DATA AND MEDIA: GOVERNANCE BY RESTRICTIONS

By Adil Nussipov

CENTER FOR MEDIA, DATA AND SOCIETY
CEU SCHOOL OF PUBLIC POLICY

Published by CEU Center for Media, Data and Society
January 2021

Table of Contents

Introduction

Our social activity is becoming increasingly quantified, that is, captured, recorded, stored and organized in the forms of terabytes of data. Data flows across borders, allowing scientific discovery, technological advancement and business innovations – and in general, transforming the way the present and the future look like. During the pandemic that started in March 2020, data flows have allowed governments to trace and identify Covid-19-infected persons and medical scientists to analyze and test potential vaccine models, enabled students to continue their education remotely, supported small businesses to adapt to new economic realities and connected millions of quarantined people to each other.

At the same time, however, these data flows pose risks that we do not fully understand yet. In essence, while providing benefits, data flows also project significant influence and control over our lives. They allow our personal information to be captured, stored, analyzed. They filter what we see and what we do not see in news. They target us to affect our behavior. Scenarios where it could go wrong are not hard to imagine. As with anything new, however, before rushing to regulate data, we need to have a better understanding of the effects of data on our lives.

Media and data are intertwined. Thanks to technological convergence, we can hardly differentiate today where media turns into data and where data turns into media. It is not surprising then that regulating data also means to some extent regulating media. Media organizations, whether television channels, newspapers, news portals and aggregators, blogs, radio stations and podcast providers, heavily rely on data flows. Yet, we do not know what kind of impact data regulation is going to have on media freedom.

This study aims to provide a new context for media organizations on how to think about cross-border data flows and better understand what impact these flows can have on their freedom. It aims to understand how and why restrictions on cross-border data flows affect media freedom. The study is not aimed at giving a definite solution on how data should be regulated, but rather to offer a framework to think about data regulation and how that affects media freedom.

No systematic study quantitatively investigating the relationship between data and media regulations has been conducted thus far. This study provides a quantitative measure of data and media restrictions in 64 countries. Secondly, it offers a novel framework for analyzing a variety of data governance models across countries. Lastly, it offers a novel theoretical framework for understanding how these data governance models affect media freedom.

The key findings of the study are: first, that restrictions in cross-border data flows are positively associated with restrictions on media freedom (especially true for countries with a state model of data governance); and secondly that media organizations are not prepared to handle the impact of data regulation on media freedom as they do not engage in the data policymaking, have low data literacy and lack a toolkit to fight against misuse of various forms of data regulation.

The first section of the paper provides a conceptual background for data flows, introducing three main data governance models and illustrating how they may affect media freedom. The second section presents the results of a statistical investigation of the relationship between data and media restrictions. The third section substantiates these results with case studies. The last section puts forward a series of policy recommendations and presents a few concluding remaking.

I. Governing Cross-Border Data Flows

Before defining data governance, it is important to understand what data flows are. At the most basic level, data flow is the process of transferring data from point A to point B. As most data transfers take place over the internet networks, by default data flows have a transnational character as flows constantly cross borders (Kuner 2013).

To predict exactly where data has crossed borders is impossible. Because it is technologically complex, tracking how data crossed borders is costly, thus, it makes sense from a regulatory perspective to consider all data flows as cross-border transfers (Kuner 2013). The same issue makes it problematic to consider data flows as trade exchange. First, data flows can be both seen as trade in services and goods. Secondly, suppliers and consumers of data services do not need to be in the same physical location to trade (Aaronson 2018). The transaction takes place in nanoseconds and it is impossible to track whether a particular data transaction is an export or import, or who supplies or controls the data at any point in time because, on its pathway, transacted data may cross several borders before reaching the final user and/or data collector (Kuner 2013). Thus, data governance can be defined as a domestic regime of rules employed by governments to regulate cross-border data flows.

It makes sense to expect data flows by default to be governed at the global level, since their inherent cross-border nature calls for it. However, except for a few regional initiatives (such as OECD, Council of Europe, APEC and the EU), there is no global regulatory framework for cross-border data flows. Global data governance is deeply fragmented on political grounds. Governments around the world are divided in their domestic approaches toward governing data. Already in 2018, United Nations Conference on Trade and Development (UNCTAD) found that 107 states adopted laws that fully or partially addressed data regulation.

Data governance is a multi-sectoral, multi-stakeholder and dynamic political process. It is dynamic because, as technology is constantly changing and evolving, data regulations always lag behind technological advancements and innovation. In turn, it requires rules on data flows to constantly adjust to the technological change. Most of international rules on data that exist now were drafted in 1980s and 1990s when they were based on much simpler technological processes than those that we have now. It is a multi-sectoral process because data governance covers many issues including data protection, data privacy, access to information, cybersecurity and national sovereignty, digital trade, economic dependencies, and digital rights. Its effect is present across a range of sectors and, as such, multiple stakeholders are involved: it is not only data protection authorities or digital trade experts, but also representatives of human rights organizations, national security experts, technology experts, and private companies among many others that are involved in this process. Such a complex nature is partly the reason why we have not seen any success in creating global rules on data: doing so would require a political balance between sometimes too divergent policy interests.

Without such global rules, states around the world have developed a variety of domestic data regimes.

An Introduction to Data Governance Models

In this paper, governance model is understood as the framework of rules on governing a particular issue or area with a set of policy tools and political goals that are common across countries. In essence, it means that the governance model consists of two parts (both generalized as being common across several countries): political goals and policy tools to achieve them. The data governance model is a framework of rules aimed at governing cross-border data flows, with goals and tools that are common across countries. In this section, I will describe the goals and tools that are part of data governance models.

According to Leblond and Aaronson (2019), a core function of data governance is to solve the Data Trilemma of the following political goals: allowing free cross-border flows of data (1), ensuring that data are protected at the national level (2); and providing citizens with a trusted data environment (3). Among these three, governments can combine only two at the same time. Free flows of data are needed for economy as data flows are the essence of trade in digital goods and services. On the other hand, without national regulation, free flows of data pose a number of political risks for governments ranging from foreign interventions into elections, foreign propaganda and misinformation campaigns to using data dependency as economic leverage in international negotiations. At the same time, regulating data too much risks restricting data flows, which negatively affects economic competitiveness and innovation (Francesca, Ferracane and van der Marel 2019). Finally, the free flow of data and protection of data will not be valuable anymore if citizens no longer trust the digital environment and stop generating data via their online activities. If data is a resource, or a capital, or a labor, citizens are the ones who produce it. For this, governments need to provide citizens with an environment where they are protected from commercial and political exploitation of personal data and structural discriminations of AI algorithms, and where their rights to data privacy are secured.

Three political goals are at the core of the three existing governance models: the state model, focused on national data protection; the market model, focused on the free flows of data; and the citizen model, focused on creating trusted data environments for citizens. In the state model, governments prioritize national data protection over free data flows and centralize control over data flows. Centralization usually takes place via data localization requirements, when cross-border transfer of data is allowed only under certain conditions. Although this model may provide citizens with an overly protected data environment, it also significantly increases costs for international business operations. For example, China bans any cross-border data transfers unless specifically allowed after a security assessment carried out by the government. That makes the entrance of foreign companies into the Chinese market extremely costly and the expansion of domestic companies internationally very difficult.

