Subterranean Fault Lines in Cyberspace

An Analysis of Mid-level Internet Layers Across States

Meicen Sun*

Abstract

This empirical section explores differences in cyberspace border orientation among states. It analyzes three features of the middle layers of the Internet, which I term Connectivity, Concentricity, and Circularity. A cluster analysis reveals that the resulting grouping cannot be accounted for by either regime type or economic development. While further data collection and analyses remain underway, the preliminary results suggest that fault lines among states in the physical world do not neatly map onto the cyberspace, which calls for new frameworks and empirical analyses in this direction.

*Postdoctoral Scholar – Program on Democracy and the Internet, Stanford University; Research Affiliate – Initiative on the Digital Economy, Massachusetts Institute of Technology; Assistant Professor – School of Information Sciences, University of Illinois Urbana-Champaign (August 2023-). Email: mc27@stanford.edu.

1 Introduction

1.1 Motivation

This research project seeks to address two gaps in the study of cyberspace politics. First, the largest proportion of literature on the state control of the Internet deals with the top, application-layer of the Internet that directly interfaces with the end user (Roberts, 2018; King, Pan, & Roberts, 2017; Nabi, 2014; Munger, Bonneau, Nagler, & Tucker, 2019). This is arguably the most visible, palpable form of control, yet it also means that the more insidious forms of control exercised through the remaining, subterranean layers have not received nearly as much attention from political scientists.[1]

Second, this project is also motivated to shed new light on the rise of so-called digital authoritarianism (Guriev & Treisman, 2022; Dragu & Lupu, 2021). In particular, persistent tension within democracies on such issues as digital sovereignty and Internet governance hints at factors beyond regime type that may have implicated their ability to cooperate against the autocratic surge in cyberspace (DeNardis, 2014; Farrell & Newman, 2016; Pohle & Thiel, 2020; Floridi, 2020). In response to this, this project aims to contribute, beyond the substantive findings, a methodological approach that defers first and foremost to data in revealing underlying variations among states in their cyberspace behavior that may have thus far been overlooked. Doing so allows us to explore the possibility of emerging fault lines in cyberspace that are yet to be fully captured by theories derived from politics in the physical world.

1.2 Framework

This project takes stock of existing literature on the politics of sovereign borders (Simmons, 2005; Carter & Poast, 2017). Specifically, it borrows the "border orientation" framework as a theoretical guide for its empirical analyses (Simmons & Kenwick, 2021). Importantly, it reaffirms "filtering" as a central objective of border orientation in cyberspace as in the

[1]These are the, from the bottom up, physical, data link, network, transport, session, and presentation layers (OSI model), or the network access, Internet, and transport layers (TCP/IP model).

physical space, and that border orientation as a "compound concept" entails a composition of attributes. Additionally, the latent nature of border orientation as articulated in the framework informs this project's focus on subterranean layers of the Internet.

With this, this draft examines three features of middle-layer Internet infrastructure, each of which addresses a salient dimension of state border orientation in cyberspace:

Connectivity: This variable measures how connected a state's Internet infrastructure is to the rest of the world's. It captures the degree of *exclusion*, or to what extent a state decides to detach its Internet infrastructure from other states'.

Concentricity: This variable measures how centralized the state's Internet infrastructure is. It captures the degree of direct *control* the state assumes over its Internet networks, instead of delegation to either domestic private actors or foreign actors.

Circularity: This variable measures how robustly the Internet functions within state borders. In conjunction with the preceding two variables, Circularity complements the state's border orientation profile in highlighting, if any, *contrast* between internal and external information flows.

In the sections to follow, I outline data sources used to code these variables, before presenting the empirical results thus far obtained with a discussion to follow on substantive significance and next steps.

2 Data

Border orientation in cyberspace concerns a multitude of attributes. This draft zeroes in on the subterranean, below-application layers of the Internet infrastructure. While the final project will include all layers below this level, this draft examines those between the application- (top) and the physical- (base) layers. The three variables that have thus far been coded are Connectivity, Concentricity, and Circularity, illustrated below in turn.

