

FUNDAÇÃO GETULIO VARGAS
SÃO PAULO SCHOOL OF BUSINESS ADMINISTRATION

UNIVERSITÀ DELLA SVIZZERA ITALIANA
INSTITUTE OF COMMUNICATION AND PUBLIC POLICY

FABIANO ANGÉLICO

**OPEN GOVERNMENT DATA AS AN ANTICORRUPTION INTERVENTION IN
GLOBAL CITIES: DO LOCAL-LEVEL ACCOUNTABILITY ACTORS ACTUALLY
USE DATA?**

SÃO PAULO AND LUGANO

2023

FABIANO ANGÉLICO

**OPEN GOVERNMENT DATA AS AN ANTICORRUPTION INTERVENTION IN
GLOBAL CITIES: DO LOCAL-LEVEL ACCOUNTABILITY ACTORS ACTUALLY
USE DATA?**

Thesis submitted to the Fundação Getúlio Vargas's São Paulo School of Business Administration of São Paulo (FGV-EAESP) and to the Institute of Communication and Public Policy of Università della Svizzera Italiana (ICPP-USI) in partial fulfillment of the requirements for the degree of Doctor in Public Administration and Government (FGV EAESP) and Doctor in Communication Sciences (ICPP-USI)

Area of Concentration: Analysis,
Administration, Information Technology, Public
Policy, Communication

Supervisor: Prof. Dr. Maria Alexandra Cunha
Co-Supervisor: Prof. Dr. Jean-Patrick
Villeneuve

SÃO PAULO AND LUGANO

2023

FABIANO ANGÉLICO

**OPEN GOVERNMENT DATA AS AN ANTICORRUPTION INTERVENTION IN
GLOBAL CITIES: DO LOCAL-LEVEL ACCOUNTABILITY ACTORS ACTUALLY
USE DATA?**

Thesis submitted to the Fundação Getúlio Vargas's São Paulo School of Business Administration of São Paulo (FGV-EAESP) and to the Institute of Communication and Public Policy of Università della Svizzera italiana (ICPP-USI) in partial fulfillment of the requirements for the degree of Doctor in Public Administration and Government (FGV EAESP) and Doctor in Communication Sciences (ICPP-USI)

Areas of Concentration: Public Administration and Government, Communication Sciences, Government and Civil Society in a Subnational Context, Public Integrity

Supervisor: Prof. Dr. Maria Alexandra Cunha
Co-Supervisor: Prof. Dr. Jean-Patrick Villeneuve

Date: 23/02/2023

Examination committee

Prof. Dr. Maria Alexandra Cunha
(supervisor) FGV-EAESP

Prof. Dr. Jean-Patrick Villeneuve
(co-supervisor) ICPP-USI

Prof. Dr. Marco Antonio Carvalho Teixeira
FGV-EAESP

Prof. Dr. Paula Chies Schommer
ESAG-UDESC

Prof. Dr. Elina De Simone
Università Roma Tre

ACKNOWLEDGEMENTS

Academic theses bring the name of one sole author. However, no one writes such a piece of knowledge alone. This thesis results from years of professional and academic experience, a trajectory facilitated by many generous masters. Firstly, I wanted to thank my supervisors, Prof Maria Alexandra Cunha, from FGV EAESP, and Prof. Jean-Patrick Villeneuve, from USI ICPP, for their invaluable support during this doctorate. This process was a long and winding road in the middle of a global pandemic. During the whole period, my two supervisors were supportive, collaborative, and patient — my huge thank you to both of you.

At USI and FGV, several professors have generously shared their time and knowledge with me. Thank you very much, Giulia Mugellini, Marlen Heide, Alberto Bittonti, Jolanta Drzewiecka, Jean-François Savard, Isabelle Caron, Annamaria Astrologo, Andre Lecours. Also huge thanks to Mario Aquino, José Antonio Puppim, Marco Antonio Carvalho Teixeira, Marlei Pozzebon, Luís Paulo Bresciani, Eduardo Grin, André Guzzi, Gabriella Lotta, Maria Rita Loureiro, Marta Farah, Claudio Couto.

I also have to acknowledge two Universities that, before FGV and USI, contributed to my academic and professional trajectory: UFMG, in Brazil, where I completed my major in Journalism, and the University of Chile, where I dedicated to postgraduate studies on transparency, accountability, and anticorruption. In this regard, I thank my undergrad supervisors Carmen Vieira and Elton Antunes (UFMG) as well as Constanza Toro, Silvana Lauzán, and the late Prof. Jose Zalaquett (Universidad de Chile).

