

Is remote measurement a better assessment of internet censorship than expert analysis? Analyzing tradeoffs for international donors and advocacy organizations of current data and methodologies

Terry Fletcher^{1,*}  and Andria Hayes-Birchler²

¹Millennium Challenge Corporation, Department of Policy and Evaluation, Washington, District of Columbia 20005, USA

²Independent Consultant for the Millennium Challenge Corporation, Department of Policy and Evaluation, Washington, District of Columbia, 20005, USA

*Corresponding author. E-mail: fletcherata@mcc.gov

Received: 24 January 2022; **Revised:** 01 December 2022; **Accepted:** 04 February 2023

Key words: expert analysis; internet censorship; internet filtering; internet shutdown; remote sensing



Abbreviations: ACLU, American Civil Liberties Union; AN, access now; DNS, domain name system; FH, freedom house; HTTP, hypertext transfer protocol; HTTPS, hypertext transfer protocol secure; IP, internet protocol; ISP, internet service provider; ITU, International Telecommunications Union; OONI, open observatory of network interference; ONI, OpenNet initiative; SGD, Sustainable Development Goal; UN, United Nations; USAID, United States agency for international development; VD, Varieties of Democracy; VPN, virtual private network; V-Dem, Varieties of Democracy

Abstract

Donor organizations and multilaterals require ways to measure progress toward the goals of creating an open internet, and condition assistance on recipient governments maintaining access to information online. Because the internet is increasingly becoming a leading tool for exchanging information, authoritarian governments around the world often seek methods to restrict citizens' access. Two of the most common methods for restricting the internet are shutting down internet access entirely and filtering specific content. We conduct a systematic literature review of articles on the measurement of internet censorship and find that little work has been done comparing the tradeoffs of using different methods to measure censorship on a global scale. We compare the tradeoffs between measuring these phenomena using expert analysis (as measured by Freedom House and V-Dem) and remote measurement with manual oversight (as measured by Access Now and the OpenNet Initiative [ONI]) for donor organizations that want to incentivize and measure good internet governance. We find that remote measurement with manual oversight is less likely to include false positives, and therefore may be more preferable for donor organizations that value verifiability. We also find that expert analysis is less likely to include false negatives, particularly for very repressive regimes in the Middle East and Central Asia and therefore these data may be preferable for advocacy organizations that want to ensure very repressive regimes are not able to avoid accountability, or organizations working primarily in these areas.

Policy Significance Statement

This research is essential for policy makers that are attempting to develop indicators to measure the prevalence of internet censorship in a country, either for the purpose of conditioning development assistance on good

  This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

governance or measuring project progress toward the goals of a freer and more open internet. Expert analyses (such as Freedom House or V-Dem) are more effective at identifying censorship in very repressive regimes where verifiable information is limited and may be more useful to civil society and advocacy organizations, while remote sensing (such as Access Now or the OpenNet initiative) data are more verifiable and objective and may be more useful to donors.

1. What is Internet Censorship?

Since the 1990s, the internet has spread around the world, reaching 3.8 billion people in three decades and fundamentally changing the way information is produced, disseminated, and consumed (Cohen-Almagor, 2013; International Telecommunication Union, 2020; Shahbaz and Funk, 2020). Policymakers, civil society, and academics have praised the internet as a tool for encouraging freedom of speech and information globally (*Reno v. ACLU*, 1997; Howard et al., 2011; Corduneanu-Huci and Hamilton, 2018; USAID, 2020). The Arab Spring in the early 2010s is often cited as an example of how the internet can help facilitate information sharing across civil society and hasten transitions to democracy (Howard et al., 2011; Roberts et al., 2011; Stepanova, 2011; Farrell, 2012). However, just as quickly as information has spread across the digital world, governments have found ways to restrict access through various forms of internet censorship (Zittrain et al., 2017; Lakshmana, 2018; Gopaldas, 2019; Zeleke, 2019; Chun-Chih and Thung-Hong, 2020).

We define internet censorship as any method used to intentionally prevent information or services from reaching users over the internet. We focus on government censorship, as censorship by internet service providers (ISPs) is rare and often directed by the government (Zittrain et al., 2017; Taye, 2020). As opposed to traditional censorship, which often involves arresting or attacking members of the media to stop content production (McColm et al., 1991; Karatnycky et al., 2003), internet censorship requires new tools from repressive governments, who often cannot stop the global production of information,¹ prevent it from entering their country, or stop their citizens from engaging with it (Clark et al., 2017; Zittrain et al., 2017). In place of these traditional methods, governments often censor the internet through internet filtering and internet shutdowns.²

1.1. Internet filtering

Internet filtering is used to restrict users' access to specific websites, domains, or IP addresses through technical blocks, including but not limited to DNS poisoning, hypertext transfer protocol (HTTP) filtering through middleboxes, and IP filtering (Zittrain et al., 2017; Yadav and Chakravarty, 2018). Governments may deploy internet filtering software themselves, or they may compel ISPs to block or filter certain content within their country (Puyosa and Chaguaceda, 2017; Zittrain et al., 2017).

Governments block content for a variety of reasons. Some governments want to restrict the flow of information by blocking e-mail, social media, or video calling services (Carsten, 2014; Zittrain et al., 2017). Other governments block online content that expresses certain political views, such as content

¹ Websites with content which is illegal throughout the world (such as child pornography, copyright infringement, and scams) are often taken down or seized by authorities (usually in the developed world) instead of being blocked (Sisario, 2010; Immigration and Customs Enforcement, 2018; Farivar and Blankstein, 2019). Taking down websites that are illegal around the world is not considered censorship in these datasets and is not the focus of this paper. This is also generally not an option for authoritarian governments as domains must be registered in the country that is taking them down (Kravets, 2012).

² There are additional methods of censorship, including raising the price of internet above market-value to restrict access (through government monopolies or excessive fees); passing restrictive laws on online activities; utilizing internet surveillance, harassment, arrests, legal action or attacks to intimidate, punish or induce self-censorship in content producers or consumers; and pressuring ISPs to engage in censorship (Zuckerman, 2010; Freedom House, 2019; Feldstein, 2021; Shen and Truex, 2021). Due to this paper's focus on comparing remote measurement to expert analysis, however, it examines only censorship methods that can be measured through remote, machine-based methods.

from opposition parties, civil society, human rights advocates, or specific minority groups (Zittrain et al., 2017; Chun-Chih and Thung-Hong, 2020; Shahbaz and Funk, 2020). Some restrict content for social, cultural, or religious reasons, such as content related to sexuality, gambling, drugs, alcohol, or other content that is perceived to be offensive (Zittrain et al., 2017). Governments may block content continuously, or only at specific times, such as around an election (Zittrain et al., 2017; Anthonio, 2020; Taye, 2020). Governments may restrict information around a specific event to rewrite history to hide repressive practices, spread disinformation, or the eliminate cultural heritage of minority groups (Anonymous, 2021; Berg, 2021; Bontridder and Poulet, 2021; Cook, 2022). They may be transparent— noting that access to certain sites is not permitted—or they may try to disguise the filtering so that the lack of access appears to be a technical problem—such as displaying “file not found” on a restricted website (Dalek et al., 2015; Zittrain et al., 2017; Taye, 2020).

In recent years, improvements in security protocols and circumvention tools have made filtering challenging. Encryption makes it difficult for censors to see which portions of a website a user is attempting to access (Clark et al., 2017; Rahimi and Gupta, 2020). Hypertext transfer protocol secure (HTTPS) in particular has made it challenging for governments to restrict certain pages without censoring the entire website (Clark et al., 2017; Zittrain et al., 2017; Rahimi and Gupta, 2020). This leads governments to restrict either the entirety of a website or none of it (e.g., all of Wikipedia or none of it, instead of just select pages). Circumvention tools like virtual private networks (VPNs) can also get around this selective filtering but are ineffective against full internet shutdowns (Al-Saqaf, 2015).

1.2. Internet shutdowns

In part due to the increasing difficulty of filtering select content, governments are more often turning to blunt censorship tools, such as dramatically slowing the speed of the internet (also known as throttling) or shutting down the entire internet (Al-Saqaf, 2015; Taye, 2020). Internet shutdowns were rare in the early 2010s (Rydzak, 2018; Subramanian, 2012; Roberts et al., 2011) but have become increasingly common (CIPESA, 2019; Taye, 2020), often occurring around specific events such as an election or large protest (Zittrain et al., 2017; Anthonio, 2020; Taye, 2020). Governments often cite concerns about violent protest or instability as a reason for shutting down the internet (Zittrain et al., 2017; Taye, 2020), although studies have demonstrated that shutting down the internet tends to increase the likelihood of violence, rather than decrease it (Rydzak, 2018, 2019). Like filtering, shutdowns may be done in a way that makes it difficult to differentiate between intentional shutdowns and technical issues. Internet shutdowns may be country-wide or targeted, so that only certain regions are shut down, and they may last only a few hours or months (Taye, 2020). Internet shutdowns are often cited as more harmful than internet filtering since they impact the entire internet economy (West, 2016; Raveendran and Leberknight, 2018; NetBlocks, 2020; Woodhams and Migliano, 2020).