2.1 Connectivity: Density of Internet Exchange Points

Unlike submarine cables on the physical level, Internet exchange points (IXPs) are less geographically constrained in their placement (Kurbalija, 2016; DeNardis, 2012). State actors thus exercise a considerable degree of discretion in their implementation without regard to, say, whether they are landlocked. This makes IXP density a meaningful indicator of a state's border orientation in cyberspace. Essentially, it measures how dense the web of peering points with the global Internet that the state maintains.[2] The thinner the web, the more it resembles an intranet than an Internet, and the lesser degree of the state's connectivity with the global Internet (Gregori, Improta, Lenzini, & Orsini, 2011; Chatzis, Smaragdakis, Feldmann, & Willinger, 2013).

To measure a state's Internet connectivity, I obtain primary data on the number of active IXPs for each state from the Packet Clearing House.[3] I then divide the number by the total population of the state and log it to calculate the state-level density of IXPs.[4]

2.2 Concentricity: Restriction on Internet Service Providers

A large and growing proportion of Internet control that occurs on the application layer is now being carried out by Internet service providers (ISPs) rather than government actors themselves (Seltzer, 2008; Sun & Zhao, 2022; Land, 2019). The delegation to ISPs obscures the act of Internet control by having ostensibly private actors perform tasks desired by

[2] "Oracle: China's Internet Is Designed More Like An Intranet".
[3] "Internet Exchange Point Datasets".
[4] "Population, Total", *The World Bank*.

the government actor (Tăbucă et al., 2010; Wachs, Schanzenbach, & Grothoff, 2014). An examination of Internet control should thus examine below-application layers at which Internet control takes place via the state control of ISPs.

To measure a state's Internet concentricity, which is a composite concept, I make use of the 2022 *Freedom on the Net* (FoN) scores compiled by the Freedom House. Specifically, I use the score to question A4 under the category of "Obstacle to Access" in the FoN. It asks, "Are there legal, regulatory, or economic obstacles that restrict the diversity of service providers?"[5] As the research methodology details, criteria include the presence of monopoly on ISPs, and various related legal, regulatory, or economic requirements. Both *de jure* and *de facto* hurdles to a free and competitive ISP market bear on the final score. To my knowledge, this is the best approximation for the degree to which the state restricts foreign ISPs from operating within state borders as well as the degree of control it exerts over private ISPs.[6]

2.3 Circularity: Broadband Internet Speed

Finally, the circularity of a state's Internet can be proxied simply by how fast data flows through the Internet within its state borders. The two preceding variables – IXP density and state restriction on ISPs – do in some part affect the Internet speed in the state. At the same time, it has been shown that Internet speed reflects primarily the requisite infrastructure, such as cables and fibers, into which the state decides to invest.[7] This means that we would be able to glean from Internet speed an aspect of cyberspace border orientation that is distinct from either IXP placements or ISP restrictions.

To measure a state's Internet circularity, I log the 2023 broadband Internet downloading speed data[8] as obtained from the Speedtest Global Index[9]. Broadband Internet speeds are chosen over mobile Internet speeds for availability for a greater number of states.

[5]"Freedom on the Net Research Methodology".

[6]I reverse the sign of this 0-6 ordinal svariable in my results for a more intuitive interpretation where a higher value indicates a higher degree of Concentricity.

[7]See, for example, Carmen Ang, "Mapped: The Fastest (And Slowest) Internet Speeds in the World", *Visual Capitalist*, September 29, 2021.

[8]"Internet Speeds by Country 2023", *World Population Review*.

[9]Speedtest Global Index.

2.4 Political and economic variables

For analysis of main results, I further collect data on regime type from the 2018 Polity Project[10], and on GDP per capita from the World Bank for 2021, for all states in the sample.[11] Summary statistics for the three Internet infrastructure variables and the two political and economic variables are presented in Table 1 for the 57 states.

Table 1: Summary statistics for Internet infrastructure and political/economic attributes

Statistic	N	Mean	St. Dev.	Min	Max
Connectivity (Logged IXP/population)	57	−16.110	1.268	−18.671	−12.715
Concentricity (ISP restriction index)	57	-3.684	-1.311	0	-6
Circularity (Logged broadband speed)	57	3.974	0.933	2.069	5.545
Logged GDP/capita	57	8.781	1.333	6.453	11.195
Polity score	57	3.596	6.478	−10	10

[10] "The Polity Project", *Center for Systemic Peace*.
[11] "GDP per capita (current US$)", *The World Bank*.

