages on a domain that is reachable only via HTTPS, the filters would need to prevent or break the encryption between the domain and the user. This may be done by a number of different methods, almost all of which are visible to an informed user using a modern browser. To our knowledge, no country has attempted this strategy at a broad scale.

³⁹ Common Crawl, <https://commoncrawl.org/>

the responses into two groups: those that took place over HTTP and those that took place over HTTPS. More than half of the websites support both protocols, so for each site we looked at how the share of HTTPS requests and responses changed from month to month over our period of analysis.

In total, we analyzed data on approximately 136 million successful Common Crawl requests to the 2,046 domains that constituted our global test list. About 106 million of those requests were fulfilled over HTTP and the remaining 30.4 million over HTTPS; about 22 percent of all requests took place over HTTPS. When the results are broken down by month, we can see a clear trend line indicating the increasing use of HTTPS across these 2,046 websites:



In March of 2015, HTTPS accounted for 13.6 percent of all requests, but by the beginning of May 2017, HTTPS usage had grown to 31.5 percent.

To further examine the current state of HTTPS connectivity, we attempted to access each domain on the global test list through both HTTP and HTTPS. Of the 2,046 websites on the list, there are 827 domains (40.4 percent) that are available only through HTTPS connections; users attempting to connect to the websites in this group through HTTP are automatically redirected to a secure HTTPS connection. For example, users attempting to connect to <http://www.nytimes.com> are redirected to <https://www.nytimes.com>. A few examples of other popular websites that require HTTPS connections are theguardian.com, theverge.com, amazon.com, and google.com.

The global list includes 1,151 websites (56.2 percent) that allow users to connect through HTTP. Among these 1,151 domains, 161 websites (7.9 percent) are inaccessible through HTTPS and automatically redirect users to an unsecure, HTTP connection. For example, users attempting to connect to <https://www.bbc.com> will be redirected to <http://www.bbc.com>. Other popular websites that do not yet offer access through HTTPS include slate.com, cnn.com, and npr.org.

These findings point in two directions: one, unsurprisingly, is that HTTPS use is on the rise, and two, HTTPS adoption still has a long way to go, even among some of the world's most trafficked websites.

The implications of the broadening adoption of HTTPS on filtering are mixed. The use of HTTPS by many websites, especially those hosting user-generated content, poses a challenge to the censors. This leaves them with the options of blocking either everything or nothing. The adoption of HTTPS by highly visited websites such as Wikipedia has resulted in greater accessibility for those sites in many cases.⁴⁰ As of August 2011, Saudi Arabia was blocking specific Wikipedia entries such as one on the theory of evolution and individual Twitter accounts such as those of Egyptian activist Wael Ghonim and Gamal Eid, the director of a Cairo-based regional human rights NGO who often posts tweets critical of the record of freedom of expression in Saudi Arabia and other Arab countries. The website of Eid's human rights organization remains blocked.⁴¹ Also, Saudi Arabia used to block the Facebook pages of politically objectionable individuals.⁴² We have no evidence that Saudi Arabia is now blocking any individual accounts on Twitter, Wikipedia, or Facebook.

In 2013, researchers found that Iran was blocking access to nearly 1,000 Persian-language Wikipedia articles. The articles contained content related to politics, sex, religion, human rights, arts and culture, media and journalists, academia, profanity, drugs, and alcohol. Approximately a quarter of the articles were biographies of people the government had arrested, detained, or killed.⁴³ Internet Monitor's 2017 study of Wikipedia accessibility uncovered evidence that accords with the findings of the 2013 study. Furthermore, the 2017 study found that individual articles subject to censorship in the past have been receiving increased levels of traffic since Wikipedia transitioned to HTTPS in June 2015, suggesting that the transition has made it more difficult for Iran to censor selected Wikipedia content.⁴⁴

Other government censors choose to block entire domains and platforms. For example, online activists, writers, and blocked websites in Egypt resorted to using the online publishing platform medium.com to disseminate their content, exploiting the fact that the censors cannot ban individual accounts since Medium uses HTTPS by default. Medium.com became a preferred space for many local and regional longform authors to bypass blocking. This tactic, however, did not work for long. In June 2017, the censors blocked access to the entire website, which resulted in massive overblocking, with access denied to all of the millions of posts on Medium. In a similar situation, the censors in Malaysia blocked the entire platform in January 2016 after the company did not comply with a government request to take down objectionable content related to a corruption case.⁴⁵ In another similar case, Turkey banned access to all of Wikipedia because the censors could not block individual offensive entries.⁴⁶ As was confirmed by our research, China and Iran continue to block Twitter and Facebook.