Today, both internet filtering and internet shutdowns are widespread practices, with some sources estimating that some form of internet censorship currently exists in more than half of countries in the world (Bischoff, 2020; Mechkova et al., 2020). Internet filtering has been widespread for many years, but the number of internet shutdowns has increased dramatically each year since the mid-2010s (Zittrain et al., 2017; Selva, 2019). Estimating the exact number of governments that utilize internet filtering or internet shutdowns is challenging, since many governments attempt to hide or disguise their internet censorship, and technical failures can be mistaken for censorship (Crandall et al., 2015; Gueorguiev et al., 2017; Pearce et al., 2018; VanderSloot et al., 2018). This paper explores two main methods of measuring internet censorship—expert analysis and remote measurement—and examines the pros and cons of each. We compare the findings from four of the most accessible datasets on internet censorship and discuss the tradeoffs faced by policy-makers, civil society, and academics that use these data.

2. Measuring Internet Censorship

Government censorship of the internet is inherently focused on the removal and obfuscation of information. Governments often work to hide both the content of the internet from their citizens and the

methods they are using to hide that content (Gueorguiev et al., 2017; VanderSloot et al., 2018). This means that measurement of internet censorship can be both challenging and dangerous as governments can target citizens that are attempting to uncover censorship (Crandall et al., 2015; Narayanan and Zevenbergen, 2015; Pearce et al., 2018; VanderSloot et al., 2018; Weinberg, 2018). However, having accurate measures of internet censorship is important for a range of stakeholders, including users attempting to subvert it, academics attempting to better understand it, and donors or advocates attempting to address it or incentivize policies that limit it.

2.1. Literature review

Despite a need for accurate measures of internet censorship, we find that almost no work has been done empirically comparing the consistency of methodologies for measuring internet censorship. We conduct a systematic review of the literature on internet censorship using Google Scholar. We search the full text of all articles containing the terms “internet censorship,” “internet filtering,” or “internet shutdowns” and choose four datasets that are among the most often cited: Freedom House’s Freedom of the Net; Varieties of Democracy’s Digital Society Project (V-Dem); OpenNet Initiative (ONI); and Access Now’s #KeepItOn data. We focus on datasets that may be useful to donor and advocacy organizations that prioritize public and accessible data with broad country coverage as described in Section 2.2. Other datasets identified in the literature include data from Howard et al. (2011), Censored Planet (2020), Open Observatory of Network Interference (2020), and ICLab (Niaki et al., 2020).

We then repeat the search with the same terms as well as the name of each dataset: searching each combination of one of the initial search terms (internet censorship, internet filtering, and internet censorship) and each dataset (“Freedom on the Net,” “V-Dem” OR “Varieties of Democracy,” “OpenNet Initiative,” and “Access Now”). Then we repeat the process with each possible combination of datasets with each of the initial search terms to identify other works comparing these datasets. We review all articles that include more than one dataset and one of the initial search terms somewhere in the full text to determine whether these works are comparing measurement methods between these datasets. Table 1 depicts the findings of each of these searches, as well as the results of searches from an earlier iteration of this analysis conducted in July 2020 (Fletcher and Hayes-Birchler, 2020).

We find that at least one of these datasets is featured in 23% of works including the words “internet shutdowns,” 18% of the works including the words “Internet filtering” and 12% of works including the words “internet censorship,” indicating that these datasets are widely used in the literature. In the case of the Freedom House and V-Dem datasets, often the articles are citing qualitative results from their reports or using variables other than those that measure internet censorship (this is the case with the only paper

Table 1. Number of google scholar results on internet censorship, filtering, and shutdowns by dataset

	Internet shutdown		Internet censorship		Internet filtering	
	Jan 2022	July 2020	Jan 2022	July 2020	Jan 2022	July 2020
All articles	1,530	813	15,100	12,200	8,910	8,770
Any dataset	347	210	1,800	1,550	1,620	1,460
Freedom on the Net	132	72	754	613	400	340
V-Dem	38	6	112	70	49	13
Access Now	160	76	174	125	94	72
OpenNet Initiative	128	108	1,040	972	1,310	1,220
Two datasets	127	66	316	248	270	198
Three datasets	15	6	21	10	19	9
All four datasets	1	0	1	0	1	0

Note. Few articles reference one or more dataset, and those that do rarely compare datasets and never compare overall methodologies.

that cites all four datasets: Joshi (2021)). Despite the wide use of these data, we find that only two works compare the results of two of these datasets (Frantz et al., 2020 compare ONI and V-Dem, and Feldstein, 2021 compare V-Dem and Access Now), and no articles compare any three or all four.

Frantz et al. (2020) compare V-Dem's data on internet shutdowns and internet filtering with three other datasets, two datasets produced for specific articles (Howard et al., 2011; Rydzak, 2018), and ONI. The Howard et al. (2011) dataset is largely an expert analysis created using a database of historical news sources and interviews with experts. The Rydzak dataset is a hybrid dataset gathered from expert reviews of historical events and remotely sensed data from Google's Traffic Disruptions tool, Oracle's Internet Intelligence, and Access Now. Unlike this paper, Frantz et al. (2020) do not convert scores into binary variables, but rather look at the correlations of the individual scores. Frantz finds that V-Dem is somewhat correlated with the Howard et al. (2011) and Rydzak datasets, but minimally correlated with ONI data.

Feldstein (2021) briefly compares three datasets on internet shutdowns from V-Dem, Access Now, and NetBlocks (Woodhams and Migliano, 2020); however, his primary concern is with the different kinds of information provided by these datasets as opposed to testing these datasets when they ask the same questions. He notes that the length of a shutdown (as measured by Access Now) is not necessarily predictive of its cost (as measured by NetBlocks) or its political impact (as measured by V-Dem). He concludes that Access Now's data may be less effective at describing the depth of impact of a shutdown compared with the other two datasets but does not comment on their comparative accuracy when measuring whether a shutdown took place.

While he does not directly compare the datasets or concepts focused on here, Kawerau (2021) does compare remote sensing and expert analysis with respect to cyber-security. Specifically, he compares V-Dem's data on "Does the government have sufficiently technologically skilled staff and resources to mitigate harm from cyber-security threats?" with remotely sensed data on the susceptibility of IP addresses to cybersecurity threats from Shodan.io using the Common Vulnerability Scoring System (CVSS). Kawerau finds that while there is a statistically significant relationship between these two measurements, V-Dem finds more vulnerability than the remote measurement-based indicator.

As an aside, it is notable that the literature appears to be tracking the shift from internet filtering to internet shutdowns as a method of internet censorship. The number of articles featuring the phrase "internet shutdowns" has nearly doubled in the last year and a half increasing by over 700 articles, while the number of articles on internet filtering has barely moved, increasing by fewer than 200 articles. Additionally, datasets that continue to update (Freedom House, V-Dem, and Access Now) have continued to see stronger growth in their usage, despite ONI still being used in more articles overall.

2.2. Consumers and producers of internet censorship data

We find there are many reasons consumers seek data on which governments censor the internet. Academics have an interest in understanding trends in internet censorship and its relationship to other phenomena (e.g., Howard et al., 2011; Freyburg and Garbe, 2018; Sagir and Varlioglu, 2020; Sutterlin, 2020). Some consumers are technical experts and internet users working to circumvent censorship practices (e.g., Roberts et al., 2011; Leberknight et al., 2012; Al-Saqaf, 2015). Other consumers are advocacy or donor organizations that use the data to pressure governments to stop internet censorship (e.g., Millennium Challenge Corporation, 2019; Parks and Thompson, 2020; Sayadi and Taye, 2020; SK, 2020).

An indicator measuring internet censorship might be of use to assess the United Nation's (UN) Sustainable Development Goals (SDGs), particularly target 16.10 "Ensure public access to information and protect fundamental freedoms, in accordance with national legislation and international agreements" (UN General Assembly, 2015). The current targets for this goal do not include any reference to internet access, rather focusing on illegal attacks on journalists and freedom of public information laws (UN General Assembly, 2021). Incorporating a target focused on internet censorship could better align this indicator with its purported focus on fundamental freedoms, update this indicator for a world where more and more people get their information from online sources, and address concerns that have dogged

the SDGs from their inception that they fail to promote accountable, transparent, and democratic governance (Smith, 2015; Smith and Gladstein, 2018; International Telecommunication Union, 2020).

While all consumers of these data value an accurate reflection of the world, they may place more or less value on other characteristics of a dataset. An academic researcher may value a dataset that includes many years of historical data for the purpose of running regressions. Several academic papers reviewed that use these data focus on the creation of historic datasets that may look decades backward (Howard et al., 2011; Rydzak, 2018; Frantz et al., 2020). A user of circumvention tools in an authoritarian country might value data that is constantly updated. Some tools are particularly geared for these users, such as OONI or NetBlocks which are updated on a daily or even hourly basis (Open Observatory of Network Interference, 2020; NetBlocks, 2022b). Some donors are interested in a dataset with broad country coverage, publicly available and accessible data, and measurements explicitly linked to governance (Millennium Challenge Corporation, 2019; USAID, 2019; Tilley and Jenkins, 2020; Fletcher, 2021; Carneades, 2022). Donors and advocates can use these data to target development aid to countries that are pursuing good governance reforms, identify countries that may need to reform their internet regulations, or pressure autocratic regimes to reduce oppression (Wagner, 2013). We focus on datasets and criteria of interest to global donor and advocacy organizations.