3 Methodology

As stated, the objective of this project is to discover underlying patterns in cyberspace border orientation among states that may be distinct from currently understood fault lines. To this end, I employ a k-means clustering method, which is a type of unsupervised algorithm, using data on Internet Connectivity, Concentricity, and Circularity. This allows for a subsequent analysis with respect to regime type and economic development without letting either dictate the generation of the patterns themselves. As a first step, I use the elbow method to determine the optimal number of clusters to be generated based on the within-cluster sum of square (WCSS), as shown in Figure 1.

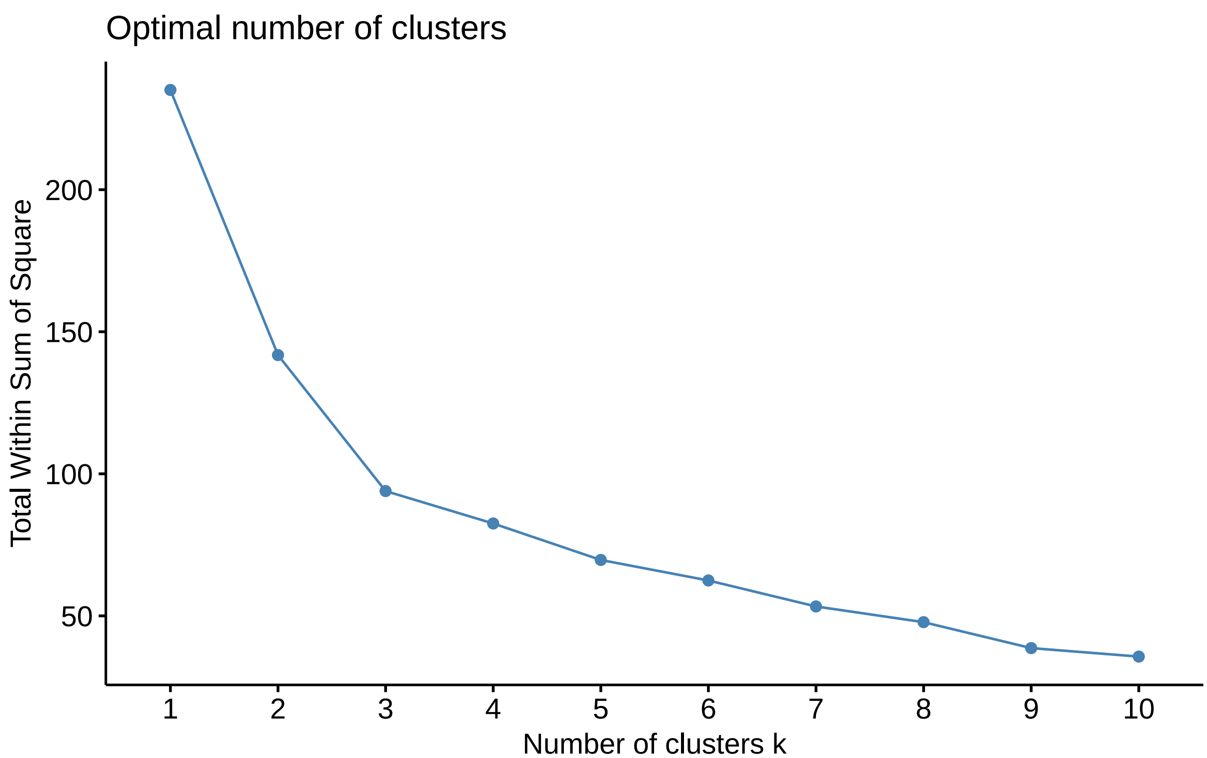

Figure 1: Elbow plot based on within-cluster sum of square with data on Connectivity, Concentricity, Circularity

The plot indicates that the optimal number of clusters in this case is 3. I then conduct a k-means cluster analysis on the 57 states in three clusters based on the three Internet infrastructure attributes.

4 Results

4.1 Descriptive Statistics by Cluster

The k-means clustering algorithm yields three clusters of 23, 24, and 9 states, respectively. Table 2 below lists the mean values of the three attributes for each of the three clusters.

Table 2: Descriptive Statistics for Attribute Means by Cluster

Cluster	N	Connectivity	Concentricity	Circularity
1	23	-14.98911	-4.541667	4.482371
2	24	-17.07740	-3.750000	3.389486
3	9	-16.51739	-1.222222	4.175564

Based on these, we see that the three clusters display fairly distinct cyberspace border orientations as proxied by the three Internet infrastructure attributes: Cluster 1 is characterized by a high degree of Connectivity, a low degree of Concentricity, and a high degree of Circularity. Substantively, states in this cluster generally enjoy a dense network of peering points, a lightly regulated ISP market, and high Internet speeds.