⁴⁰ Justin Clark, Robert Faris, and Rebekah Heacock Jones, "Analyzing Accessibility of Wikipedia Projects Around the World," Berkman Klein Center for Internet and Society,

<https://cyber.harvard.edu/publications/2017/04/WikipediaCensorship>

⁴¹ The Arabic Network for Human Rights Information, <http://anhri.net/?lang=en/>

⁴² "Saudi Arabia," Freedom House, <https://freedomhouse.org/report/freedom-net/2012/saudi-arabia>

⁴³ Nima Nazeri and Collin Anderson, "Citation Filtered: Iran's Censorship of Wikipedia," Center for Global Communication Studies, Annenberg School for Communication, University of Pennsylvania, Nov. 2013, http://www.global.asc.upenn.edu/fileLibrary/PDFs/Citation_Filtered_Wikipedia_Report_11_5_2013-2.pdf

⁴⁴ Justin Clark et al., "Analyzing Accessibility of Wikipedia Projects."

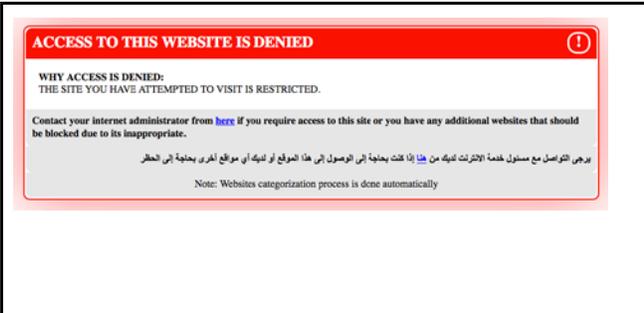
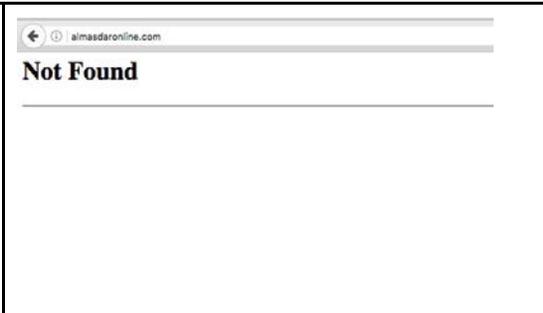
⁴⁵ Amanda Connolly, "Medium stands by investigative journalists as Malaysia blocks the site," *The Next Web*, https://thenextweb.com/media/2016/01/27/medium-stands-by-investigative-journalists-as-malaysia-blocks-the-site/#.tnw_cKBNz3om

⁴⁶ Hande Atay Alam, Merieme Arif, and Joe Sterling, "Turkey blocks Wikipedia over what it calls terror 'smear campaign,'" *CNN*, <http://www.cnn.com/2017/04/29/europe/turkey-wikipedia/index.html>

Transparency in Internet Filtering

Countries have varying degrees of transparency in Internet filtering. In the MENA region, the censors are generally more transparent about filtering social content on religious and cultural grounds, and they often provide commentary and statistical reports about their efforts to the media. They are less likely to acknowledge or comment on political filtering. In an unprecedented case of transparent political filtering, the government authorities in Bahrain, the UAE, Saudi Arabia, and Egypt publicly announced in June 2017 that they were blocking websites based in or affiliated with Qatar. This announcement came after the countries severed their diplomatic relations with the state of Qatar over allegations that Qatar supports the Muslim Brotherhood and militant groups.