In contrast, in the market model, instead of directly regulating data flows, the government creates a basic regulatory environment, in which data flows are self-regulated by market players and industries. Essentially, in this model, the government relies on selective sector-specific data regulation, corporate and sectoral data standards and industry self-regulation instead of creating a comprehensive domestic regulatory framework. By doing so, the government ensures free flows of data and economic benefits associated with it. However, it also has to decide between instituting a national data protection strategy or creating a solid legal environment for the protection of citizen rights. The U.S. is an example of a country with the market model of data governance. It ensures free cross-border flows of data and does not have any national data strategy. Rather, the government protects the rights of citizens via sector-specific laws on data flows such as regulations of financial data and health data. The U.S. also includes data protection-related provisions in its international free trade agreements in order to ensure the international protection of national security, consumer privacy and rights.

Lastly, in the citizen model of data governance, the government puts a strong emphasis on data protection and privacy, in two ways: the government can shield citizens from both the commercial and political misuse of their personal data by companies and foreign (or their own) governments; or the government can empower citizens to keep companies and organizations accountable for how they handle their data. In this model, although the government can combine maintaining a high-trust data environment with national data strategy or with free data flows, trust always will be a founding principle, with one of the other two factors playing only a secondary role. As such, the government may create a high-trust data environment as well as a national data protection strategy, but the latter will be done to protect the former and not to establish a state-wide surveillance. Similarly, trust can be combined with free data flows, but only under binding international rules that will protect the citizens' interests. The case of members of the European Union (EU) is illustrative. On their own, EU member-states do not have their own national data strategies but rather comply with the EU's General Data Protection Regulation (GDPR), which allows them to benefit from free cross-border flows of personal data among themselves. At the same time, for data flows outside the EU, the union requires trading partners to comply with a set of requirements, which in practice means that a trading country must be compliant with the GDPR's rules. By focusing on citizens' trust, member states give up national data strategies for the EU-wide regulation and free data flows.

Data Governance and Media Freedom

How are data governance and media connected? Digitization of information, consisting of packaging of this information into data and distribution of these packages via digital networks, sits at the core of technological advancement that the world has faced in the last 20 years (Iosifidis 2011). As a result, media experienced a convergence of different types of media forms. In essence, convergence refers to "the delivery" of different media types "via the same transmission platform." What this means for media is that when previously video, audio and text were transmitted via different channels, thanks to digital technology, they are all now accessible on the same channel at the same time – the internet (Iosifidis 2011, 172). Today, media rely on constant cross-border flows of data, more so than any other industries. Media organizations, whether they operate online or offline, transfer a vast amount of data across borders, including personal and non-personal data.

Data restrictions measures such as data localization, intermediary liability and content access contribute to a restrictive media environment.

First, data localization is commonly understood as a set of requirements that "either mandate data to be kept locally or impose conditions to transfer data cross-border" (Ferracane, Lee-Makiyama and van der Marel 2019). There are five different data localization regimes: unconditional flow regime, conditional flow regime, local storage requirement, local processing requirement and ban on data transfers (Chander and Ferracane 2019). In the unconditional flow regime, cross-border data flows are allowed without any restrictions. In the conditional regime, data flows are allowed only if certain conditions are met by the data sharing company and/or data receiving country. In local storage regime, companies are required to locally store the copies of data that are being transferred abroad, while in the local processing regime, companies should either build data centers or employ third parties to process all data within a country. In the last regime, data transfers abroad are banned.

Second, intermediary liability refers to the legal responsibility of intermediary platforms for "the illegal and harmful activities performed by their users even when the platforms are unaware of these activities" (Ferracane, Lee-Makiyama and van der

Marel 2019, 102). The concept is based on an important distinction between “content producer” that is a party that creates and posts content, and “intermediary” that is a company responsible for mediating between producer and the internet, such as “Internet Service Providers (ISPs), web hosting providers, social media platforms and search engines” (Ferracane, Lee-Makiyama and van der Marel 2019, 102). This concept refers only to those countries that make intermediaries legally responsible for illegal and harmful activities carried out by their users. In contrast, the safe harbor model is the opposite concept where governments, under certain requirements, grant intermediaries “broad or conditional immunity” for the activities of their users (Ferracane, Lee-Makiyama and van der Marel 2019, 102).

Lastly, content access refers to policy measures that regulate access to online content (Ferracane, Lee-Makiyama and van der Marel 2019). It refers to online censorship, content blocking and filtering, discriminatory application of technical standards and practices of controlling network bandwidth such as prioritizing certain content or certain providers, or slowing down foreign websites (Ferracane, Lee-Makiyama and van der Marel 2019).

In regard to media, in combination with harsh intermediary liability and content access, rules on data localization may help governments to easily locate, control and shut down media websites and platforms that are politically unaligned. Having all data stored in local data centers and having access to them makes targeted surveillance, internet shutdowns and slowdowns, content blocking and censorship easier to implement by governments. As a reciprocal effect, media organizations and workers operating in such environments are forced to self-censor. While data localization laws ensure that governments control data flows and know where to press to switch them off, intermediary liability and content access rules provide governments with legal justifications for doing so under the umbrella of national security, anti-terrorism protection, and data protection, among others.

Strict data localization requirements also add legal and financial costs to the operations of media organizations, distorting competition in the media market and hurting media pluralism. It is a significant financial burden to build and sustain data centers to locally process data, especially for non-commercial media organizations. In such situations, facing the need to process data locally, media organizations either turn to third-party solutions and spend significant amount of resources on complying with data rules, or are simply forced out of the market. As such, data localization requirements can create unequal market conditions where only large media companies or state-managed media organizations can survive, driving out less wealthy and less government-friendly media. Finally, data localization laws may also create opportunities for governments to engage in regulatory licensing practices where the government-controlled regulators use technical licensing to drive opposition media out of the market or to disrupt their operations.

Thus, analyzing the logic of data governance models and comparing it with media restrictions, it appears that countries with the state model of data governance also have a more restricted media environment than countries with the market or citizen models. In the state model, governments impose discriminatory access rules and strict legal responsibilities on online platforms and providers, and require security assessments before allowing data transfers. In the market model, these tools are either not used or are fine-tuned to ensure free market operations under minimal regulation. Lastly, in the citizen model, although the forms of these tools and the extent to which they are used vary, they do not reach extreme levels.

Variation in how data localization measures are used across models highlights how different data governance models affect media freedom differently. Governments that regulate data via the state model tend to employ the harshest forms of data localization, whether by banning cross-border data transfers altogether, especially for industries dealing with sensitive data (like news organizations) or requiring organizations to store the majority of data (or at least a copy of each data bit transferred abroad) in data centers located inside the country. States that govern data flows via the citizen model tend to set up conditional regimes on data localization, meaning that data can be transferred abroad only if certain requirements are fulfilled, especially in relations to personal data. For example, the EU requires states willing to share or access data with EU member-states to be compliant with GDPR requirements. Finally, countries with the market model tend to employ the minimum amount of data localization, if any at all.

The purposes of localizing data also vary across data governance models. While the citizen model localizes data in order to ensure the protection of human rights of citizens, the state model does so in order to protect the government's interests, national security and sovereignty. In countries with the state model, data localization is also an additional policy tool to censor, surveil and control groups with opposing views. Lastly, in the market model, data localization presents an unnecessary economic cost to the free flow of capital, goods, services, and data itself.

In summary, restrictions on cross-border data flows have negative implications for media environment. As such, I aim to test the two following hypotheses:

H1: Higher scores on data restrictions measures are positively associated with media restrictions scores;

H2: Countries with the state model of data governance tend to have a more restricted media environment than countries with the market and citizen models.