We find that there are two broad methods of measuring internet censorship referenced in the literature: expert analysis and remote measurement. We define expert analysis as a process where one or more experts answer specific questions, which are used to create quantitative scores about internet censorship in a country. Remote measurement uses software and user reports to sense and catalog specific censorship events, often with human oversight. The datasets we use from Freedom House and V-Dem are expert analyses (Freedom House, 2019; Pemstein et al., 2020). The datasets we use from ONI and Access Now are remotely measured (Faris and Villeneuve, 2008; Access Now, 2017). While some work has been done to compare individual datasets or new tools with existing tools for validity (Frantz et al., 2020; Raman et al., 2020; Feldstein, 2021), we find no work comparing methodologies as we do here and in Fletcher and Hayes-Birchler (2020).

2.2.1. *Expert analysis*

The methodology for expert analyses involves periodically surveying one or more experts and aggregating that information into a quantitative measure. These analyses are published regularly, usually on an annual basis. Sometimes they include disaggregated data on certain responses or narratives explaining the rationale for score changes. These data are used by researchers to provide a general context for country policy environments (e.g., Maréchal, 2017), variables in regressions (e.g., Sagir and Varlioglu, 2020), and to determine funding and incentivize reform (Millennium Challenge Corporation, 2019). While the reports produced by these organizations can provide helpful context for censorship, there are some drawbacks to expert analyses, which do not document specific events nor provide information as to exactly how the internet was censored in particular instances (Roberts et al., 2011). Given that these datasets are only produced once a year, they are less useful for users attempting to actively circumvent government censorship in real time.

Examples of expert analyses with questions on internet censorship include Freedom House's Freedom on the Net report (which includes the Key Internet Controls report), Reporters Without Borders' Press Freedom Index, and V-Dem's Digital Society Project. In this report, we focus on V-Dem and Freedom House's Key Internet Controls, as these have disaggregated data, which examines the same questions on internet censorship. These two datasets use different methods of expert analysis: Freedom House trains a single expert or organization in their methodology and how to create a narrative report (Freedom House, 2019). The expert answers specific questions on over 100 issues, and then meets with other experts to normalize ratings around the world (Freedom House, 2019). V-Dem surveys multiple experts and then aggregates their responses into a single score (Pemstein et al., 2020). V-Dem uses bridge coding, overlap coding, and anchoring vignettes, where experts code the same situations to normalize responses to a common scale (Coppedge et al., 2021). Other expert analyses are focused on creating historical datasets for use by researchers (e.g., Howard et al., 2011). These are not particularly useful to donors or civil society due to the fact that they are not regularly updated (Millennium Challenge Corporation, 2019).

2.2.2. Remote measurement

Remote measurement of internet censorship involves sensing and cataloging specific instances of censorship (such as certain pages that were blocked or moments when the internet was shut down in a particular place). We divide remote measurement into three categories: no oversight, manual oversight, and automated oversight. No oversight methods generally involve a program testing for a particular type of censorship in each country, with the raw data being made available for use by other researchers. Examples include OONI, which uses software installed on computers of volunteers around the world to sense censorship instances, or ICLab (Niaki et al., 2020; Open Observatory of Network Interference, 2020). However, without some degree of oversight, the raw data produced by these methods are prone to false positives, false negatives, and other technical challenges (Weaver et al., 2009; Pearce et al., 2018; Weinberg, 2018; Yadav and Chakravarty, 2018).³ Yet these data can be useful for users attempting to circumvent censorship because they are published daily or hourly and users can ground truth any potential false positives/negatives themselves.

To mitigate these challenges, many datasets turn to some type of oversight. Manual oversight methods are those which involve some level of human testing or aggregation of instances of censorship. This may involve a machine identifying a possible instance of censorship and a human checking to see if it can be confirmed, or a human reviewing a series of automated tests and aggregating them into a single score. The two remotely measured datasets reviewed here both use manual oversight. ONI has volunteers download software on their computers that test a list of potentially censored pages. These automated results are then reviewed by humans and aggregated into a score for each of four policy categories (Faris and Villeneuve, 2008). Access Now uses both volunteer reports and machine sensing methods to detect potential shutdowns and then uses local volunteers, internet companies, and the media to manually confirm shutdowns (Access Now, 2017). Unlike the other datasets considered here, Access Now publishes their data at the level of individual instances of censorship (e.g., a specific internet shutdown), with additional information about the context for that instance (Taye, 2020).

A comparatively new method for detecting internet censorship includes both automated sensing and oversight, where various methods are used to alleviate the challenges of automated remote sensing without requiring human oversight or in-country volunteers (Sfakianakis et al., 2011; Pearce et al., 2018; VanderSloot et al., 2018; Weinberg, 2018; Hoang et al., 2019; Raman et al., 2020). These methods are lauded as being more efficient and ethical, as they do not endanger in-country volunteers who may be subject to government reprisals for assisting in identifying government censorship (Crandall et al., 2015; Pearce et al., 2018; VanderSloot et al., 2018). However, despite the promise for academics and users of circumvention tools, the current forms of these data are too inaccessible and disaggregated to be useful to donors or advocates, and therefore we do not include any in our analysis.

3. Comparing Censorship Data

Given the importance of these data for researchers, donors, policymakers, and civil society, it is vital they be as accurate as possible. Without omniscience, we cannot know whether any of these data are perfectly accurate (in that they capture all and only instances of internet censorship). However, it is possible to assess the likelihood that datasets include false positives (they capture censorship that did not actually occur) or false negatives (they do not capture censorship when it occurs)⁴ by examining their methodology, as well as comparing how often and where they agree or disagree with one another. While we

³ Note that some of these organizations publish reports on instances of censorship that are confirmed with oversight (e.g., Basso et al., 2022; Filastò et al., 2022; Netblocks, 2022a), but these reports do not claim to be a systematic or complete categorization of all instances of censorship occurring or detected.

⁴ False negatives occur when a dataset covers a country but does not identify any of the censorship that actually occurs in a given year. This is distinct from not covering a country. Note that, for this analysis, we are focusing on whether a country government censored the internet in a particular way in a given year. Some methods may be prone to false negatives at lower levels that do not impact the overall assessment of whether a country censors the internet in a particular year (that is a remote system may not detect censorship on every test of a network, or may not catch every instance of censorship in a given year, but might detect censorship on other tests or instances, these false negatives are not counted).

cannot determine with certainty whether false positives or false negatives occur in any given dataset, the findings from our empirical analyses, combined with each dataset's methodology, and our broader literature review, all suggest that remotely measured data with manual oversight are less likely to contain false positives, but may be more vulnerable to false negatives. Conversely, some expert analyses appear more likely to include false positives but may be less vulnerable to false negatives. Recognizing this may help consumers of these data identify tradeoffs when selecting which datasets to utilize. In [Section 4](#) we discuss some potential explanations for these differences, while in [Section 5](#) we discuss the implications for policymakers and other users of these data.

3.1. *Methods*

In order to compare datasets, we focus on three concepts covered by multiple datasets: (1) did a country's government filter political content on the internet in a given year?, (2) did a country's government block social media in a given year?, and (3) did a country's government shut down the internet in a given year? The exact questions asked by each source, as well as the scales used to score them, are described in [Table 2](#). As noted in [Section 3.3](#), we conduct sensitivity testing on the process of converting these scores into binary values. We use Stata to compare answers from each dataset for the same countries and years. We use a binary comparison instead of taking regressions of the original data because this allows us to isolate the more objective question of whether or not a government censored the internet from other questions of depth or pervasiveness of censorship which may be more subjective or dependent on the scale of the data. Additionally, simplified data can be more useful to donor organizations that are trying to simply answer the question of whether or not a government censored the internet in a given year.

Due to a lack of overlap in the years and concepts covered by these datasets, it is not possible to compare all variables across all datasets. V-Dem is the only dataset that overlaps temporally with ONI, but ONI does not contain any measure of internet shutdowns (Coppedge, 2020). Therefore V-Dem and ONI are compared on the two internet filtering questions (political content and social media) from 2007–2012. Freedom House, V-Dem, and Access Now overlap temporally from 2016–2020, but Access Now does not contain information on the filtering of political content (note that this is an update to the data published in Fletcher and Hayes-Birchler, 2020). Therefore, Freedom House, V-Dem, and Access Now are compared on the concepts of social media blockages and internet shutdowns from 2016–2020. For all comparisons, only countries and years covered by all datasets are included. This includes 325 observations compared between Freedom House, V-Dem, and Access Now and 74 observations compared between V-Dem and ONI.

Since these datasets are on different scales, we first convert all of the scores into binary yes/no responses for the three questions, except for Freedom House's dataset, which is already binary. In the Access Now dataset, any country listed as having a "full-shutdown" in a particular year is counted as shutting down the internet. If a country is listed as having a "service-based" shutdown it is counted as having blocked social media (Access Now, 2017). If a country is listed as having "full and service-based shutdowns" this indicates the government both shutdown the internet for some location or period of time and also blocked social media at another location or period of time; as such, it is counted in both categories. In order to convert V-Dem data to binary values we consider any response other than "Never or almost never" as censorship occurring in the country. Similarly, for ONI, any score other than "No evidence of filtering" is counted as censorship occurring in the country.

We then compare these binary scores across each relevant dataset to determine whether responses for each variable are the same across datasets. In other words, if V-Dem states that a given country filtered political content, blocked social media, and shutdown the internet in a given year, do the other datasets agree with this assessment? Where datasets disagree, we then categorize how often each dataset uniquely identified instances of censorship (indicating potential false positives) and how often each dataset uniquely identified instances of non-censorship (indicating potential false negatives).