On the technical level, some of the ISPs return explicit block pages, while others serve error messages. We have identified block pages in Bahrain, Hungary, India, Indonesia, Iran, Kuwait, Malaysia, Oman, Pakistan, Qatar, Russia, Saudi Arabia, Singapore, South Korea, Sudan, and Thailand.⁴⁷ Some ISPs such as those in the UAE and Saudi Arabia link from their block pages to information about Internet filtering regulations in their respective countries. They also provide email addresses for the public to request whitelisting or blacklisting of websites. ISPs in Turkey and India mention on their block pages court orders based on which blocking is taking place. The Yemen National ISP YemenNet serves an explicit block page for social content and Internet tools. It, however, serves a 404 "Not Found" page for political content. An in-depth analysis by Citizen Lab at the University of Toronto has concluded that the 404 message is a disguised block page.⁴⁸ Apparently, the ISP attempts to mislead users who seek to access political content into believing the websites they are trying to reach are not available. The ISPs in Egypt serve timeout messages for the blocked websites.

	
Explicit block page for social content	Error message for political content

Screenshots of two block pages served by Yemen's national ISP YemenNet.

⁴⁷ See more examples of block pages on our country profile pages: <https://thenetmonitor.org/research/2017-global-internet-censorship/irn>

⁴⁸ Jakub Dalek et al., "Information Controls during Military Operations: The case of Yemen during the 2015 political and armed conflict," Citizen Lab at the Munk School of Global Affairs, <https://citizenlab.org/2015/10/information-controls-military-operations-yemen/>

Governments Exert Greater Influence on Public Discourse Online

Another trend is that state actors are increasingly engaging in debates and discussion on platforms conventionally used by non-state early adopters. However, the playing field is not level, as state actors—frequently self-identified as such—typically have greater leverage. They take part in conversations but block access to content they do not approve of when they are able to do so, and arrest or harass contributors of content they deem objectionable. For example, social media state accounts from the Gulf countries continue to make the case against Qatar over its diplomatic row with Saudi Arabia, the UAE, and Bahrain. However, these states have threatened their citizens with jail sentences and financial fines under cybercrime laws should they express sympathy toward Qatar on social media.⁴⁹ In June 2017, Bahrain followed through on threats to arrest citizens who express sympathy toward Qatar.⁵⁰ In Egypt, state media are able to support the government's decision to cede sovereignty of islands from Egypt to Saudi Arabia. However, the censors there have blocked a website that argues and mobilizes against the move.⁵¹ Some of these strategies do not require the cooperation of the content hosting platforms—for example, directly engaging with users or threatening citizens who are openly linked to social media activity. Without the help of the platforms, removing content and unmasking anonymous users is more difficult. Communications between governments and hosting companies are typically private; we do not know the extent to which platforms cooperate with different governments in removing content and identifying account holders.

Social Media Censorship

Internet content providers are increasingly migrating to social media platforms. Many political and human rights groups located in the MENA region no longer maintain websites and instead use social media to disseminate content. Others redirect traffic from their websites to Facebook pages. In Libya, for example, we find that the websites of many political groups, advocacy organizations, and news outlets are defunct and have been replaced by Facebook pages.⁵²

According to Freedom House's *Freedom on the Net* report, censorship of social media platforms and communication apps reached an all-time high across the globe in 2016. Freedom House's study measured the level of Internet and digital media freedom in 65 countries based on an examination of local laws relevant to the Internet, website availability testing, and interviews from in-country sources. More governments than ever before targeted social media platforms and messaging apps like WhatsApp and Telegram in an attempt to control the digital flow of information.