II. Measuring Data and Media Restrictions

To allow cross-country analysis of data regulations, I use the Digital Trade Restrictiveness Index dataset by Ferracane, Lee-Makiyama and van der Marel (2019) that examines the restrictiveness of data-regulating laws across 64 countries. At the moment, this is the most comprehensive dataset that provides in-depth measures of regulations related to cross-border data flows.

The variable of our interest, Data Restrictiveness Index is a cluster of three groups of policy estimates: data localization, intermediary liability and content access (Ferracane, Lee-Makiyama and van der Marel 2019). The Index for Data localization measures the degrees of requirements for storing and processing data locally, data retention, consent of data subjects for data collection, government's access to personal data and penalties for noncompliance. The Index for intermediary liability measures the presence of safe harbor clauses, user identity and monitoring requirements, and notice and takedown regimes. Lastly, the content access index measures the degrees of web content blocking and filtering, discriminatory use of licensing schemes, prioritization of certain network bandwidth and slowing down of foreign websites.

These measures are combined into the final Data Restrictiveness Index that ranges from 0 to 1, 0 indicating unconditional free flow of data and 1 indicating extremely restricted regime on data flows. The five above mentioned measures do not weight equally in the final index: data localization, being the cornerstone of data governance, comprises 40% of the index in accordance with its importance whereas intermediary liability and content access are allocated 30% each.

The effects of data localization on media freedom are less likely to be observed in formal laws and regulations, and more in the experiences and perceptions of media organizations. Regulations of cross-border data flows rarely emphasize the type of organizations that they aim to regulate or make any direct references to the media industry. As such, it is difficult to establish any direct law-based links between data localization and media freedom. On the other hand, indices that measure the perceptions and experiences of media workers are useful in that sense. These indices help us understand the real-life impact that data localization regulations may have on the freedom of media organizations and professionals.

In that sense, Reporters Without Borders' World Press Freedom Index fits the purpose well. The index is compiled based on surveys of media experts across seven indicators: pluralism, media independence, environment and self-censorship, legislative framework, transparency, infrastructure and abuses (Reporters Without Borders 2020). Each indicator measures the respective aspects of media freedom, such as the degree of representation of different opinions, independence of media from government and business, political and legislative environment, transparency of institutions that regulate media, quality of infrastructure and number of violent acts against journalists in a given period (Reporters Without Borders 2020). The resulted index is a combination of seven indicators that ranges from 0 to 100, where 0 indicates more free media environment, and 100 indicates less free media environment. For the purposes of cross-variable comparison, the final Media Restrictiveness Index was rescaled to 0 to 1 (as a continuous variable), with 0 indicating less media restrictions, and 1 indicating more media restrictions. The resulted dataset contains observations for 64 countries (see full table in Appendix I). As Digital Trade Restrictiveness Index started only in 2019 and does not contain data for previous years, the resulted dataset is the first one that simultaneously addresses both data governance and media freedom.

As shown on Figure 1, results of statistical regression illustrate support for the first hypothesis. Based on the given data, we observe positive and statistically significant association between restrictions on cross-border data flows and restrictive media environment among the 64 countries. Results of the base model indicate that a one-point increase in data restrictions is associated with a 0.68 increase in media restrictions. Essentially, that means that if a country implements any form of data localization measures, this increase in data restrictions may be associated with an increase in media restrictions by 0.68 points. Results of regression are summarized in Table 1 (see Appendix II).

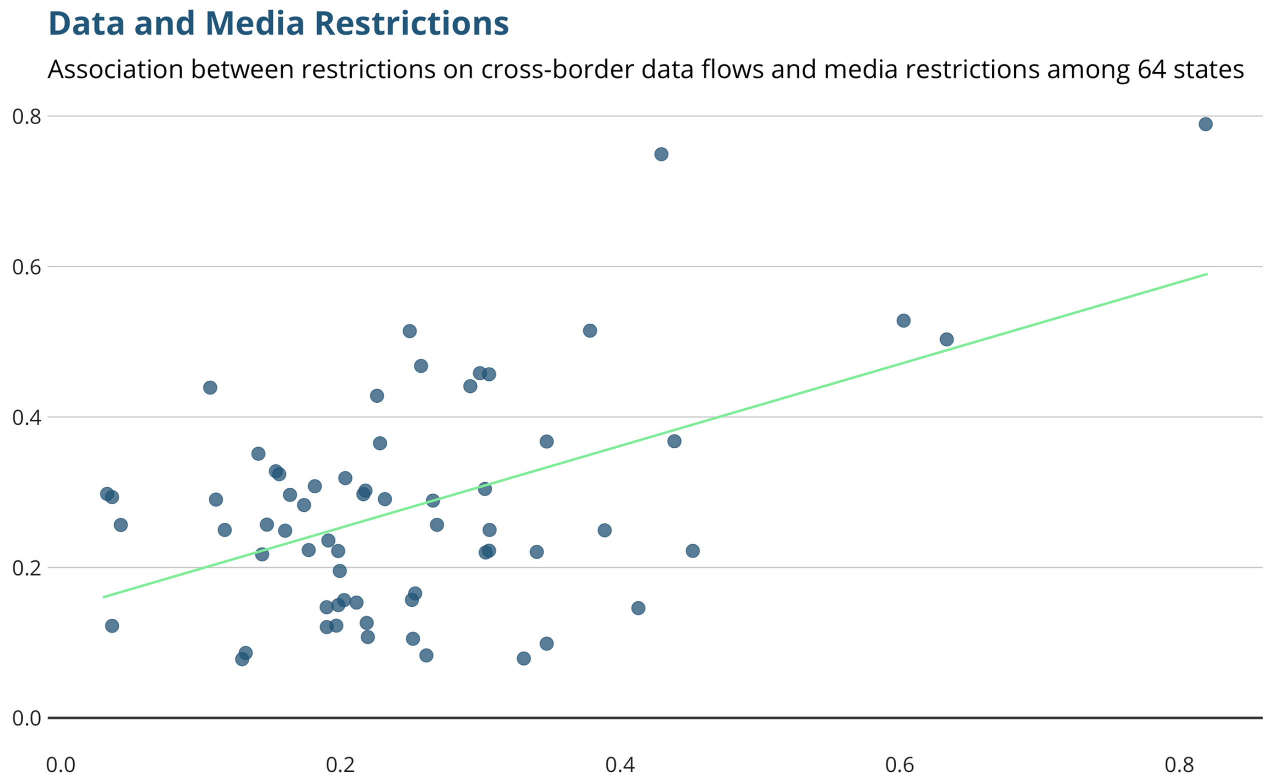


Figure 1

Results keep their statistical significance even when we change the statistical models. In the one where we substituted Data Restrictiveness Index with its logarithmic form to attribute for a gap between countries with the state models and the rest, the association between data and media restrictions remain positive. When testing for individual data governance models, the effect of the Data Restrictiveness Index remains positive.

Figure 2 illustrates how data governance models score on media restrictions. In the figure, Data Restrictiveness Index is assigned to x-axis while Media Restrictiveness Index is assigned to y-axis. As it can be seen, the general trend is that the state model of data governance scores high both on data restrictions and media restrictions in support of the second hypothesis. China is located on the far-right upper corner followed by Vietnam (light green dot), Russia and Turkey (blue dots).