The literature suggests that remote measurement without oversight is likely to result in both false positives and false negatives (Weaver et al., 2009; Pearce et al., 2018; Weinberg, 2018; Yadav and

Table 2. *Data sources for analysis*

Institution	Freedom house	V-Dem	ONI	Access now
Dataset	Key Internet Controls	Digital Society Project	ONI Censorship Data	#KeepItOn
Collection method	Expert Analysis	Expert Analysis	Remote Measurement	Remote Measurement
Years covered	2015–2020	2000–2020 (pre-2018 data is retroactive)	2007–2013	2016–2020
Countries covered	65 every year	173 every year (174 after 2011 with South Sudan). Includes all countries covered by all other datasets	74 observations total. Between 3 and 37 observations per year	All countries are covered every year. Only countries with censorship are listed
Internet filtering question (political content)	(Freedom House’s data contains a question on this, but it is not used in our analysis as it does not overlap temporally with a remotely measured dataset that asks this question)	How frequently does the government censor political information (text, audio, images, or video) on the Internet by filtering (blocking access to certain websites)?	How much does the government censor web sites that express views in opposition to those of the current government	N/A
Internet filtering question (social media)	Are entire apps or key functions of social media, messaging, and calling platforms temporarily or permanently blocked to prevent communication and information sharing?	How often does the government shut down access to social media platforms?	How much does the government censor web sites that provide e-mail, Internet hosting, search, translation, Voice-over Internet Protocol (VoIP) telephone service, and circumvention methods	“Service-based shutdowns” where a government blocks social media or communication platforms
			N/A	

(Continued)

Table 2. Continued

Institution	Freedom house	V-Dem	ONI	Access now
Internet Shutdown question	Does the government intentionally disrupt the internet or cellphone networks in response to political or social events, whether temporary or long term, localized or nationwide?	How often does the government shut down domestic access to the Internet?		Full shutdowns of the internet or cellphone networks, whether temporary or long term, and localized or nationwide
Scoring scale	Yes/No	Ordinal Data used: 0–4 with 0 meaning “Extremely Often” and 4 meaning “Never or almost never.”	0–4 with 0 meaning “No evidence of internet filtering” and 4 meaning “pervasive internet filtering”	Lists specific instances of filtering or shutdowns

Note. The datasets ask nearly identical questions on internet filtering and internet shutdowns, but measure these questions using very different methodologies.
Abbreviation: ONI, OpenNet initiative.

Chakravarty, 2018). However, we anticipate that the remotely measured datasets we examine will result in fewer false positives than remote measurement without oversight due to the manual oversight and emphasis on verifiability in their methods. In an attempt to guard against the false positives common to the automated elements of remote measurement, these datasets establish for themselves a burden of proof to verify specific instance of filtering or shutdowns. The same burden of proof does not apply to the expert analysis methodology. We anticipate that this burden of proof may result in more false negatives in these remote measured datasets, as they may believe a country is censoring its internet but cannot verify it and therefore do not count it.

3.2. Findings

Our findings support the hypothesis that remotely measured datasets are less likely to contain false positives than expert analyses. Table 3 depicts the findings of the analysis of Freedom House (FH), V-Dem (VD), and Access Now (AN) on the concepts of social media blockage and internet shutdown. For each column, each cell is mutually exclusive and all the rows in a column are jointly exhaustive.

While all three datasets agreed as to whether a country shut down the internet in a given year in the majority of cases (64%), they disagree in at least a third of cases. The disagreement is more pronounced for social media blockages, where the three sources agree only slightly more than half the time (58.15%). For both concepts, Access Now has the fewest instances (1.23% and 0.62%) of uniquely identifying censorship. The fact that in 99% of cases for both types of censorship, at least one of the expert analyses agrees with Access Now in identifying censorship—combined with the verifiable evidence Access Now publishes for each occurrence—indicates comparatively few false positives in the Access Now dataset (i.e., if there is a false positive in the Access Now dataset, it is unlikely it could be avoided by using an expert analysis instead).

Conversely, V-Dem is the sole dataset to identify censorship in 19.38% of instances for internet shutdowns and 14.15% of instances for social media blockages. This combined with the lack of verifiable evidence for its scores may indicate a higher rate of false positives. It may alternatively suggest that V-Dem finds instances of censorship missed by other datasets, and the other datasets include some false negatives. An analysis comparing V-Dem's expert analysis and remote measurement of data on cyber security in Kawerau (2021) finds similarly that V-Dem was more likely to identify that there are vulnerabilities than a remote measurement system.

To investigate the issue of false negatives, we examine the cases where a dataset was the only one *not* to list a country as censoring the internet. As hypothesized, Access Now is the most likely to omit a country from its list of censors when both other datasets find that censorship occurred (9.23% of cases for internet shutdowns and 17.23% for social media blockages). V-Dem and Freedom House each have low shares of

Table 3. Freedom house, V-Dem, & access now comparison

	Internet shutdowns (%)	Social media blockages (%)
FH only finds as censorship	2.77	6.15
VD only finds as censorship	19.38	14.15
AN only finds as censorship	1.23	0.62
FH and VD find censorship but AN does not	9.23	17.23
FH and AN find censorship but VD does not	1.85	1.54
VD and AN find censorship but FH does not	1.54	2.15
All datasets identify Censorship	13.23	8.92
All datasets identify No Censorship	50.77	49.23

Note. The expert analyses (particularly V-Dem) are more likely to uniquely identify censorship, while the remotely measured data is more likely to uniquely identify non-censorship. Comparisons are based on 325 country/year-level observations.

Table 4. *V-Dem & ONI comparison*

	Political filtering (%)	Social Media filtering (%)
V-Dem only finds censorship	30.14	19.18
ONI only finds censorship	2.74	8.22
Both datasets find censorship	42.47	24.66
Neither dataset finds censorship	24.66	47.95

Note. V-Dem uniquely identifies censorship much more often than ONI for both political and social media filtering. Comparisons are based on 74 country/year-level observations.

Abbreviation: ONI, OpenNet initiative.

cases where they uniquely identified a country as not censoring the internet (1.9% and 1.5% for V-Dem, and 1.5% and 2.15% for Freedom House), indicating potentially fewer false negatives.

Table 4 presents the findings of the analysis of V-Dem and ONI on the two issues they both covered: social media blockages and political filtering. The two data sources agree in over two-thirds of cases, but once again V-Dem is more likely to uniquely identify censorship (in 19.8% of cases for social media blockages and 30.14% of cases for political filtering). Although there are also cases where ONI uniquely identified censorship, they were far less frequent (8.22% of cases for social media blockages and 2.74% of cases for political filtering.)

3.3. Sensitivity testing

To assess the robustness of these conclusions, we run three sensitivity tests to determine how much our conclusions are dependent on the underlying assumptions we have made. First, we test a second model of V-Dem treating both “rarely” and “never or almost never” as instances of non-censorship. Second, because the Freedom House dataset is offset from the calendar year, we match the exact months of the Freedom House dataset with the exact months of the Access Now dataset for one year. Finally, we compare just V-Dem and Access Now (with 895 observations) to determine whether the small sample size may be skewing the data.

V-Dem aggregates responses from many experts on a five-point scale. In our initial analysis, we only considered instances of censorship whenever the ordinal response was in the highest category: “Never or almost never”. However, respondents may be prone to central tendency bias, making them less likely to select the extreme values even when they have no evidence of censorship. We therefore re-ran the same analyses but considered both “rarely” and “never or almost never” as non-censorship in V-Dem’s data. Under this model, V-Dem uniquely identifies many fewer instances of censorship, but in three of the four comparisons, still uniquely identified more instances of censorship than the remote sensing dataset. Only comparing ONI and V-Dem on social media filtering with this alternative model did V-Dem uniquely identify fewer instances of censorship: V-Dem uniquely identified 5.5%, while ONI uniquely identified 8.2%. It is notable that using this alternative model for V-Dem increased the overall agreement of all the datasets in all four comparisons by anywhere from 6% to 14%. This indicates that V-Dem might benefit from instituting some controls for central tendency bias in their data to avoid mislabeling instances of non-censorship as censorship, as this puts it more in line with the results of other datasets, but as no such measures have been implemented, and these data are still documenting at least some censorship, our general conclusions still hold.

Freedom House’s data is offset from the calendar year. Since censorship events can last several months, (Taye, 2020) this may not impact the comparison substantially, but to check the impacts of this offset, we match the months of the Freedom House report exactly using the Access Now data for the 2017 Freedom House data. Using these parameters, Access Now identifies shutdowns that Freedom House does not in 6% of cases (compared to 3% in the original model), while Freedom House identifies shutdowns that Access Now does not 12.3% of the time (compared to 12% in the original model). In terms of social media blockages, Access Now identifies blockages that Freedom House does not in 4.6% of cases (compared to

2.8% in the original), and Freedom House identifies blockages that Access Now does not in 27.8% of cases (compared to 23.4% in the original model). Given that there is minimal variation and Access Now continues to identify censorship much less frequently, the differences in these datasets are not wholly due to the differences in time periods covered and would be unlikely to change our conclusion.

Finally, the sample size of these data is somewhat small due to the need to have overlapping time periods and countries. We test the impact of this small sample size by also comparing just V-Dem and Access Now, using all 895 observations shared by both datasets. We find that V-Dem continues to uniquely identify instances of censorship more than Access Now. V-Dem identifies shutdowns that Access Now does not in 24.24% of all cases using this larger dataset, while Access Now uniquely identifies shutdowns in 1.67% of all cases. In terms of social media, V-Dem identifies blockages that Access Now does not in 24.58% of cases, while Access Now identifies blockages V-Dem does not in just 1.01% of cases. All of these sensitivity tests support the initial conclusion that generally expert analyses are more prone to over-identifying censorship that may not exist, while remote sensing is more prone to under-identifying censorship when it may actually exist.