Freedom House's study found that 24 governments slowed or cut off access to social media and communication apps between May 2015 and May 2016—an increase from 15 countries the previous year. Messaging apps such as Telegram, Viber, Facebook Messenger, LINE, IMO, Google

⁴⁹ See for example, Gulf News, Mariam M. Al Serkal, "Qatar sympathisers to face fine, jail," June 7, 2017, <http://gulfnews.com/news/uae/government/qatar-sympathisers-to-face-fine-jail-1.2039631>

⁵⁰ Dana Khraiche, "Man Detained in Bahrain for Opposing Anti-Qatar Action Online," Bloomberg Politics, <https://www.bloomberg.com/politics/articles/2017-06-14/man-detained-in-bahrain-for-opposing-anti-qatar-action-online>

⁵¹ Tiran w Sanafir, tiranwsanafir.com

⁵² For example, the National Centrist Party of Libya's Facebook Page, <https://www.facebook.com/NCP.Libya/>, and Libya Al Ahrar TV's Facebook Page, <https://www.facebook.com/LibyaAlAhrarTV>



Hangouts, and WhatsApp were blocked by multiple countries; 12 of the 65 countries blocked the most popular messaging app, WhatsApp. Ten countries blocked access to voice over Internet protocol (VOIP) platforms that enable video chat such as Skype or Google Hangouts.⁵³

Commercialization of Specialized Tools for Targeted Surveillance

There has been a notable increase over the past several years in the use of malware to conduct targeted surveillance of political dissidents, activists, and journalists. Research from the Citizen Lab at the Munk School of Global Affairs, University of Toronto, has revealed numerous cases from around the globe of the use of digital spying tools designed for criminal investigations and counterintelligence to target journalists, human rights advocates, and activists.⁵⁴ Researchers there document a growing number of incidences of misuse of spyware for political ends worldwide. Furthermore, they argue that the increase in use of "lawful intercept" spyware to target political opponents and human rights defenders is evidence that such spyware has significant abuse potential.⁵⁵ The Lab concludes that there is evidence of an informal "principle of misuse" and that the misuse of government-exclusive spyware is a global problem, especially because network breaches reveal that some customers of malware manufacturers such as Gamma Group (FinFisher) and Hacking Team include a global list of government customers in countries known for their poor human rights records.⁵⁶

A Wider Array of Strategies to Suppress Information Flows

While filtering remains a mainstay of content restriction policies, the continued rise of alternative approaches to clamping down on free expression is well documented. In addition to the unprecedented level of direct social media platform censorship by way of shutdowns of applications in 2016, governments often used law enforcement to silence citizens. The governments of 38 countries arrested individuals as a result of their social media activity, according to Freedom House.

For example, Turkish authorities arrested 1,656 people in the second half of 2016 and, as of December 2016, were investigating over 10,000 others for insulting officials on social media or for their alleged support of terrorist organizations. According to Turkey's Interior Ministry, the charges included "provoking hatred among the people, praising terrorist organizations, insulting statesmen, and targeting the indivisibility of the state or safety of citizens."⁵⁷

In February 2016, a court in Saudi Arabia sentenced a man to 10 years in prison and 2,000 lashes for expressing his atheist beliefs on Twitter.⁵⁸ A branch of Iran's armed forces, the Islamic Revolutionary Guards Corps (IRGC), "summoned, detained, and warned" 450 or more administrators of social media groups in August 2016. According to a website associated with the

⁵³ "Freedom on the Net 2016," *Freedom House*.

⁵⁴ <https://citizenlab.org>

⁵⁵ <https://citizenlab.org/2016/08/million-dollar-dissident-iphone-zero-day-nso-group-uae/>

⁵⁶ <https://citizenlab.org/2017/06/reckless-exploit-mexico-nso/>

⁵⁷ "Turkey arrests 1,656 social media users since summer," Associated Press, Dec. 24, 2016, <http://bigstory.ap.org/article/ed3e585528e144f2b8fceab5b450eb97/turkey-arrests-1656-social-media-users-summer>

⁵⁸ "Saudi Arabia Sentenced a Man to 10 Years in Prison and 2,000 Lashes for Atheist Tweets," *VICE News*, February 27, 2016, <https://news.vice.com/article/saudi-arabia-sentenced-man-to-10-years-in-prison-and-2000-lashes-for-atheist-tweets>

IRGC, the social media users "were carrying out immoral activities, insulted religious beliefs, or had illegal activities in the field of fashion."⁵⁹ Bloggers critical of governments were at risk for publishing their work online as well. For example, a top Vietnamese blogger who writes under the pseudonym Mother Mushroom was arrested in October 2016 for criticizing the government over its handling of a toxic chemical spill earlier that year.⁶⁰