Data Governance Models and Media Restrictions

How data governance models score on media restrictions among 64 states

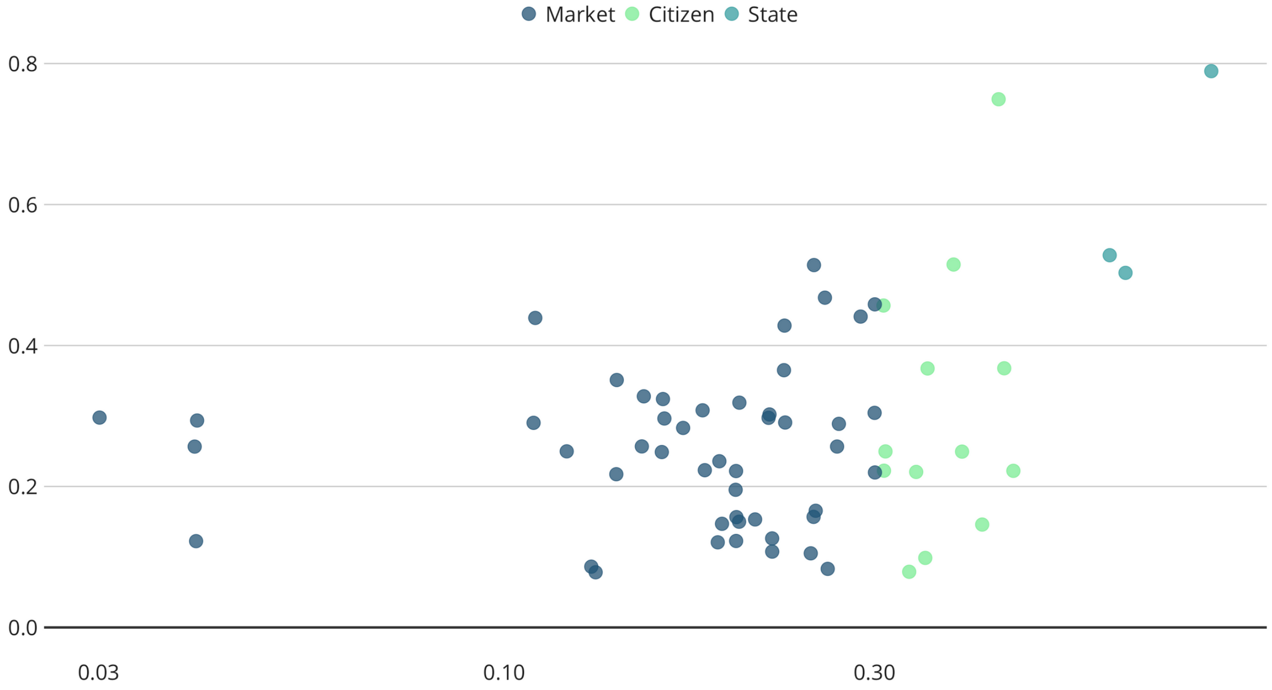
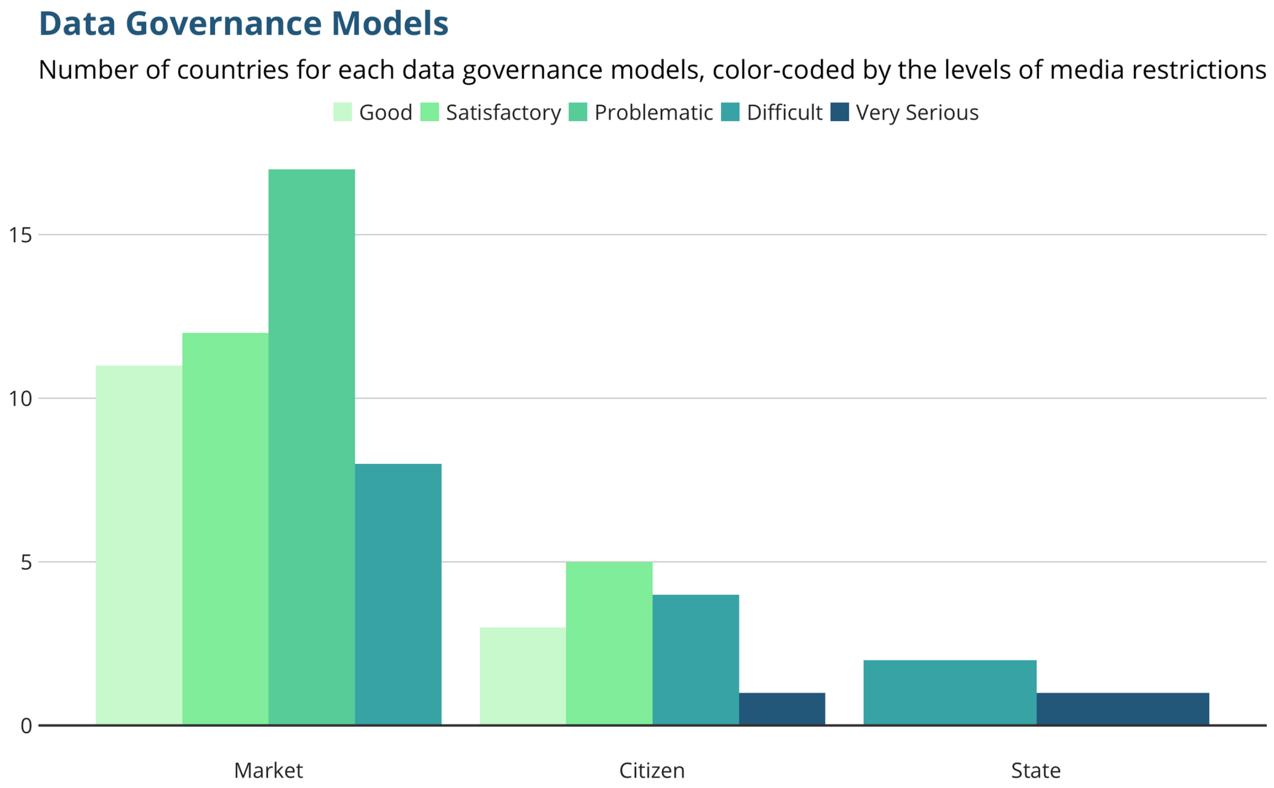


Figure 2

To better see the association with the state model and media restrictions, Figure 3 breaks down the data governance models by number of countries and color-codes the levels of media restrictions, based on the six-cluster taxonomy used by Reporters Without Borders (2020). As it can be seen, countries with the state model score exclusively 35 points and higher, covering categories such as "problematic," "difficult" and "very serious." In contrast, countries with the market model do not score higher than 55. Lastly, countries with the citizen model show mixed results.



Data Governance Models

Number of countries for each data governance models, color-coded by the levels of media restrictions

Figure 3

III. Case Studies

This section provides several empirical cases on how data restrictions affected freedoms of media organizations. Experiences of journalists and media organizations across countries largely support propositions regarding how different data governance models affect media freedom differently. We will start with the state model. In the dataset, three countries score high both on data and media restrictions: China, Russia and Turkey.