Public sector and not-for-profit leaders have also been crucial to my professional and academic development. I have to thank Mario Vinicius Spinelli, who trusted me for strategic positions at Sao Paulo City Hall and taught me a lot about accountability institutions. Similarly, I need to acknowledge the late Claudio Weber Abramo, who guided me in learning about transparency and integrity from the non-profit perspective.

I must also thank CAPES, a government agency within the Ministry of Education in Brazil. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Last but not least, I have to thank my family for all their support.

ABSTRACT

Transparency and open data are regarded as means for improving governance, including control of corruption. For this reason, there has been an explosion of access to information laws and open data portals in the last 25 years. However, a crucial aspect of the interrelation between open government data and corruption control is still understudied: data use. This thesis investigates the uses and barriers to the use of open government data to counter corruption in a megacity — São Paulo, the largest city in Brazil and South America — focusing on a dataset that the local government made publicly available for six years (2016-2022): urban land and real estate ownership. This thesis builds upon various source, including documentation and semi-structured interviews. From document analysis and direct experience, followed by validation with knowledgeable actors, it has been noted that accountability organizations, both from civil society and government, have not systematically used the urban land real estate dataset as anticipated. In-depth interviews with 20 top-level accountability actors, totaling nearly 16 hours, have been conducted to investigate these professionals' perceptions about data access and use. The results show that the disclosure of the dataset informing details about urban land and real estate, such as the name of property owners, has had limited use on corruption control. Building upon the concept of “publicity and accountability condition” (Lindstedt & Naurin, 2010) to assess the relationship between transparency and corruption added to elements of historical institutionalism, I propose the concept of “socio-political and techno-legal conditions” for analyses of open data use on producing desired outcomes toward corruption control. The disclosure of data will not bring about expected and desired effects unless there are: i) social consensus around the value of integrity and political incentives towards anticorruption institutional building and ii) technical and legal conditions for effective data use for legal and administrative sanctioning mechanisms. This thesis contributes to theory building in three ways: the emerging literature on open data; the academic debate on the conceptualization and mis-conceptualization of corruption; and the scientific knowledge on local-level corruption. Regarding practice, this thesis argues for more sophisticated designs of transparency and open data policies for accountability, particularly at the subnational level.

Keywords: open data; transparency; open government data; accountability; data use; corruption; public sector; data for development; local governments.

RESUMO

Transparência e dados abertos são considerados meios para a boa governança, incluindo o controle da corrupção. Por esta razão, houve uma explosão de leis de acesso à informação e de portais de dados abertos nos últimos 25 anos. Contudo, um aspecto crucial da relação entre dados abertos governamentais e controle da corrupção é pouco estudado: o uso dos dados. Esta tese investiga os usos e barreiras para usos de dados governamentais abertos para combater a corrupção numa megacidade — São Paulo, a maior cidade do Brasil e da América do Sul — com foco em um conjunto de dados que a Prefeitura local disponibilizou publicamente durante seis anos (2016-2022): propriedade de imóveis e de terras urbanas. Esta tese baseia-se em várias fontes, incluindo documentação e entrevistas semiestruturadas. Da análise documental e experiência direta, seguida de validação com atores experientes, observa-se que órgãos e instituições de controle, tanto da sociedade civil como do estado, não utilizaram sistematicamente este conjunto de dados como previsto. Foram realizadas entrevistas aprofundadas com 20 profissionais experientes para investigar as percepções desses atores sobre acesso e utilização de dados. Os resultados mostram que a divulgação do conjunto de dados informando detalhes sobre terrenos urbanos e imóveis, tais como o nome dos proprietários, teve uso limitado para controle da corrupção. Com base no conceito de "condições de publicidade e responsabilização" (Lindstedt & Naurin, 2010) acrescida a elementos do institucionalismo histórico, proponho o conceito de "condições sociopolíticas e tecno-jurídicas" para analisar o papel da abertura de dados no controle da corrupção. A divulgação de dados não produzirá efeitos esperados e desejados sem i) consenso social em torno do valor da integridade e incentivos políticos para o desenvolvimento de um arcabouço institucional anticorrupção e ii) condições técnicas e jurídicas para o uso eficaz de dados em mecanismos sancionatórios. Esta tese contribui para a construção teórica de três formas: a literatura emergente sobre dados abertos, a conceitualização de corrupção, e o conhecimento científico corrupção em nível local. No que diz respeito à prática, esta tese sugere a relevância de se construir políticas públicas de transparência e de dados abertos mais sofisticadas para fins de controle, particularmente em nível subnacional.

Palavras-chave: dados abertos; transparência; dados governamentais abertos; responsabilização; uso de dados; corrupção; setor público; dados para o desenvolvimento; governos locais.