4. Analysis

While we cannot determine conclusively which dataset is most accurate, our findings suggest that remote measurement with human oversight results in fewer false positives than expert analysis, although it may be more vulnerable to false negatives. Conversely, our analysis also suggests that expert analyses, V-Dem in particular, include more false positives than remote measurement, though they may provide a more complete picture of internet censorship.

4.1. Individual case analysis

At least some of the instances of censorship identified by the expert analyses but not by Access Now appear to be accurate, as they are confirmed by other sources, such as a social media blockage in Kazakhstan in 2019, or social media blockages in the U.A.E. in 2020 (Human Rights Watch, 2020; Turak, 2020). While others such as Cuba's 2020 internet shutdown or Libya's 2020 internet shutdown, cannot be verified by any other source and so may well be false positives from the expert analyses.

One explanation for the apparent false negatives found in remote measurement data could be that these methods are constrained by the number of in-country volunteers or journalists who can manually confirm each instance of censorship. In countries with significant limitations on civil liberties or press freedoms journalists and civil society organizations may lack the capacity to confirm censorship. This is borne out in the data, as the majority of censorship identified by Freedom House and V-Dem but not Access Now come from North Africa, the Middle East, and Central Asia. These regions have some of the strictest limitations on the media in the world (Reporters Without Borders, 2020). Of the 30 instances of shutdowns identified by Freedom House and V-Dem, 20 are in one of these regions. Azerbaijan, Kazakhstan, and Uzbekistan, alone account for nine of these instances, with Kazakhstan being identified as shutting down the internet by both expert analyses but not Access Now in 2016, 2017, 2019, and 2020. A similar pattern can be seen in the social media blockages. Of the 56 instances of social media blockages identified by the two expert analyses but not Access Now, 30 are in the Middle East or Central Asia.

There are several possible explanations for the many instances of censorship that are identified by V-Dem, but no other datasets, beyond the concerns about central tendency bias that have already been discussed. Looking at news reports from the countries where V-Dem uniquely identifies internet censorship, it may also be the case that some experts are conflating social media blockages with full internet shutdowns. This is supported by instances such as Venezuela in 2019, where, according to news reports, the government blocked access to social media platforms, but did not shut down the internet (Gold, 2019), but V-Dem identifies this as a shutdown. Saudi Arabia (Dahan, 2019), and Cuba (Amnesty International, 2017) appear to be in the same situation. Experts may also conflate civil liberties in general with internet censorship as appears to have happened in the Philippines where V-Dem identifies

a shutdown despite no other evidence for one (Engagemedia and Sinar Project, 2018). This would explain the high degree of internal correlation in V-Dem's indicators noted by Frantz et al. (2020): if experts are conflating different types of censorship, the data may be picking up on the general level of freedom in a country instead of the specific question of an internet or social media shut down. This would explain why V-Dem's data shows that shutdowns were almost as prevalent in 2000 as they are today, in spite of the literature suggesting they were very rare before 2011 (Roberts et al., 2011; Subramanian, 2012; CIPESA, 2019) and very prevalent today (Rydzak, 2018; Taye, 2020).

However, there do appear to be instances where V-Dem is accurately picking up on censorship that is not captured in the other two datasets, such as in Lebanon where independent news sources confirm that social media apps were blocked in 2019, but only V-Dem identified censorship (Hall, 2019). This might also be the case in Rwanda where some sources report government control of social media and communication apps in the lead-up to the 2017 elections (AFP, 2017; McDevitt, 2017) and only V-Dem identified social media blockages, which suggests that instances of V-Dem uniquely identifying censorship are likely a mix of false positives *and* V-Dem picking up on censorship that the other datasets miss.

A similar pattern emerges with Freedom House's data. Though there are fewer cases of censorship uniquely identified by Freedom House than V-Dem, both still identify many more than Access Now. There are several potential explanations for this. Freedom House may be more willing to call something government censorship even if it is only suspected but not corroborated. For example, shutdowns in Zambia in 2016 and 2020, were captured only by Freedom House, and the evidence is inconclusive if they were intentional or simple technical outages (Gambanga, 2016; Mwango, 2020). In other situations, Freedom House may be conflating the shutdowns of other media with internet shutdowns, such as when it was the only source to document an internet shutdown in the Philippines in 2020 (there was no evidence of an internet shutdown, but the government did shut down the major broadcasting network and phone communications (Ballaran, 2018; Ramos, 2020)). Some of the differences may arise from which platforms the datasets classify as "social media," with Freedom House including smaller sites like Telegram and Medium (leading to Freedom House uniquely identifying censorship in Bahrain and Malaysia (Marczak, 2016; Freedom House, 2017), which other datasets might classify as messaging apps (in the case of Telegram) or news sites (in the case of Medium). While there are fewer instances than V-Dem, Freedom House's uniquely identified censorship appears to also be a mix of false positives/uncorroborated censorship, and accurately identified cases of censorship that other datasets miss or would classify differently. One advantage of Freedom House and Access Now over V-Dem is that they include citations of these instances, even if Freedom House has a lower bar for confirming censorship than Access Now.

4.2. Limitations

There are several limitations to the methodology used in this paper. It is possible, and even likely that the expert identified by Freedom House is one of the experts that fill out V-Dem's survey, given the limited number of country experts for certain countries. However, given the fact that V-Dem surveys over 20,000 experts each year, including around 100–200 experts per country, it is unlikely that a single expert that fills out Freedom House's survey would shift the final V-Dem results substantially (Coppedge et al., 2021). However, there is a potential that experts might communicate with each other and form similar opinions about a given country, increasing the correlation between these surveys.

The internet has changed rapidly since it was created and will likely continue to change (Cohen-Almagor, 2013). This means that some of our results may be dependent on the particular years that are being measured. As noted above, many of V-Dem's likely false positives are from older years where V-Dem found internet shutdowns, but experts agree that they were not particularly prevalent (Roberts et al., 2011; Subramanian, 2012; CIPESA, 2019). This relates to the concern that, given that there are only two datasets for each method, these conclusions may not be generalizable. As the internet continues to change, and technology for censoring and detecting censorship changes, these findings may not continue

to hold. However, these conclusions are valid for donor organizations making decisions right now, and for the current universe of available datasets on internet censorship.

Further, as noted above, this analysis only captures certain components of internet censorship, it does not capture components such as a government requesting that a social media company self-censor, take down particular types of content or posts themselves, or make them unavailable in a particular country (Zittrain et al., 2017). Self-censorship is challenging to measure, and a review of the policies of social media companies may be more effective to measure this type of censorship than either remote sensing or expert analysis, but more research is needed (Zhong et al., 2017; Zittrain et al., 2017; Shen and Truex, 2021).

5. Conclusion

There are a range of considerations that go into determining which dataset is the best measure of internet censorship for a given purpose. While perfect accuracy is desirable, given the high levels of disagreement it is unlikely that any one dataset is completely accurate. Consumers of these data should therefore consider whether they prefer a higher likelihood of false positives or false negatives. Would they rather falsely accuse a country of censoring the internet when it does not, or allow a country that does censor the internet to claim that its internet is free unopposed?

Some consumers may prefer that their data be fully verifiable: that there should be a high burden of proof to show a country has censored their internet, even if that means missing some instances. Our analysis indicates that such consumers would be best served by a remotely measured dataset with manual oversight, using methods similar to Access Now or ONI, where each instance of censorship has been verified, and where an average of 96.8% of cases of censorship are supported by two or more datasets. Other consumers may prefer that as many instances of censorship are captured as possible, even if that means including some countries which may not have engaged in censorship. Given that many governments try to hide their censorship—and that many are quite sophisticated in doing so—these consumers may worry that the burden of proof of remote measurement methods with human oversight leaves them vulnerable to missing too many cases of censorship, particularly in especially repressive regimes. These consumers would be advised to use an expert analysis dataset instead, such as Freedom on the Net or V-Dem.

Based on our review of the literature, we would expect that multilateral and bilateral donor organizations may prefer data that is verifiable and therefore may be better served by remotely sensed data than expert analysis when measuring internet censorship. Honig (2019) notes an incentive for international development organizations to prefer clear measurable data to ensure accountability given the challenges of managing projects from a distance. Chattopadhyay (2016) notes the importance of verifiable data for goals like the SDGs, and the political challenges with basing such goals on subjective assessments. These organizations may also simply be influenced by political and diplomatic pressure to not give or take away development aid on the basis of subjective unverifiable assessments. In fact, in 2020 the U.S. development agency, the Millennium Challenge Corporation shifted from using Freedom House's Freedom on the Net data to the #KeepItOn data from Access Now to inform their assessment of which countries are well governed enough to receive assistance (Millennium Challenge Corporation, 2019, 2020). Other donors have expressed a desire for objective, verifiable, and actionable datasets to inform similar measures of country performance (European Commission, 2016; USAID, 2019).