State Model and Censorship

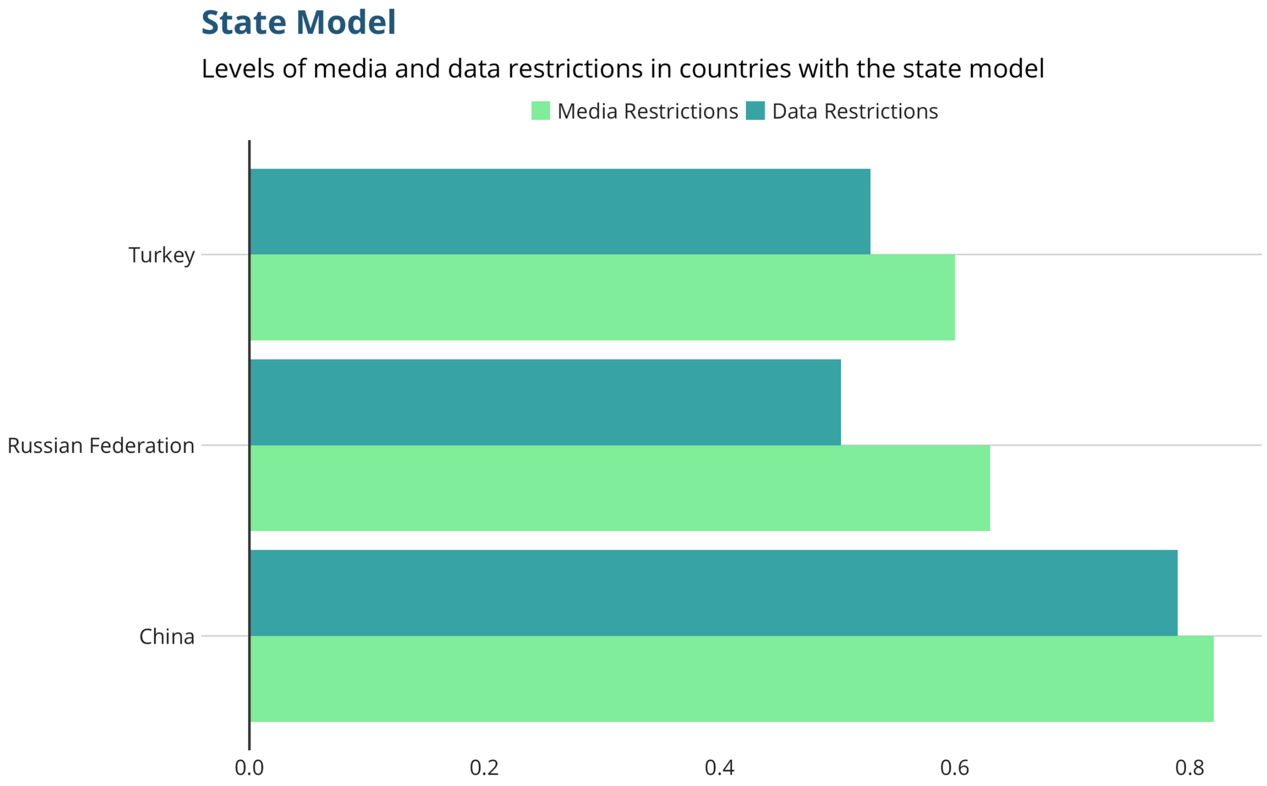


Figure 4

China's data governance framework is based on the Cybersecurity Law, which imposes strict data localization measures and real name registration, and also requires operators and providers to cooperate with government bodies in criminal and national security-related investigations (Shahbaz, Funk and Hackl 2020). Vaguely defined concepts of "important data" and "key information infrastructure operators" provide to a lot of room for the government to interpret these terms in a political manner. Intermediaries are not protected by safe harbor policy and platforms are required to remove the content immediately after receiving a notice from the government (Ferracane, Lee-Makiyama and van der Marel 2019). On top of the Cybersecurity Law, media sector regulations have own sector-specific data localization requirements. For example, since 2017, any website or social media network are legally required to get the approval of the Cyberspace Administration if they want to get involved in news production.

Russia's Federal Law No. 242-FZ requires all processing and storage of personal data of Russian citizens to be done in data centers located inside the country. Data localization are also found in sector-specific regulations, such as in the media and financial sectors. Particularly, Russia's Blogger Law (the Federal Law # 97) requires "organizers of information distribution in the internet" to store on Russian territory information on "facts of receiving, transfer, delivery and/or processing of voice information, texts, images, sounds and other electronic messages and information about users during six months from the end of these actions" (Ferracane, Lee-Makiyama and van der Marel 2019, 56).

Turkey follows the trend. The Turkish government has the right to access user data and ban content without a warrant as per the Law on Regulating the Internet (Ferracane, Lee-Makiyama and van der Marel 2019). Moreover, on 21 July 2020, Justice and Development Party (AKP) of Turkey proposed a law to Parliament that provides the government more control over social media platforms (CPJ 2020). The proposal includes data localization measures for platforms that have more than one million users in Turkey and obliges them to have local representative offices. It also includes measures on intermediary liability with a notice and takedown regime that will require platforms to remove content considered to violate personal rights and privacy within 48 hours upon receiving such an order from the government (CPJ 2020). Penalty for not complying with these data localization and notice and takedown requirements will be in the form fines (around US$ 4.4m) and blocking and/or slowing down of internet traffic by internet providers (CPJ 2020).

According to Gulnoza Said from CPJ Europe and Central Asia Program, the true intention of the proposed law is to censor journalists via international social media platforms (CPJ 2020). In other words, by regulating social media platforms such as Facebook and Twitter, the government is forcing the tech companies to shut down and censor journalists whose publications do not comply with the newly proposed law. Although the law is aimed only at platforms, it indirectly affects online journalism in the country, and that's what the government aims for. Indeed, according to CPJ, the Turkish government for a long time has been pushing social media platforms to block the publication of journalists who were critical of the government's actions (CPJ 2020).