Cluster 2, meanwhile, is characterized by a low degree of Connectivity, a medium degree of Concentricity, and a low degree of Circularity. States in this cluster have many fewer peering points per capita, the slowest Internet speeds of all, and an ISP market that displays an intermediate level of state control.

Finally, Cluster 3 is characterized by a medium degree of Connectivity, the highest degree of Concentricity, and a fairly high degree of Circularity. Here, states enjoy a moderate density of peering points, fairly high Internet speeds, and by far the tightest state control of ISPs.

Next, I analyze these results with respect to both the regime types and the levels of economic development for the states in question. From there, we will see how cyberspace border orientation attributes do or do not correlate with the political and economic orientations of the states.

4.2 Correlation with political regime type

Table 3 below depicts the three clusters of states in their cyberspace border orientation and each state's regime type: Color blue denotes a democracy, color purple an anocracy, and color red an autocracy, as per definitions in the Polity Project.

Table 3: Correlation of State Clustering with Regime Type

Cluster	1	2	3
States	Argentina	Angola	Belarus
	Armenia	Bangladesh	China
	Australia	Colombia	Kazakhstan
	Bahrain	Ghana	Lebanon
	Brazil	India	Myanmar
	Cambodia	Indonesia	Saudi Arabia
	Canada	Iraq	United Arab Emirates
	Costa Rica	Jordan	Uzbekistan
	Ecuador	Kenya	Vietnam
	Estonia	Malawi	
	France	Mexico	
	Gambia	Morocco	
	Georgia	Nigeria	
	Germany	Pakistan	
	Hungary	Philippines	
	Italy	Rwanda	
	Japan	South Africa	
	Malaysia	Sri Lanka	
	Serbia	Sudan	
	Singapore	Thailand	
	South Korea	Tunisia	
	Ukraine	Uganda	
	United Kingdom	Zambia	
	United States	Zimbabwe	

A few distinctive observations emerge from the results: First, autocracies are overrepresented in Cluster 3, making up almost 80% of the total. This compares with just one autocracy in each of the two other clusters

Second, in contrast, there are no appreciable differences between Cluster 1 and Cluster 2 in terms of the regime types being represented. Both are largely comprised of democracies, along with a small number of anocracies. Whereas autocracies display highly similar cyberspace border orientation, there is an even split among democracies into two clusters of rather different orientations.

Taking into account the aforementioned cluster characteristics, one gains further insights. The concentration of autocracies in Cluster 3 suggests that citizens in these states generally enjoy fairly high Internet speeds within their borders as the states maintain a tight grip on the ISPs. At the same time, the split of democracies into the two clusters suggests a divergence in experiences for these citizens. While those in Cluster 1 enjoy both very high Internet speeds, very little state control, and a high density of peering points with the rest of the world, those in Cluster 2 fare remarkably worse on all three fronts. Notably, except for the somewhat lesser state control of ISPs, citizens in Cluster 2 seem to be all in all worse off than their autocratic counterparts as far as the daily Internet experience is concerned.

Meanwhile, the results hint at the economic dimension that may be at play, where the clustering may correlate more strongly with the level of development than with regime type. To this I now turn.

4.3 Correlation with economic development

Table 4 below depicts the three clusters of states and their respective classifications of economic development: Color olive denotes a high-income economy, color teal a middle-income economy, and color cyan a low-income economy, as per country classifications in the United Nations' World Economic Situation and Prospects 2022 (Belsey, 2022).

The results suggest a tentative but imperfect relationship of cyberspace border orientation and economic development. First, the fact that all Cluster 2 states are middle- or low-income economies indicates the likely role of state capacity in dictating a state's cyberspace border orientation. The lesser degree of Concentricity for this cluster when compared to Cluster 3, for example, may be due as much to a lack of capability as to that of willingness.