Advocacy and civil society organizations, on the other hand, may prefer data that ensures all methods of censorship are captured. Organizations that are focused on advocating for fundamental freedoms may be less concerned if the indices are capturing a general climate of civil liberties (as V-Dem appears to be) since they may be working on multiple fronts and would be concerned if a country that was very restrictive was shown to be free because instances of restriction could not be validated. Additionally, organizations that work in the Middle East, North Africa, and Central Asia may prefer expert analyses that appear to capture restrictions in these areas more consistently.

There are of course other, logistical considerations that users consider when choosing a particular dataset. Users that want universal country coverage may avoid Freedom House's data, which only covers 65 countries. Users looking to create a time series may prefer V-Dem's data, as it starts in 2000. Some users may prefer data at the incident level as opposed to the country level, which makes Access Now's dataset more appealing. Others may prefer that the data include a clear index and ranking of countries to better "name and shame" to leverage policy change. Some users might prefer a dataset that goes deep on specific questions, such as Access Now on internet shutdowns, while others may be more interested in binary answers to many different questions, such as those provided by Freedom House's Key Internet Controls.

Progress has been made to create remotely measured datasets with automated oversight, which may be more accurate than either of the methods reviewed here (Pearce et al., 2018; VanderSloot et al., 2018; Weinberg, 2018; Hoang et al., 2019; Raman et al., 2020). However, the current versions of these datasets fail to meet many of the logistical considerations above. They are often too technical or disaggregated to be useful to donors and advocacy organizations. Therefore, there is an opportunity for future work in the aggregation of these, potentially more accurate, datasets into annually ranked indices of censorship that are more accessible to donor and advocacy organizations.

Acknowledgments. Thanks to Daniel Barnes, Jennifer Sturdy, Cindy Sobieski, Maité Hostetter, and Alexandra Berry for reviewing. An earlier conference version of this paper can be found here (<https://doi.org/10.5281/zenodo.7384697>). This current paper includes updates on the data and expands upon the analysis of the original conference paper.

Competing Interests. The authors are or have been employed by the Millennium Challenge Corporation (MCC). The views expressed herein are those of the authors and should not be construed as an express or implied endorsement of those views by the MCC nor the U.S. Government.

Author Contributions. Conceptualization: T.F., A.H-B.; Data curation: T.F.; Methodology: T.F., A.H-B.; Revision for publication: T.F.; Writing original draft: T.F., A.H-B. All authors approved the final submitted draft.

Data Availability Statement. Replication data and code can be found at Zenodo: <https://zenodo.org/record/7384697#.Y-GljK3MJPY>.

Funding Statement. The initial research underpinning this work was funded by the Millennium Challenge Corporation. The views expressed herein are those of the authors and should not be construed as an express or implied endorsement of those views by the MCC nor the U.S. Government.

References

- Access Now** (2017) Shutdown tracker optimization project. Available at <https://www.accessnow.org/cms/assets/uploads/2017/09/How-to-view-the-Access-Now-Internet-Shutdown-Tracker-2017.pdf> (accessed 4 May 2020).
- AFP** (2017) Diplomats concerned over Rwanda social media controls. *Daily Mail*. Available at <https://www.dailymail.co.uk/wires/afp/article-4556156/Diplomats-concerned-Rwanda-social-media-controls.html> (accessed 4 May 2020).
- Al-Saqaf W** (2015) Internet censorship circumvention tools: Escaping the control of the Syrian regime. *Media and Communication* 4(1), 39–50.
- Amnesty International** (2017) Cuba's Internet paradox: How controlled and censored internet risks Cuba's achievements in education. *Amnesty International*. Available at <https://www.amnesty.org/en/latest/news/2017/08/cubas-internet-paradox-how-controlled-and-censored-internet-risks-cubas-achievements-in-education/> (accessed 4 May 2020).
- Anonymous** (2021) You shall sing and dance: Contested 'safeguarding' of Uyghur intangible cultural heritage. *Asian Ethnicity* 22(1), 121–139. <https://doi.org/10.1080/14631369.2020.1822733>.
- Anthonio F** (2020) A Shutdown taints Togo's 2020 presidential election: What happened and what's next. *Access Now*. Available at <https://www.accessnow.org/a-shutdown-taints-togos-2020-presidential-elections-what-happened-and-whats-next/> (accessed 4 May 2020).
- Ballaran J** (2018) Group criticizes gov't move shutting down cellphone signals during events. *Inquirer.net*. Available at <https://newsinfo.inquirer.net/963786/group-criticizes-govt-move-shutting-down-cellphone-signals-during-events-signal-jamming-cell-sites-fina-media-group-shutdown> (accessed 4 May 2020).
- Basso S, Xynou M, Filastò A and Meng A** (2022) Iran blocks social media, app stores and encrypted DNS amid Mahsa Amini protests. *OOONI*. Available at <https://ooni.org/post/2022-iran-blocks-social-media-mahsa-amini-protests/> (accessed 20 October 2022).

- Berg A** (2021) Problems of digital control. The digital repression, religious persecution and genocide of the Uyghurs in Xinjiang by the Communist Republic China.
- Bischoff P** (2020) Internet censorship 2020: A global map of internet restrictions. *Comparitech*. Available at <https://www.comparitech.com/blog/vpn-privacy/internet-censorship-map/> (accessed 4 May 2020).
- Bontridder N and Poulet Y** (2021) The role of artificial intelligence in disinformation. *Data & Policy* 3, E32. <https://doi.org/10.1017/dap.2021.20>.
- Carneades** (2022) *Are All Lives Equal? Why Cost Benefit Analysis Values Rich Lives More and How Philosophy Can Fix It*. Washington, DC: Carneades.org. Available at https://www.researchgate.net/publication/361535357_Are_All_Lives_Equal_Why_Cost-Benefit_Analysis_Values_Rich_Lives_More_and_How_Philosophy_Can_Fix_It.
- Carsten P** (2014) Google's gmail blocked in China. *Reuters*. Available at <https://www.reuters.com/article/us-google-china/google-gmail-blocked-in-china-idUSKBN0K70BD20141229> (accessed 4 May 2020).
- Censored Planet** (2020) Censored planet raw data. Available at <https://censoredplanet.org/data/raw> (accessed 4 May 2020).
- Chattopadhyay S** (2016) What gets measured, gets managed challenges ahead for UN's data-driven development agenda. *Overseas Development Institute*. Available at <https://cdn.odi.org/media/documents/11230.pdf> (accessed 4 May 2020).
- Chun-Chih C and Thung-Hong L** (2020) Autocracy login: Internet censorship and civil society in the digital age. *Democratization* 27(5), 874–895. <https://doi.org/10.1080/13510347.2020.1747051>.
- CIPEA** (2019) State of Internet Freedom in Africa 2019. *Collaboration on International ICT Policy for East and Southern Africa*. Available at https://cipesa.org/?wpfb_dl=307 (accessed 4 May 2020).
- Clark J, Faris R and Jones RH** (2017) *Analyzing Accessibility of Wikipedia Projects Around the World*. Cambridge, MA: Berkman Klein Center for Internet & Society Research Publication.
- Cohen-Almagor R** (2013) Internet history. In Luppicini R (ed.), *Moral, Ethical, and Social Dilemmas in the Age of Technology: Theories and Practice*. Hershey, PA: IGI Global, pp. 19–39. <http://doi:10.4018/978-1-4666-2931-8.ch002>.
- Cook S** (2022) Countering Beijing's media manipulation. *Journal of Democracy* 33(1), 116–130. <https://doi.org/10.1353/jod.2022.0008>.
- Coppedge M, Gerring J, Knutsen CH, Lindberg SI, Teorell J, Altman D, Bernhard M, Fish MS, Glynn A, Hicken A, Luhrmann A, Marquardt KL, McMann K, Paxton P, Pemstein D, Seim B, Sigman R, Skaaning S, Staton J, Wilson S, Cornell A, Alizada N, Gastaldi L, Gjerløw H, Hindle G, Ilchenko N, Maxwell L, Mechkova V, Medzihorsky J, von Römer J, Sundström A, Tzelgov E, Wang Y, Wig T and Ziblatt D** (2020) V-Dem Dataset v10. Varieties of Democracy (V-Dem) Project. <https://doi.org/10.23696/vdemds20>.
- Coppedge M, Gerring J, Knutsen CH, Lindberg SI, Teorell J, Marquardt KL, Medzihorsky J, Pemstein D, Alizada N, Gastaldi L, Hindle G, Pernes J, von Römer J, Tzelgov E, Wang Y and Wilson S** (2021) V-Dem Methodology v11.1. Varieties of Democracy (V-Dem) Project.
- Corduneanu-Huci C and Hamilton A** (2018) Selective Control: The Political Economy of Censorship. *World Bank Group*.
- Crandall JR, Crete-Nishihata M and Knockel J** (2015) Forgive us our SYNs: Technical and ethical considerations for measuring internet filtering. In *NS ethics '15: Proceedings of the 2015 ACM SIGCOMM Workshop on Ethics in Networked Systems Research*. New York, NY: Association for Computing Machinery. Available at <https://dl.acm.org/doi/abs/10.1145/2793013.2793021>.
- Dahan N** (2019) Internet, interrupted: How network cuts are used to quell dissent in the Middle East. *Middle East Eye*. Available at <https://www.middleeasteye.net/news/internet-interrupted-how-middle-east-countries-use-network-restrictions-clamp-down-dissent> (accessed 4 May 2020).
- Dalek J, Deibert R, McKune S, Gill P, Senft A and Noor N** (2015) *Information Controls During Military Operations The Case of Yemen During the 2015 Political and Armed Conflict*. Toronto, ON: University of Toronto. Available at <https://citizenlab.ca/2015/10/information-controls-military-operations-yemen/> (accessed 4 May 2020).
- Engagemedia & Sinar Project** (2018) Internet censorship monitoring: Duterte's drug war. Available at <https://sinarproject.org/digital-rights/updates/internet-censorship-monitoring-dutertes-drug-war> (accessed 4 May 2020).
- European Commission** (2016) Methodology for country allocations: European development fund and development cooperation instrument 2014–2020.
- Faris R and Villeneuve N** (2008) Measuring global internet filtering in R. In Deibert JP, Rohozinski R and Zittrain J (eds), *Access Denied: The Practice and Policy of Global Internet Filtering*. Cambridge: MIT Press. Available at http://opennet.net/sites/opennet.net/files/Deibert_02_Ch01_005-028.pdf (accessed 4 May 2020).
- Farivar C and Blankstein A** (2019) Feds take down the 'world's largest dark web child porn marketplace. *NBC News*. Available at <https://www.nbcnews.com/news/crime-courts/feds-take-down-world-s-largest-dark-web-child-porn-n1066511> (accessed 4 May 2020).
- Farrell H** (2012) The consequences of the internet for politics. *Annual Review of Political Science* 15, 35–52.
- Feldstein S** (2021) *The Rise of Digital Repression: How Technology is Reshaping Power, Politics, and Resistance*. New York: Oxford University Press.
- Filastò A, Geybulla A and Xynou M** (2022) Azerbaijan and Armenia block TikTok amid border clashes. *OONI*. Available at <https://ooni.org/post/2022-azerbaijan-and-armenia-blocks-tiktok/> (accessed 20 October 2022).
- Fletcher T** (2021) Strengthening financial inclusion on MCC's scorecard. *Millennium Challenge Corporation*. Available at <https://www.mcc.gov/blog/entry/blog-101921-financial-inclusion> (accessed 20 October 2022).