Citizen Model and Right To Be Forgotten

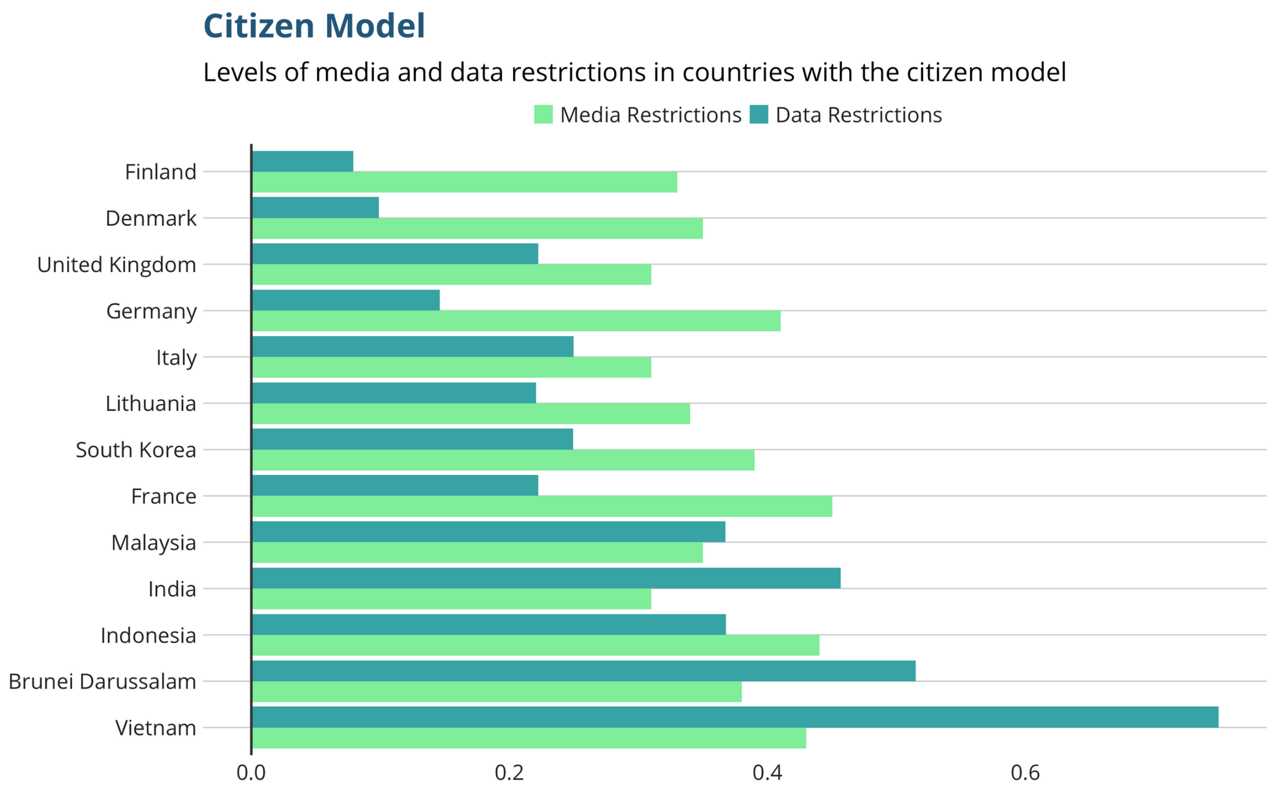


Figure 5

In the citizen model, rights of citizens and media freedom may come to a clash with each other. Although governments pursuing this model do not aim to control media, their efforts in ensuring that civil rights are protected may unexpectedly limit freedoms of media organizations. The challenge of balancing between the two is illustrated via GDPR's right to be forgotten, which allows citizens to request websites to remove their personal data (it is not surprising that the majority of EU member states in the dataset fall into the citizen model category). The right was established in Google Spain v AEPD and Mario Costeja González case where the European Court of Justice ruled that search engines are obliged to consider requests from citizens to remove content related to their personal data, and that, if engines reject such requests, citizens may take the case to national or European authorities. The rule raised strong criticism by media advocacy groups and news outlets such as The Guardian and CPJ, among them. From the citizen rights perspective, the right gives citizens control over personal data held and published on the internet. From media freedom perspective, however, this right can be used to erase historical facts that otherwise are considered legitimate news stories, a practice that essentially limits freedom of expression and access to information (King 2015). For example, reporting on crimes and horrors of the past can be distorted if parties involved in those crimes want to clean their names by using the right to be forgotten. The risk is that tech and social media platforms may start to erase requested data without proper investigation as a way to avoid fines, other penalties and legal problems for non-compliance with GDPR.

Vietnam presents another interesting case of the citizen model's negative effect on media freedom. Vietnam is a country with the citizen model of data governance, but it is also the second country in World Press Freedom Index dataset labeled as having "Very Serious" media environment – meaning a high level of restrictions on media. Vietnam has its own version of Cybersecurity Law (similar to China's), that makes online speech as well as platforms and intermediaries where online speech is observed liable for criminal investigations (Shahbaz, Funk and Hackl 2020). Intermediaries are required to cooperate with the government in removing content such as "state opposition undermining national security and social order; conducting propaganda; propagating obscenity, pornography and harming national traditions and customs; providing information offending organisations or individuals; and advertising banned goods and services, banned newspapers, works and publications" (Ferracane, Lee-Makiyama and van der Marel 2019, 60).

Market Model and Press Freedom

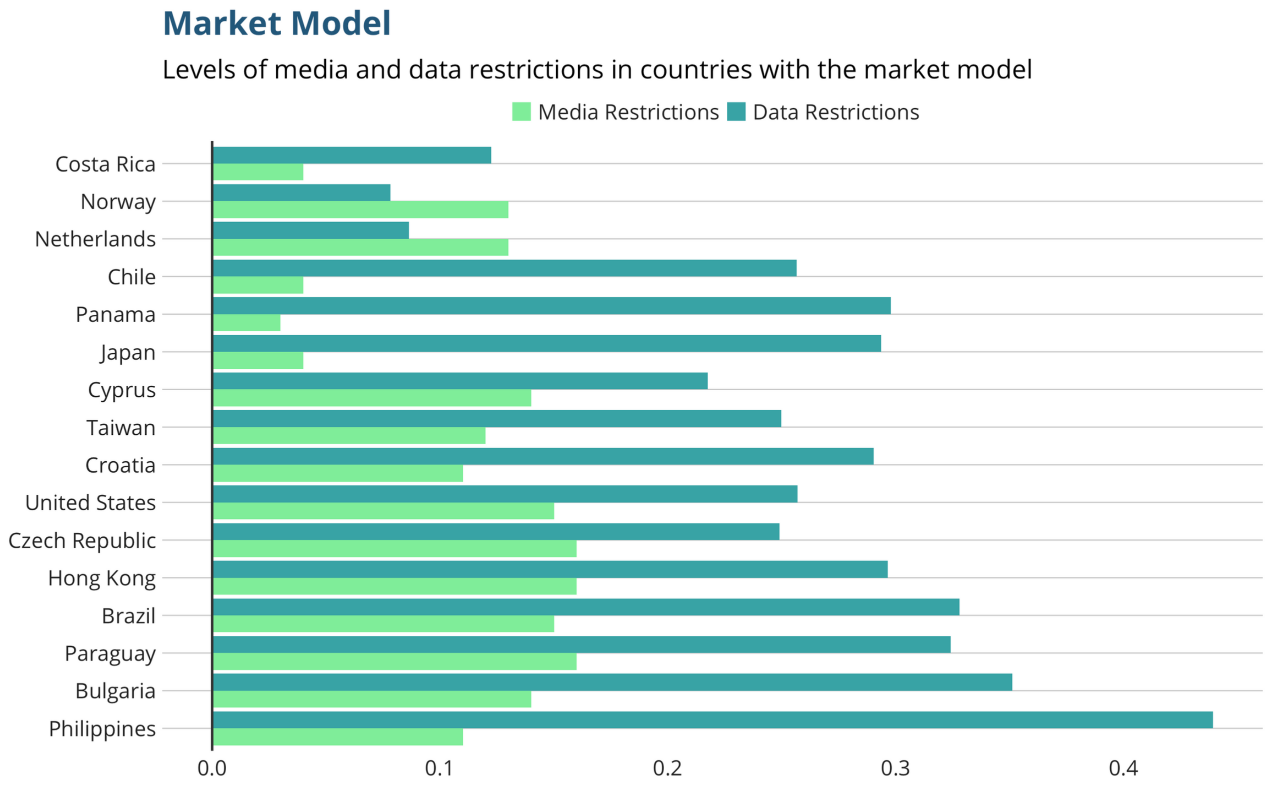


Figure 6

In the market model, the market-driven regulation of data flows is associated with a selective, industry-specific regulatory approach. The case of the new California Consumer Privacy Act (CCPA) and press freedom in the U.S. is a good example. CCPA introduces the first privacy-focused legislation in the U.S., although it focuses only on the state of California. It was the result of negative reactions of American citizens toward massive data breach scandals happening in the country, the most prominent being the case of Cambridge Analytica, which used the profiles of 87 million Facebook users to affect the outcomes of elections around the world, including the U.S.

CCPA provides special exemptions for journalistic activities. In general, media organizations are exempted from CCPA compliance if such compliance threatens these organizations' "non-commercial activities" (Zwillinger, Parsons and Anderson 2019). This means CCPA's data access and deletion requirements do not apply to news organizations' non-commercial journalistic activities, such as "cultivating sources, conducting interviews and investigations, taking notes, taking photographs or making audio or video recordings, collaborating with other members of the press, protecting the identity of reporters in undercover investigations, preparing materials for publication, and publishing news or opinions for public consumption" (Zwillinger, Parsons and Anderson 2019).