LIST OF FIGURES

Figure 1 – The expected outcome of the disclosure of data in São Paulo	21
Figure 2 – Open data for development – the World Bank’s approach.....	49
Figure 3 – Diagram of the systematic literature review on “open data” and “corruption”, according to the PRISMA approach	55
Figure 4 – National and state level anticorruption “communities of practice” relevant to the context of São Paulo City	81
Figure 5 – Reproduction of the official CGM-SP webpage announcing the disclosure of the IPTU dataset as an anticorruption initiative	86
Figure 6 – Reproduction of an official document from the São Paulo City Hall’s IT Department (translation below)	89
Figure 7 – Challenges for open data to more likely produce expected outcomes on corruption control in megacities	140

LIST OF TABLES

Table 1 – Historical summary of the case.....	18
Table 2 – Exploratory literature review on corruption	24
Table 3 – Theories of corruption	29
Table 4 – Systematic literature review on corruption and HI – summary.....	39
Table 5 – Number of publications meeting the search criteria – open data & corruption....	55
Table 6 – Journals that published the 20 selected articles	57
Table 7 – Articles on open data and corruption arguing for likely positive outcomes of open government policies	58
Table 8 – Articles on open data and corruption establishing feasible ambitious for open government policies.....	60
Table 9 – Summary of the literature review on Open Data and Corruption.....	62
Table 10 – Transparency International’s CPI (Corruption Perception Index) for the G20 countries of the “Global South”.....	63
Table 11 – Theoretical perspectives used in the thesis.....	66
Table 12 – List of accountability actors interviewed (encoded).....	72
Table 13 – The “publicity condition”, according to data-driven journalists and CSO representatives	110
Table 14 – The “accountability condition”, according to data-driven journalists and CSO representatives	112
Table 15 – Persistence of corruption, according to the perceptions of diagonal accountability actors	118
Table 16 – The “publicity condition”, according to the perceptions of horizontal accountability actors	127
Table 17 – The “accountability condition”, according to the perceptions of horizontal accountability actors	128
Table 18 – Persistence of corruption, according to the perceptions of horizontal accountability actors	137

ACRONYMS

ATI	Access to Information
CGE-MG	Controladoria Geral do Estado de Minas Gerais (Office of the Comptroller General of Minas Gerais State)
CGE-SP	Controladoria Geral do Estado de São Paulo (Office of the Comptroller General of São Paulo State)
CGM-SP	Controladoria Geral do Município de São Paulo (Office of the Comptroller General of São Paulo City)
CGU	Controladoria Geral da União (Office of the Comptroller General of the Union)
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Coordination for the Improvement of Higher Education Personnel)
ENCCLA	Estratégia Nacional de Combate à Corrupção e à Lavagem de Dinheiro (National Strategy to Fight Corruption and Money Laundering)
FoccoSP	Fórum de Combate à Corrupção e Lavagem de Dinheiro do Estado de São Paulo (Forum to Fight Corruption and Money Laundering of the State of São Paulo)
FATF	The Financial Action Task Force
FS	Agência Fiquem Sabendo (Get to Know – journalism agency specialized in ATI)
GDP	Gross Domestic Product
HI	Historical Institutionalism
IPTU	Imposto Predial, Territorial e Urbano (Property, Land and Urban Tax)
MPF	Ministério Público Federal (Federal Prosecutors' Office)
MPSP	Ministério Público do Estado de São Paulo (São Paulo State Prosecutors' Office)
NGO	Non-Governmental Organization
OGP	Open Government Partnership
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSM	Public Sector Motivation
TCM-SP	Tribunal de Contas do Município de São Paulo (São Paulo City Court of Auditors)
TJSP	Tribunal de Justiça de São Paulo (São Paulo State Court)

SUMMARY

1	INTRODUCTION	11
1.1	Theme, Research Problem, and Objectives	14
1.2	The Case of São Paulo, Brazil	17
1.3	Diagonal and Horizontal Accountability	21
2	THEORETICAL-EMPIRICAL FOUNDATION	23
2.1	Corruption: a Literature Review	23
2.1.1	<i>Systematic Literature Review – Corruption and Historical Institutionalism</i>	31
2.2	Transparency and Accountability – “Transparency Dynamics”, Targets and “More Complex Accountability”	42
2.3	Open Data for Development – Not All Data Are The Same	48
2.4	Open Data and Corruption – a Systematic Literature Review	53
2.5	Open Data, Accountability and Corruption: “The Publicity and Accountability Conditions” and “Corruption Persistence”	63
3	METHODOLOGY	67
3.1	Epistemological Positioning	67
3.2	Research Design	67
3.3	Data Collection	69
3.4	Procedures for Data Analysis	75
3.5	Validation Criteria	77
4	PRESENTATION OF THE RESULTS	78
4.1	The Transparency and Anticorruption “Communities of Practice” in São Paulo	78
4.2	Horizontal Accountability Institutions: What They Proactively Inform About Data Use	84
4.3	Civil Society and Media: Greater Accountability, Better Policy Analysis or Increased Business Opportunities?	96
4.4	The Perceptions of Diagonal Accountability Actors	98
4.5	The perceptions of horizontal accountability actors	118
4.6	From “publicity and accountability conditions” to “techno-legal and socio-political conditions”	138