Nonetheless, economic development does not tell the full story of cyberspace border orientation. Gambia, for instance, has adopted an orientation akin to the vast majority of the most developed states in the world, its own low level of development notwithstanding. On the other hand, the richest countries find themselves in both Clusters 1 and 3 with highly

Table 4: Correlation of State Clustering with Economic Development

Cluster	1	2	3
States	Argentina	Angola	Belarus
	Armenia	Bangladesh	China
	Australia	Colombia	Kazakhstan
	Bahrain	Ghana	Lebanon
	Brazil	India	Myanmar
	Cambodia	Indonesia	Saudi Arabia
	Canada	Iraq	United Arab Emirates
	Costa Rica	Jordan	Uzbekistan
	Ecuador	Kenya	Vietnam
	Estonia	Malawi	
	France	Mexico	
	Gambia	Morocco	
	Georgia	Nigeria	
	Germany	Pakistan	
	Hungary	Philippines	
	Italy	Rwanda	
	Japan	South Africa	
	Malaysia	Sri Lanka	
	Serbia	Sudan	
	Singapore	Thailand	
	South Korea	Tunisia	
	Ukraine	Uganda	
	United Kingdom	Zambia	
	United States	Zimbabwe	

contrasting cyberspace border orientations. This is indicative of a possible interaction effect of regime type and economic development.

4.4 Discussion

Together, the preliminary findings make for two observations. First, cyberspace border orientation for a given state is not wholly dictated by either political orientation or state capacity. Second, the resulting pattern from the cluster analysis suggests a possible interaction between regime type and state capacity. One can interpret it as primarily determined by state capacity, which splits states into either Cluster 2 (low state capacity) or Clusters 1 and 3 (high state capacity). Thereafter, the regime type dictates which orientation that a high-capacity state adopts. Alternatively, one may also interpret it as primarily determined by regime type, which splits states into either Cluster 3 (autocratic) or Clusters 1 and 2 (democratic). Thereafter, the level of development dictates which orientation that a democratic state adopts.

Whichever interpretation, the findings attest to a high degree of heterogeneity among democracies, consistent with what motivated this project. A finer understanding of this heterogeneity – its sources, correlates, and dimensions along which it occurs – would be worthwhile amid the ongoing calls for uniting democracies to counter digital authoritarianism worldwide (Cohen & Fontaine, 2020; Donahoe & Polyakova, 2020). Certain fault lines within democracies on cyberspace have been duly noted, notably the transatlantic divide (Cole & Fabbrini, 2016; Chander & Sun, 2022; Cooper, 2019). More than the substantive findings, this project seeks to show that a data-driven approach helps uncover additional dimensions along which democracies diverge that may have eluded public discourse. The next section outlines the next steps for this project.

5 Next steps

Data collection and analyses are underway for the following two groups of variables.

5.1 Additional attributes of cyberspace border orientation

Two additional attributes of subterranean cyberspace border orientation are currently being coded. First, I aim to collect data on the number, positioning, and density of submarine and overland cables for each state. These variables will complement the existing attributes in rendering a fuller profile of cyberspace border orientation that includes the bottom-most, physical layer.

The second attribute being coded is the aggregate routing history of Internet traffic for a given state. In particular, I plan to measure travel outside of the state for a given packet before it reaches a user within the state's borders. Whether a state's Internet traffic traverses the world before reaching the domestic user, or it never leaves its own borders, would make for vastly different cyberspace border orientations. In coding this variable I will build upon existing measures in computer science that have been compiled, such as hop counts, in a variety of technical and geographical contexts (Begtasevic & Van Mieghem, 2001; Gupta et al., 2014; Obar & Clement, 2012; Shah, Fontugne, & Papadopoulos, 2016).

5.2 Additional correlates of cyberspace border orientation

The second group of variables concerns correlates of cyberspace border orientation besides regime type and economic development. In addition to other state-level variables, such as geographical regions, I aim to examine the correlation between clustering in cyberspace border orientation and other clusterings that have been studied in political science. These include, to name a few, trade blocs, interstate conflict, and defense alliances. Border orientation is an inherently geospatial concept. A fuller profile of cyberspace border orientation thus necessitates a look beyond individual state-level attributes.

References

Begtasevic, F., & Van Mieghem, P. (2001). Measurements of the hopcount in internet. In Pam2001, a workshop on passive and active measurements, amsterdam, the netherlands, april 23-24, 2001.

Belsey, M. A. (2022). Content type. World Economic Situation and Prospects, 2022.

Carter, D. B., & Poast, P. (2017). Why do states build walls? political economy, security, and border stability. Journal of conflict resolution, 61(2), 239–270.

Chander, A., & Sun, H. (2022). Sovereignty 2.0. Vand. J. Transnat'l L., 55, 283.