However, this exemption applies only to the extent that compliance with CCPA would disrupt these non-commercial journalistic activities. This contrasts with GDPR, according to which media organizations are exempt from GDRP requirements if they believe that a "publication" which GDPR attempts to regulate falls under the otherwise vague category of "public interest" (Pinto 2019). In regard to commercial activities, media organizations should be in full compliance. Despite that, the market regulation approach taken by the U.S. has its own risks. Political advertising and targeting, disinformation campaigns and fake news thrive in this self-regulated environment.

IV. Policy Recommendations

The study argues that on average, data restrictions and media restrictions tend to converge with one another. Whether directly or indirectly, restrictions on data flows are associated with restrictions on media freedom. Countries with the state model tend to restrict data and media more than countries with citizen and market models.

That does not necessarily mean that citizen or market models are better for media environment than the state model. Data governance models are theoretical constructs that aggregate, approximate and simplify a much more complex reality. Indeed, many countries do not clearly fall into one of three categories and there is significant variation among them. For example, both Taiwan and Singapore have introduced Data Protection Acts, but they fall within the market model category. Even between them, Singapore regulates disinformation via a law that some may consider to be censorship while Taiwan prefers the self-regulatory approach (CPJ 2019a; CPJ 2019b). Still, their data protection acts are not even close to GDPR.

Lastly, even within countries we see variations. For example, the U.S. lacks a national a data protection law, but the state of California has its own that can be compared to GDPR – yet, the U.S. still falls within the market model category.

Taking into accounts these caveats, data governance models still present a useful way of thinking about data governance and media freedom. At the core of any governance system is a set of political values that elicit themselves in the ways governments regulate. Data governance models are also built on these political values. This idea of having political values at the core of technological systems transforms data governance from something too technical and too complex to understand into something we are already familiar with. In other words, behind technicalities and complexities of data flows there are the same political, cultural and social values that are present in every human activity.

As such, data governance models help media organizations to better understand and engage with policymaking on data-related issues. Having said that, based on the research carried out for this study, I am putting forward a set of recommendations for media organizations, policymakers and advocates.

Media organizations, policymakers, advocates and other stakeholders should actively engage into policymaking processes on data flows.

Media freedom-related aspects of data governance are not part of the mainstream debate and media organizations do not push this agenda forward in the main national and international decision-making bodies. To protect its interests, the media industry should actively advocate for inclusion of media-specific aspects of data governance into in the global and national policymaking on data and request to participate in these processes.

Media organizations, policymakers, advocates and other stakeholders should increase their data literacy.

Media organizations and advocates seem to be lost in the technical complexities of data debate, which is not surprising given that even top policymakers could barely understand the answers of Mark Zuckeburg (Facebook's head) during the hearing in the U.S. Senate and at the European Parliament. The level of data literacy among public policymakers remains generally low. This needs to change as it directly affects the contribution of media organizations to the debates on data regulation. At the same time, journalists should also improve their data literacy either by continuing to practice data journalism (this could be an effective way to increase data literacy) or by gaining data-coding skills and knowledge about data regulations.

Media organizations, policymakers, advocates and other stakeholders should develop a toolkit to sustain the effects of data restrictions.

One very surprising and fundamentally important conclusion of this study is how under-protected and unprepared media sector is in the face of increasing data regulations. Journalists and media organizations around the world seem to be taken by surprise when what seemed to be a form of technical, non-media-related regulation starts to affect their work. This recommendation builds on the previous one about the need to develop data and technology literacy, by calling on media advocacy groups to design a set of tools to sustain the expansion of data regulation and protect the freedom of media from its effects. This may include, but not limited to: creating a global network of experts working on media-related aspects of data governance who can give expert opinion on new data laws and how they affect media and journalism; media organizations should provide more on-job training on legal aspects related to working with data; more media research organizations need to focus more on better understanding the effects of data regulation on journalism.

National and international policymakers working on data governance should include media organizations and other stakeholders into the data regulation debate at both national or international levels.

Technological convergence means media will play an even more important role in the lives of people in the coming years. Integration of technology into our everyday lives affects what people see, hear and perceive, and media has an important role in that. Excluding media organizations from policy debates about data regulations means increasing the risks of disinformation, fake news and propaganda. With the integration of technology into people's lives, people are bombarded with content from a sheer amount of news sources, which sometimes can prevent them from getting accurately informed. Disinformation has destabilizing consequences for peace and economy, especially at times of crises. As such, engaging media advocates and organizations and accommodating their concerns in the global data debate is essential.

REFERENCES

Aaronson, Susan Ariel. 2018. “Data is Different Why the World Needs a New Approach to Governing Cross-border Data Flows.” CIGI Papers No. 197, Center for International Governance Innovation.

Committee to Protect Journalists. 2019a. Singapore ‘fake news’ legislation endangers press freedom. Accessed on September 20, 2020 at https://cpj.org/2019/04/singapore-fake-news-legislation-endangers-press-fr/

Committee to Protect Journalists. 2019b. Q&A: Taiwan’s digital minister on combatting disinformation without censorship. Accessed on September 20, 2020 at https://cpj.org/2019/05/qa-taiwans-digital-minister-on-combatting-disinfor/

Committee to Protect Journalists. 2020. Turkey proposes social media law, threatening press freedom. Accessed on September 10, 2020 at https://cpj.org/2020/07/turkey-proposes-social-media-law-threatening-press-freedom/

Committee to Protect Journalists. 2015. Two continents, two courts, two approaches to privacy. Accessed on September 9, 2020 at https://cpj.org/2015/04/attacks-on-the-press-two-continents-two-courts-two-approaches-to-privacy/

Ferracane, Martina Francesca, Hosuk Lee-Makiyama and Erik van der Marel. “Digital Trade Restrictiveness Index.” European Center for International Political Economy.

Iosifidis, Petros. 2011. Global Media and Communication Policy: International Perspective. London: Palgrave Macmillan UK.

King, Geoffrey. 2015. Two Continents, two courts, two approaches to privacy. Committee to Protect Journalism, accessed on September 10, 2020 at https://cpj.org/2015/04/attacks-on-the-press-two-continents-two-courts-two-approaches-to-privacy/

Kuner, Christopher. 2013. “International Regulation of Transborder Data Flows.” In Transborder Data Flows and Data Privacy Law, Oxford: Oxford University Press, 25 – 59.

Leblond, Patrick and Susan Ariel Aaronson. 2019. “A Plurilateral “Single Data Area” Is the Solution to Canada’s Data Trilemma.” CIGI Papers No. 226, Center for International Governance Innovation.

Pinto, Timothy. 2019. The Rise of GDPR in Media Law. Taylor Wessing, accessed on September 18, 2020 at https://www.taylorwessing.com/download/article-rise-of-gdpr-in-media-law.html

Reporters Without Borders. 2020. 2020 World Press Freedom Index: Detailed Methodology. Accessed on July 1, 2020 at https://rsf.org/en/detailed-methodology

Shahbaz, Adrian, Allie Funk and Andrea Hackl. 2020. “User Privacy or Cyber Sovereignty? Assessing the human rights implications of data localization.” Freedom House, available at https://freedomhouse.org/report/special-report/2020/user-privacy-or-cyber-sovereignty

Zwillinger, Marc J., Kandi Parsons and Michelle Anderson. 2019. Legal Frontiers in Digital Media. Media Law Resource Center, accessed on September 18, 2020 at https://cnpa.com/mlrc-article-tackles-thorny-california-privacy-law-issues/

APPENDIX I

Table below provides data restrictions and media restrictions score for 64 countries in 2019, based on data from Ferracane, Lee-Makiyama and van der Marel (2019) and Reporters Without Borders (2020). Sorted from highest to lowest score on data restrictions.