Government authorities often send requests to social media platforms to remove content. Although these requests sometimes have legitimate aims, such as to protect children from sexual content or to counter violent or other forms of illegal speech, governments have exploited such policies to suppress information that would otherwise be protected speech.⁶¹ For example, activists in Turkey say their government uses social media takedown policies to silence criticism and ban the spread of news related to Kurdish politics.^{62 63}

According to its transparency reports, Twitter received 5,569 court orders and other legal requests from Turkish authorities for content to be removed from Twitter in 2016. In total, Twitter withheld 2,060 tweets and suspended 512 accounts. Turkey was the country that sent by far the most requests. The country with the second most requests, Russia, was successful in withholding 271 tweets and suspending 65 accounts through a total of 2,118 requests.⁶⁴

Moreover, state authorities impose license requirements to allow news websites to operate and remain accessible within their jurisdictions. For example, in November 2016, the censors in Qatar blocked a news website for failing to register for and obtain a media license from the Ministry of Culture.⁶⁵

These non-technical content restriction strategies were once used as substitutes for Internet filtering in countries such as Russia and Egypt. Today, the overall pattern is more homogeneous. These alternative strategies are most prevalent in countries that also aggressively filter the Internet.

Internet Disruptions

As Internet tools are increasingly used for mobilizing and sharing information, authorities more often resort to disrupting the Internet to respond to situations when the threat is from the flow of information among citizens. A growing number of countries employ this tactic. For example, the authorities in Bahrain disrupt Internet service and mobile services in areas densely populated by Shiites who organize street protests against the government.⁶⁶ Access Now documented more than

⁵⁹ "Iran warns, detains 450 social-media admins, citing 'immoral' posts," *World Tribune*, August 24, 2016, <http://www.worldtribune.com/iran-warns-detains-450-social-media-admins-citing-immoral-posts/>

⁶⁰ Mike Ives, "Vietnam Arrests Mother Mushroom, a Top Blogger, for Criticizing Government," *New York Times*, October 11, 2016, https://www.nytimes.com/2016/10/12/world/asia/vietnam-arrest-blogger-mother-mushroom.html?_r=1

⁶¹ "Demystifying Social Media Censorship—in Arabic, Spanish and English," *Global Voices*, November 4, 2016, <https://globalvoices.org/2016/11/04/demystifying-social-media-censorship-in-arabic-spanish-and-english/>

⁶² *Ibid.*

⁶³ Efe Kerem Sozeri, "Uncovering the accounts that trigger Turkey's war on Twitter," *The Daily Dot*, January 31, 2015, <https://www.dailydot.com/layer8/twitter-transparency-report-turkey-censorship>

⁶⁴ "Removal Requests," *Twitter*, <https://transparency.twitter.com/en/removal-requests.html>

⁶⁵ Omar Chatrivala, "An update on Doha News being blocked in Qatar," *dohanews.co*, January 25, 2017, <https://dohanews.co/an-update-on-doha-news-being-blocked-in-qatar/>

⁶⁶ <https://bahrainwatch.org/blog/2016/10/07/100-days-since-internet-shutdown-in-duraz/>

50 Internet shutdowns in 18 countries in 2016—more than double the number of shutdowns documented in 2015.⁶⁷ India shut down the Internet 20 times in the first six months of 2017 to "prevent violence fueled by rumors circulated on social media or mobile messaging applications."⁶⁸

The authorities in Iraq and Algeria disrupt the Internet to prevent the leaking of school exams and the exchange of answers among students. Access Now documented shutdowns in at least 11 African countries in 2016, several of which took place during national elections.⁶⁹

Further complicating the matter of Internet shutdowns in Africa is that not all citizens are opposed to the idea, out of fear that free speech leads to bloodshed. The media, for example, were blamed for inciting violence that left 1,100 dead and 650,000 displaced after Kenya's 2007 presidential election.⁷⁰ Ethiopians often turn to external sources for news because of the lack of reliable local and national media. But according to some citizens, diaspora media outlets and social media activists hostile to the current regime are comfortable making claims that may incite further violence, because they are free from the consequences of their dialogue.⁷¹ According to Voice of America, some Ghanaian citizens believed before the November 2016 presidential election that police would be justified in enacting a social media ban if it would protect unarmed citizens from those who use social media to spread violence.⁷²