Chatzis, N., Smaragdakis, G., Feldmann, A., & Willinger, W. (2013). There is more to ixps than meets the eye. ACM SIGCOMM Computer Communication Review, 43(5), 19–28.

Cohen, J., & Fontaine, R. (2020). Uniting the techno-democracies: How to build digital cooperation. Foreign Aff., 99, 112.

Cole, D., & Fabbrini, F. (2016). Bridging the transatlantic divide? the united states, the european union, and the protection of privacy across borders. International Journal of Constitutional Law, 14(1), 220–237.

Cooper, Z. (2019). Bridging the transatlantic divide on china. Statement before the House Foreign Affairs Committee Subcommittee on Europe, Eurasia, Energy, and the Environment on China's Expanding Influence in Europe and Eurasia.

DeNardis, L. (2012). Governance at the internet's core: The geopolitics of interconnection and internet exchange points (ixps) in emerging markets.

DeNardis, L. (2014). The global war for internet governance. Yale University Press.

Donahoe, E., & Polyakova, A. (2020). A transatlantic effort to take on china starts with technology. Center for European Policy Analysis. Article Retrieved May, 25, 2021.

Dragu, T., & Lupu, Y. (2021). Digital authoritarianism and the future of human rights. International Organization, 75(4), 991–1017.

Farrell, H., & Newman, A. (2016). The transatlantic data war: Europe fights back against the nsa. Foreign Affairs, 95(1), 124–133.

Floridi, L. (2020). The fight for digital sovereignty: What it is, and why it matters, especially for the eu. Philosophy & technology, 33, 369–378.

Gregori, E., Improta, A., Lenzini, L., & Orsini, C. (2011). The impact of ixps on the as-level topology structure of the internet. Computer Communications, 34(1), 68–82.

Gupta, A., Calder, M., Feamster, N., Chetty, M., Calandro, E., & Katz-Bassett, E. (2014). Peering at the internet's frontier: A first look at isp interconnectivity in africa. In Passive and active measurement: 15th international conference, pam 2014, los angeles, ca, usa, march 10-11, 2014, proceedings 15 (pp. 204–213).

Guriev, S., & Treisman, D. (2022). Spin dictators: The changing face of tyranny in the 21st century. Princeton University Press.

King, G., Pan, J., & Roberts, M. E. (2017). How the chinese government fabricates social media posts for strategic distraction, not engaged argument. American political science review, 111(3), 484–501.

Kurbalija, J. (2016). An introduction to internet governance. Diplo Foundation.

Land, M. K. (2019). Against privatized censorship: proposals for responsible delegation. Va. J. Int'l L., 60, 363.

Munger, K., Bonneau, R., Nagler, J., & Tucker, J. A. (2019). Elites tweet to get feet off the streets: Measuring regime social media strategies during protest. Political Science Research and Methods, 7(4), 815–834.

Nabi, Z. (2014). Censorship is futile. arXiv preprint arXiv:1411.0225.

Obar, J. A., & Clement, A. (2012). Internet surveillance and boomerang routing: A call for canadian network sovereignty. In Tem 2013: Proceedings of the technology & emerging media track-annual conference of the canadian communication association (victoria.

Pohle, J., & Thiel, T. (2020). Digital sovereignty. Pohle, J. & Thiel.

Roberts, M. E. (2018). Censored. Princeton University Press.

Seltzer, W. (2008). The politics of internet control and delegated censorship. American Society of International Law, April.

Shah, A., Fontugne, R., & Papadopoulos, C. (2016). Towards characterizing international routing detours. In Proceedings of the 12th asian internet engineering conference (pp. 17–24).

Simmons, B. A. (2005). Rules over real estate: Trade, territorial conflict, and international borders as institution. Journal of Conflict Resolution, 49(6), 823–848.

Simmons, B. A., & Kenwick, M. R. (2021). Border orientation in a globalizing world.

Sun, T., & Zhao, Q. (2022). Delegated censorship: The dynamic, layered, and multistage information control regime in china. Politics & Society, 50(2), 191–221.

Tăbucă, S., et al. (2010). The internet access as a fundamental right. Journal of Information Systems and Operations Management, 4(2), 206–212.

Wachs, M., Schanzenbach, M., & Grothoff, C. (2014). On the feasibility of a censorship resistant decentralized name system. In Foundations and practice of security: 6th international symposium, fps 2013, la rochelle, france, october 21-22, 2013, revised selected papers (pp. 19–30).