- Fletcher T and Hayes-Birchler A** (2020) Comparing measures of internet censorship: Analyzing the tradeoffs between expert analysis and remote measurement. In *Data for Policy Conference*. <https://doi.org/10.5281/zenodo.3967398>.
- Frantz E, Kendall-Taylor A and Wright J** (2020) Digital repression in autocracies. *The Varieties of Democracy Institute*. Available at https://web.archive.org/web/20200609073358/https://www.v-dem.net/media/filer_public/18/d8/18d8fc9b-3ff3-44d6-a328-799dc0132043/digital-repression17mar.pdf (accessed 4 May 2020).
- Freedom House** (2017) Freedom on the net 2017: Malaysia. *Freedom House*. Available at <https://freedomhouse.org/country/malaysia/freedom-net/2017>.
- Freedom House** (2019) Freedom on the net research methodology. *Freedom House*. Available at <https://freedomhouse.org/reports/freedom-net/freedom-net-research-methodology> (accessed 4 May 2020).
- Freyburg T and Garbe L** (2018) Blocking the bottleneck: Internet shutdowns and ownership at election times in sub-Saharan Africa. *International Journal of Communication* 12(2018), 3896–3916.
- Gambanga N** (2016) Zambian government suspected of causing internet shutdown following outage in opposition strongholds. *MTN Zambia*. Available at <https://www.techzim.co.zw/2016/08/zambian-government-suspected-causing-internet-slowdown-shutdown-following-outage-opposition-strongholds/> (accessed 4 May 2020).
- Gold H** (2019) Information war escalates as Venezuela tries to contain uprising. *CNN Business*. Available at <https://www.cnn.com/2019/05/01/media/media-venezuela-information-war/index.html> (accessed 4 May 2020).
- Gopaldas R** (2019) Digital dictatorship versus digital democracy in Africa. SAIIA Policy Insights 75.
- Guerguiev D, Shao L and Crabtree C** (2017) Blurring the lines: Rethinking censorship under autocracy. <https://doi.org/10.13140/RG.2.2.29037.08160>.
- Hall R** (2019) Lebanon blocks Grindr in latest attack on LGBT+ community. *The Independent*. Available at <https://www.independent.co.uk/news/world/middle-east/grindr-lebanon-ban-lgbt-rights-dating-app-gay-a8933556.html> (accessed 4 May 2020).
- Hoang PN, Doreen S and Polychronakis M** (2019) Measuring I2P censorship at a global scale. In *Proceedings of the 9th USENIX Workshop on Free and Open Communications on the Internet*, Berkeley, CA: USENIX Association.
- Honig D** (2019) When reporting undermines performance: The costs of politically constrained organizational autonomy in foreign aid implementation. *International Organization* 73(1), 171–201. <https://doi.org/10.1017/S002081831800036X>.
- Howard P, Agarwal SD and Hussain MM** (2011) When do states disconnect their digital networks? Regime responses to the political uses of social media. *The Communication Review* 14, 216–232.
- Human Rights Watch** (2020) Shutting down the internet to shut up critics. Available at <https://www.hrw.org/world-report/2020/country-chapters/global-5#> (accessed 20 October 2022).
- Immigration and Customs Enforcement** (2018) Over a million websites seized in global operation. *ICE Newsroom*. Available at <https://www.ice.gov/news/releases/over-million-websites-seized-global-operation> (accessed 4 May 2020).
- International Telecommunication Union** (2020) Measuring digital development: Facts and figures. *ITU Publications*. Available at <https://www.itu.int/en/ITU-D/Statistics/Documents/facts/FactsFigures2020.pdf> (accessed 20 October 2022).
- Joshi N** (2021) Internet Shutdowns During Protests: A Practice in Digital Authoritarianism. International Master's in Security, Intelligence and Strategic Studies.
- Karatnycky A, Piano A and Puddington A** (2003) Freedom in the World: The Annual Survey of Political Rights & Civil Liberties. *Freedom House*.
- Kawerau L** (2021) Governments and the Net: Defense, Control, and Trust in the Fifth Domain [Dissertation]. Konstanz: University of Konstanz. Available at <http://kops.uni-konstanz.de/handle/123456789/55972> (accessed 20 October 2022).
- Kravets D** (2012) Uncle Sam: If It Ends in .Com, It's .Seizable. *Wired.com*. Available at <https://www.wired.com/2012/03/feds-seize-foreign-sites/> (accessed 4 May 2020).
- Lakshmana KV** (2018) A coward's political weapon: Troll armies go on settling scores. *Common Cause* 37(4), 27–29.
- Leberknight CS, Chiang M, Poor HV and Wong F** (2012) A taxonomy of internet censorship and anti-censorship.
- Marczak B** (2016) Leading Bahraini ISPs are blocking telegram traffic. *Bahrain Watch*. Available at <https://web.archive.org/web/20200320193229/https://bahrainwatch.org/blog/2016/06/28/leading-bahraini-isps-are-blocking-telegram-traffic/> (accessed 4 May 2020).
- Maréchal N** (2017) Networked authoritarianism and the geopolitics of information: Understanding Russian internet policy. *Media and Communication* 5(1), 29–41.
- McColm RB, Finn J, Payne DW, Ryan JE, Sussman LR and Zarycky G** (1991) Freedom in the world political rights & civil liberties. *Freedom House*.
- McDevitt D** (2017) Rwanda censors critical, independent media in targeted fashion. *Open Technology Fund*. Available at <https://www.opentech.fund/news/new-report-investigates-internet-censorship-during-rwandas-2017-presidential-election/> (accessed 4 May 2020).
- Mechkova V, Daniel P, Brigitte S and Steven W** (2020) Digital Society Project Dataset v2. *Varieties of Democracy (V-Dem) Project*.
- Millennium Challenge Corporation** (2019) Guide to the MCC Indicators for Fiscal Year 2020. *Millennium Challenge Corporation*. Available at <https://www.mcc.gov/resources/doc/guide-to-the-indicators-fy-2020> (accessed 4 May 2020).
- Millennium Challenge Corporation** (2020) Guide to the MCC Indicators for Fiscal Year 2021. *Millennium Challenge Corporation*. Available at <https://www.mcc.gov/resources/doc/guide-to-the-indicators-fy-2021> (accessed 20 October 2022).
- Mwango C** (2020) Southern province in internet network shut down. *Zambian Observer*. Available at <https://web.archive.org/web/20220521131238/https://zambianobserver.com/southern-province-in-internet-network-shut-down/> (accessed 4 May 2020).