	Country	Data Restrictions	Media Restrictions
1	China	0.82	0.7892
2	Russian Federation	0.63	0.5031
3	Turkey	0.6	0.5281
4	France	0.45	0.2221
5	Indonesia	0.44	0.3677
6	Vietnam	0.43	0.7493
7	Germany	0.41	0.146
8	South Korea	0.39	0.2494
9	Brunei Darussalam	0.38	0.5148
11	Malaysia	0.35	0.3674
10	Denmark	0.35	0.0987
12	Lithuania	0.34	0.2206
13	Finland	0.33	0.079
14	Italy	0.31	0.2498
16	India	0.31	0.4567
15	United Kingdom	0.31	0.2223
17	Pakistan	0.3	0.4583
19	Hungary	0.3	0.3044
18	Spain	0.3	0.2199
20	Thailand	0.29	0.441
21	Romania	0.27	0.2567
22	Poland	0.27	0.2889
24	Sweden	0.26	0.0831
23	Mexico	0.26	0.4678
28	Singapore	0.25	0.5141
26	Switzerland	0.25	0.1052
27	Canada	0.25	0.1569
25	Australia	0.25	0.1655
31	Nigeria	0.23	0.365

29	Greece	0.23	0.2908
30	Colombia	0.23	0.4282
35	Portugal	0.22	0.1263
33	Peru	0.22	0.3022
34	New Zealand	0.22	0.1075
32	Malta	0.22	0.2974
36	Austria	0.21	0.1533
37	South Africa	0.2	0.2494
42	Latvia	0.2	0.1953
39	Luxembourg	0.2	0.1566
41	Ireland	0.2	0.15
40	Estonia	0.2	0.1227
38	Ecuador	0.2	0.3188
43	Slovakia	0.19	0.2358
45	Iceland	0.19	0.1471
44	Belgium	0.19	0.1207
47	Slovenia	0.18	0.2231
46	Israel	0.18	0.308
48	Argentina	0.17	0.283
51	Paraguay	0.16	0.324
49	Hong Kong	0.16	0.2965
50	Czech Republic	0.16	0.2489
52	United States	0.15	0.2569
53	Brazil	0.15	0.3279
55	Cyprus	0.14	0.2174
54	Bulgaria	0.14	0.3511
56	Norway	0.13	0.0782
57	Netherlands	0.13	0.0863
58	Taiwan	0.12	0.2498
60	Philippines	0.11	0.4391
59	Croatia	0.11	0.2903
61	Japan	0.04	0.2936
63	Costa Rica	0.04	0.1224
62	Chile	0.04	0.2565
64	Panama	0.03	0.2978

APPENDIX II

Table below shows results of statistical regression. First column represents results for simple OLS regression with media restrictions as dependent variable, and data restrictions as independent variable. It shows positive relationships between data and media restrictions. Second column represents the same regression but with logarithmic form of data restrictions variable. Last column represents original OLS model with the inclusion of dummy variable for each data governance model.

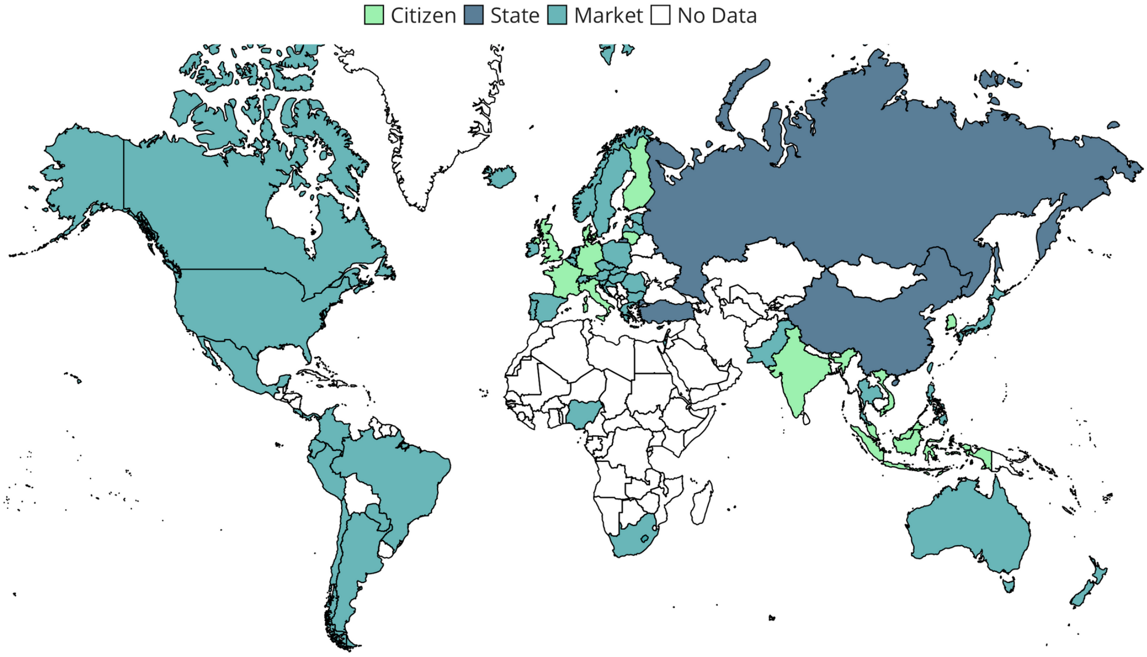
	OLS	LOG	OLS +
Intercept	0.14	0.4	0.15
	(0.03)***	(0.05)***	-0.09
Data Restrictions	0.54		0.4
	(0.11)***	–	-0.24
Log (Data)	–	0.08	–
		(0.03)**	
State	–	–	0.17
			-0.11
Market	–	–	0.02
			-0.06
Citizen	–	–	–
R-squared	0.25	0.1	0.29

*p < 0 ***; p < 0.001 **; p < 0.01 **

APPENDIX III

Data Governance across the Globe

Geographic distribution of data governance models



About the Author

Adil Nussipov is a Fellow at the Center for Media, Data and Society. He graduated with distinction from Central European University with an MA in International Relations. Before CEU, he obtained his BA degree in Political Science and International Relations at Nazarbayev University in Nur-Sultan, Kazakhstan. In the past, he acted as Senior Commissioning Editor at E-International Relations (online publication based in the United Kingdom) and Head of Research, Monitoring and Evaluation at Accountability Initiative for Reform (a non-profit organization in Kazakhstan). He is interested in global technology governance.

About CMDS

The **Center for Media, Data and Society** (CMDS) is a research center for the study of media, communication, and information policy and its impact on society and practice. Founded in 2004 as the Center for Media and Communication Studies, CMDS is part of Central European University's Democracy Institute and serves as a focal point for an international network of acclaimed scholars, research institutions and activists.

Cover photo: Shutterstock / everything possible

Center for Media, Data and Society

Tel: +36 1 327 3000 / 2609
Fax: +36 1 235 6168
E-mail: cmds@ceu.edu
cmds.ceu.edu
Postal address:
Center for Media, Data and Society
Central European University
Nador u. 9
1051 Budapest
Hungary