Summary and Conclusions

In this report we describe and document several key emerging trends that together fundamentally change the logic and application of Internet filtering. Two notable trends are closely tied: the increasing migration of content and communication to centralized platforms and apps, and the expanding use of encrypted connections by websites and platforms. Together, these two trends greatly diminish governments' ability to fine-tune filtering targets.

The prominence and influence of platforms is greatly enhanced by this process; governments that block independent websites provide a strong incentive for content producers to move their content onto larger platforms. Independent media producers then become dependent upon these platforms. For many, this offers a safe haven for their work, particularly when the platforms defend their ability to publish and resist possible incursions from state and non-state actors that exert pressure to remove controversial or politically laden content.

Several inherent vulnerabilities exist as well. Content producers become subject to the terms of the platforms and may see their content taken down with little recourse. Moreover, the platforms may submit to the pressure from governments to take content down. Additionally, adversaries are at times able to manipulate the processes put in place by platforms to flag violations of their content

⁶⁷ "#KeepItOn," *Access Now*, <https://www.accessnow.org/keepiton/>

⁶⁸ <https://www.hrw.org/news/2017/06/15/india-20-internet-shutdowns-2017>

⁶⁹ Abdi Latif Dahir, "More African governments blocked the internet to silence dissent in 2016," *Quartz*, December 31, 2016, <https://qz.com/875729/how-african-governments-blocked-the-internet-to-silence-dissent-in-2016/>

⁷⁰ Judie Kaberia, "Kenya: Too Little Action on Hate Speech?" *OpenNet Africa*, March 30, 2014, <http://www.opennetfrica.org/kenya-too-little-action-on-hate-speech/>

⁷¹ James Jeffrey, "Ethiopia: Internet shutdowns take their toll on economy," *African Business*, December 29, 2016, <http://africanbusinessmagazine.com/region/east-africa/ethiopia-internet-shutdowns-take-toll-economy/>

⁷² Peter Clotey, "Ghana Police Chief Criticized Over Proposed Social Media Ban," *VOA News*, May 27, 2016, <http://www.voanews.com/a/ghana-police-chief-criticized-proposed-social-media-ban/3349810.html>

standards and cause the takedown of content or closing of accounts. Several governments have shown a willingness to block entire platforms, despite the massive collateral damage, in order to block access to a small number of accounts or stories.

This growing dependence on large platforms is at odds with the decentralized ideal of Internet architecture and reinforces the feudal structure of the Internet. On balance, it appears to expand Internet freedom for most users that reside in restrictive Internet environments. The picture is not entirely rosy, though. One particularly striking trend is that governments are, with increasing frequency, shutting down communications infrastructure altogether for periods of time.

We describe in this report the increase of geopolitical Internet filtering, which stands in contrast to the socially minded rationales often put forward to justify filtering. Political disputes and conflicts now more frequently trigger the use of Internet censorship and spark an increase in state-to-state censorship. An interesting associated trend is the emergence of shared Internet censorship policies by political blocs.

Another shift we observe is greater engagement of government officials and state agencies in the digital public sphere. They post official statements, comment on news, and defend government policies. More frequently than in the past, governments try to influence online discourse related to public policy and compete for attention online with the non-state actors who were early adopters of digital communication tools. Moreover, pro-government actors are now more active in digital spaces than they were a few years earlier.⁷³ State actors have advantages that may enable them to hold greater sway than non-state actors. They are often able to block access to content they deem hostile and arrest or harass adversaries, while those supporting government policies are given greater latitude to communicate freely.