- Narayanan A and Zevenbergen B** (2015) No encore for encore? Ethical questions for web-based censorship measurement. *Technology Science*. Available at <http://techscience.org/a/2015121501> (accessed 4 May 2020).
- NetBlocks** (2020) Internet cut in Ethiopia amid unrest following killing of singer. *NetBlocks Mapping Net Freedom*. Available at <https://netblocks.org/reports/internet-cut-in-ethiopia-amid-unrest-following-killing-of-singer-pA25Z28b> (accessed 20 October 2022).
- Netblocks** (2022a) Internet disrupted in Sierra Leone amid anti-government protests.
- Netblocks** (2022b) The internet observatory. Available at <https://netblocks.org/projects/observatory> (accessed 20 October 2022).
- Niaki AA, Cho S, Weinberg Z, Hoang NP, Razaghpanah A, Christin N and Gill P** (2020) ICLab: A global, longitudinal internet censorship measurement platform. In *Proceedings of the 41st IEEE Symposium on Security and Privacy*. Piscataway, NJ: Institute of Electrical and Electronics Engineers.
- Open Observatory of Network Interference** (2020) OONI data explorer. Available at <https://explorer.ooni.org/> (accessed 4 May 2020).
- Parks L and Thompson R** (2020) The slow shutdown: Information and internet regulation in Tanzania from 2010 to 2018 and impacts on online content creators. *International Journal of Communication* 14(2020), 1–21.
- Pearce P, Ensafi R, Li F, Feamster N and Paxson V** (2018) Toward continual measurement of global network-level censorship. *IEEE Security & Privacy* 16(1), 24–33. <https://doi.org/10.1109/MSP.2018.1331018>.
- Pemstein D, Marquardt KL, Tzelgov E, Wang Y, Medzhorsky J, Krusell J, Miri F and von Römer J** (2020) The V-Dem measurement model: Latent variable analysis for cross-national and cross-temporal expert-coded data. V-Dem Working Paper No. 21. 5th edition. University of Gothenburg: Varieties of Democracy Institute.
- Puyosa I and Chaguaceda A** (2017) Cinco regímenes políticos en Latinoamérica, Libertad de internet y mecanismos de control. *RETOS* 7(14), 11–37.
- Rahimi N and Gupta B** (2020) A study of the landscape of internet censorship and anti-censorship in Middle East. In *EPiC Series in Computing: Proceeding of the 35th International Conference on Computers and Their Applications*, Vol. 69, pp. 60–68. Winona, MN: International Society for Computers and their Applications.
- Raman RS, Stoll A, Dalek J, Sarabi A, Ramesh R, Scott W and Ensafi R** (2020) Measuring the deployment of network censorship filters at global scale. In *Network and Distributed System Security (NDSS) Symposium 2020*. Reston, VA: Internet Society.
- Ramos D** (2020) Philippines government orders shutdown of country's leading broadcast network ABS-CBN. *Deadline*. Available at <https://deadline.com/2020/05/philippines-abs-cbn-shutdown-rodrigo-duterte-1202927654/> (accessed 20 October 2022).
- Raveendran N and Leberknight CS** (2018) Internet censorship and economic impacts: A case study of internet outages in India. In *Proceedings of the Twenty-Fourth Americas Conference on Information Systems*. Atlanta, GA: Association for Information Systems.
- Reno v. ALU** (1997) American Civil Liberties Union, 521, U.S. 811.
- Reporters Without Borders** (2020) World press freedom index. Available at <https://rsf.org/en/ranking> (accessed 4 May 2020).
- Roberts H, Zuckerman E and Palfrey J** (2011) Circumvention tool evaluation. In *The Berkman Center for Internet & Society Research Publication Series*. Cambridge, MA: The Berkman Klein Center for Internet and Society at Harvard University.
- Rydzak JA** (2018) A total eclipse of the net: The dynamics of network shutdowns and collective action responses. *University of Arizona*.
- Rydzak J** (2019) Of blackouts and bandhs: The strategy and structure of disconnected protest in India. Available at <https://ssrn.com/abstract=3330413> (accessed 4 May 2020).
- Sagir M and Varlioglu S** (2020) Explaining the relationship between internet and democracy in partly free countries using machine learning models. In *IT Research Symposium 2020*. Cincinnati, OH: University of Cincinnati, School of Information Technology.
- Sayadi E and Taye B** (2020) #KeepItOn: As Yemen's war goes online, internet shutdowns and censorship are hurting Yemenis. Available at <https://www.accessnow.org/keepiton-as-yemens-war-goes-online-internet-shutdowns-and-censorship-are-hurting-yemenis/> (accessed 20 October 2022).
- Selva M** (2019) Reaching for the off switch: Internet shutdowns are growing as nations seek to control public access to information. *Index on Censorship* 48(3), 19–22.
- Sfakianakis A, Athanasopoulos E and Ioannidis S** (2011) CensMon: A web censorship monitor. In *Proceedings of USENIX FOCI 2011*. Berkeley, CA: USENIX Association.
- Shahbaz A and Funk A** (2020) Freedom on the net 2019: The crisis of social media. *Freedom House*.
- Shen X and Truex R** (2021) In search of self-censorship. *British Journal of Political Science* 51(4), 1672–1684. <https://doi.org/10.1017/S0007123419000735>.
- Sisario B** (2010) U.S. Shuts Down Web Sites in Piracy Crackdown. *New York Times*.
- SK C** (2020) Those unspoken thoughts: A study of censorship and median freedom in Manipur, India. *Open Observatory of Network Interference*.
- Smith A** (2015) Democracy is not a mystifying western plot—It is a universal value. *The Guardian*. Available at <https://www.theguardian.com/global-development/2015/jul/03/democracy-not-mystifying-western-plot-universal-value-sustainable-development-goals> (accessed 4 May 2020).
- Smith J and Gladstein A** (2018) How the UN's sustainable development goals undermine democracy. *Quartz Africa*. Available at <https://qz.com/africa/1299149/how-the-uns-sustainable-development-goals-undermine-democracy/> (accessed 4 May 2020).
- Stepanova E** (2011) The role of information communication technologies the “Arab Spring.” *PONARS Eurasia*.

- Subramanian R** (2012) The growth of global internet censorship and circumvention: A survey. *Communication of the International Information Management Association*, 11(2).
- Sutterlin E** (2020) Flipping the kill-switch: Why governments shut down the internet. Undergraduate Honors Theses. Paper 1493.
- Taye B** (2020) Targeted, cut off, and left in the dark the #KeepItOn Report on internet shutdowns in 2019. *Access Now*. Available at <https://web.archive.org/web/20200611142256/https://www.accessnow.org/cms/assets/uploads/2020/02/KeepItOn-2019-report-1.pdf> (accessed 4 May 2020).
- Tilley A and Jenkins E** (2020) Aid transparency index 2020. *Publish What You Fund*.
- Turak N** (2020) UAE loosens some VoIP restrictions as residents in lockdown call for end to WhatsApp and Skype ban. *CNBC*. Available at <https://www.cnn.com/2020/03/26/coronavirus-lockdown-uae-residents-call-for-end-to-whatsapp-skype-ban.html> (accessed 4 May 2020).
- UN General Assembly** (2015) Transforming our world: The 2030 Agenda for Sustainable Development, *A/RES/70/1*. Available at <https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf> (accessed 4 May 2020).
- UN General Assembly** (2021) Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development. *A/RES/71/313*, *E/CN.3/2018/2*, *E/CN.3/2019/2*, *E/CN.3/2020/2*, & *E/CN.3/2021/2*. Available at https://unstats.un.org/sdgs/indicators/Global%20Indicator%20Framework%20after%202021%20refinement_Eng.pdf (accessed 20 October 2022).
- USAID** (2019) FY 2020 USAID journey to self-reliance country roadmap methodology guide. Available at <https://selfreliance.usaid.gov/> (accessed 4 May 2020).
- USAID** (2020) Digital Strategy 2020–2024. *United States Agency for International Development*. Available at <https://www.usaid.gov/usaid-digital-strategy> (accessed 20 October 2022).
- VanderSloot B, McDonald A, Scott W, Halderman JA and Ensafi R** (2018) Quack: Scalable remote measurement of application-layer censorship. In *Proceedings of the 27th USENIX Security Symposium*. Berkeley, CA: USENIX Association.
- Wagner B, Gollatz K, Calderaro A** (2013) Common narrative—Divergent agendas: The internet and human rights in foreign policy. In *Proceedings of the 1st International Conference on Internet Science*. Brussels, Belgium: NextGeneration Internet.
- Weaver N, Sommer R and Paxson V** (2009) Detecting Forged TCP Reset Packets. *International Computer Science Institute*. Available at <https://www1.icsi.berkeley.edu/pubs/networking/ndss09-resets.pdf> (accessed 20 January 2022).
- Weinberg Z** (2018) *Toward Automated Worldwide Monitoring of Network-Level Censorship*. Carnegie Mellon University. Available at <https://research.owlfolio.org/pubs/thesis.pdf> (accessed 20 January 2022).
- West DM** (2016) Internet shutdowns cost countries \$2.4 billion last year. *Center for Technology Innovation at Brookings*.
- Woodhams S and Migliano S** (2020) The global cost of internet shutdowns in 2019. *Top10VPN.com*.
- Yadav TK and Chakravarty S** (2018) Trends and patterns of internet censorship in India. *Indraprastha Institute of Information Technology Delhi*.
- Zelege TA** (2019) The quandary of cyber governance in Ethiopia. *Journal of Public Policy and Administration* 3(1), 1–7.
- Zhong Z-J, Wang T and Huang M** (2017) Does the great fire wall cause self-censorship? The effects of perceived internet regulation and the justification of regulation. *Internet Research* 27(4), 974–990. <https://doi.org/10.1108/IntR-07-2016-0204>.
- Zittrain J, Faris R, Noman H, Clark J, Tilton C and Morrison-Westphal R** (2017) The Shifting Landscape of Global Internet Censorship. *Berkman Klein Center for Internet & Society Research Publication*.
- Zuckerman E** (2010) Intermediary censorship. Available at https://tavaana.org/sites/default/files/intermediary_censorship_.pdf (accessed 4 May 2020).

Cite this article: Fletcher T and Hayes-Birchler A (2023). Is remote measurement a better assessment of internet censorship than expert analysis? Analyzing tradeoffs for international donors and advocacy organizations of current data and methodologies. *Data & Policy*, 5: e9. doi:10.1017/dap.2023.5