⁷³ Robert Faris et al., "Structure and Discourse: Mapping the Networked Public Sphere in the Arab Region," Arab Networked Public Sphere, <http://www.arabnps.org/files/2016/03/ArabNPS.pdf>

Appendix: Research Methods

Data were collected from vantage points within each of the 45 countries; these supported the analysis to determine which of the URLs on the testing lists were subject to filtering.⁷⁴ Analytical software was then used to make a preliminary determination for each URL/country test pair based on a combination of weighted metrics about whether a website was intentionally blocked in the country. The software factored in metrics including the presence of a block page, evidence of DNS tampering, and various connection errors.

To summarize the results of our work, we evaluated each country to reflect the extent of filtering in each of the four thematic areas initially developed by the OpenNet Initiative project, in which the Berkman Klein Center was an institution partner.⁷⁵

- Pervasive filtering is defined as blocking that spans a number of categories while blocking access to a large portion of related content.
- Substantial filtering is either a medium level of filtering in at least a few categories or a low level of filtering across many categories.
- Selective filtering is either narrowly defined filtering that blocks a small number of specific websites across a few categories or filtering that targets a single category or issue.

It is important to note that our evaluations are subjective assessments based upon the quantitative data we gathered through a network of vantage points in the 45 countries. A purely quantitative evaluation of the level of filtering across each thematic area would be misleading unless we were able to accurately weigh the relative importance of each website. For example, the blocking of a global news website or social media platform tells us far more about the extent of censorship in a country than the blocking of a less prominent blog.

Testing Lists

We conducted testing using two lists in each of the countries: a global list and a local list.⁷⁶ The global list is composed of the top 1,000 most trafficked websites according to Alexa Internet, Inc., plus another 1,046 websites that are deemed by experts to be of international interest. .

Internationally relevant websites with content in English make up most of the global list. The websites on the local lists contain politically and culturally sensitive content in the predominant languages of each country that are potentially subject to filtering activity. The vast majority of content in the local lists for the MENA region (80%-90%) is in Arabic. Additionally, each local list contains websites that are relevant to the local politics or culture but are less likely to be censored

⁷⁴ In this study, the in-country vantage points were provided by ICLab (<https://iclab.org/>) and Dyn (<https://dyn.com/>). The selection of the 45 countries was determined in part by the availability of vantage points. In the future, we anticipate expanding coverage to additional countries.

⁷⁵ Ronald Deibert et al. (eds.), *Access Denied*.

⁷⁶ The global and local test lists are available on our Github: <https://github.com/berkmancenter/test-lists>

(for example, news websites that report positively on the ruling government), to provide a baseline for future comparisons. The testing lists encompass a wide variety of content that fall into one of four main themes: political content, social content, content related to conflicts and national security, and Internet tools.^{77 78}

Detection of Internet Filtering

Detecting blocking is straightforward when censors deliver a block page, but the situation is more complicated when one suspects censorship is being performed through the introduction of technical errors. In the absence of a definitive block page, we instead relied on evidence gathering.

We considered a number of conditions when making our determinations of whether or not an observed technical error qualified as filtering. First, the same error should exist across multiple, unaffiliated domains. Second, the domains for which we saw errors should not be equally distributed among content categories: one or more content categories (for instance, pornography or gambling) should contain most if not all of the errors. Third, the error should be consistent: repeated requests should result in the same error. Fourth, the error should be isolated to a single request location; requests from other network locations should succeed. Not all of the errors we judged to constitute filtering fully satisfied all these conditions, but we used our best judgment to limit false positives.

We ultimately judged errors for requests from network locations in the following countries to be indicative of purposeful filtering: China, Kazakhstan, Lebanon, Turkey, Iran, and Indonesia. The natures of the errors we encountered varied slightly, but all resulted in failed requests for content. The specific errors we received depended on the client we used to make requests and the network level at which filtering took place. The following are a sample of the error messages we received: "Operation cancelled," "Connection closed," "Connection reset," and "Empty reply from server."

⁷⁷ A list of content categories and their definitions can be found here: <https://thenetmonitor.org/internet-content-categories>

⁷⁸ The methodology for URL testing list creation was initially developed by the OpenNet Initiative (ONI) and used to conduct Internet filtering testing between 2006 and 2012. See "About ONI," *OpenNet Initiative*, <https://opennet.net/about-oni>