5	DISCUSSIONS OF THE RESULTS	141
5.1	Social and Political Conditions.....	143
5.2	Technological and Legal Conditions	145
5.3	Theoretical and Policy Implications	147
6	CONCLUDING REMARKS	154
	REFERENCES	157
	ANNEXES	173
	Annex 1 – Questionnaire to Government Officials.....	173
	Annex 2 – Questionnaire to Civil Society Representatives	175

1 INTRODUCTION

Disclosing information produced or held by the governments is seen as having transformative potential toward better policy analysis, economic improvement, and greater accountability. These assumptions led to the expansion of the transparency agenda and an overwhelming increase in national laws on access to government information. In only 30 years, the number of countries adopting legislation on transparency multiplied nearly by ten: from 13 in 1990 to 127 in 2020 (UNESCO et al., 2020).

With the development of information and communication digital tools, the transparency agenda, which initially was exclusively about the right to access print documents held by government agencies, included calls for governments to publish data online. Open government data, added to the increasing access to the internet, could finally make governments and society reach development goals, including greater accountability and countering corruption.

Studies about transparency and corruption are abundant. However, the literature on open data, which includes the technological aspect (beyond the right to access printed archival documents physically), and its role on corruption control is still limited. In a study on emerging technologies and their impact on countering corruption, Adam and Fazekas found that “rigorous evidence on the impact of big and open data on transparency portals is still relatively scarce”, concluding that “we need a more nuanced and detailed understanding of how transparency portals providing open data can have an impact on corruption” (2021, p. 10).

A systematic literature review (PRISMA method) undertaken for this thesis found two initial sets of evidence from the few studies on open data and corruption published. Firstly, although there has been an extensive amount of data disclosed by governments in recent years, specific datasets crucial for monitoring corruption — such as data on procurement, on lobbying, on company registers, on land property — are rarely available in open formats, cost-free, on internet (Laboutková, 2018; Žuffová, 2020). Second, in the rare specific contexts in which some of these types of data are available, case studies bring evidence that “publicity and accountability conditions” are needed if open data is to have some impact on corruption control (Adu, 2018; Murillo, 2015; Park & Kim, 2020; Schnell, 2018).

The concept of “publicity and accountability conditions” has been put forward by Lindstedt and Naurin (2010). These authors argue that, from the principal-agent perspective, transparency is not enough to curb corruption because the information made available needs to reach the public. These actors need to be able to process such information — the publicity condition. “Furthermore, if the release of information to the public is to affect the behavior of

potentially corrupt government officials, the public must possess some sanctioning mechanism. This is the accountability condition” (2010, p. 302). Therefore, “two conditions must be present for transparency to have an effect on reducing corruption: that the principal is (a) enabled to acquire and process information, and (b) enabled to act based on the newly acquired knowledge” (Murillo, 2015, p. 44).

Adding to the notion of “publicity and accountability conditions” and discussing specifically data (rather than the broader notion of information), Hulstijn et al. (2017) argue that open data for accountability to fight against corruption is dependent mainly on a forum of actors who i) are critical; ii) have the technical knowledge to explore the data; and iii) are free to challenge authorities.

In this vein, this thesis will analyze the relationship between open government data and corruption, exploring the control and regulatory power of the former over the latter. Open government data can be defined as non-private and non-confidential data, produced with state resources, and made publicly available without restriction on use or distribution (Janssen et al., 2012). The prevalent definition of corruption is the abuse of entrusted power for private gains (Pozsgai-Alvarez, 2020).

The thesis will use a combination of two analytical lenses to analyze the role of open government data on corruption control.

Firstly, it will build upon the concept of “publicity and accountability conditions” and explore the nuances of each element separately: “publicity condition”, understood as a capacity to receive and process information, and “accountability condition”, whose core aspect relates to using evidence to activate sanctioning mechanisms.

While the concept of “publicity and accountability condition” is useful to examine the process of collecting and using data, it may not be helpful to analyse more complex challenges related to power asymmetry or social and political environment. Therefore, another analytical lens will be used in addition: historical institutionalism elements will be deployed to analyze the persistence of corruption, even in a context where transparency and access to data are present.

This thesis is an instrumental case study in this thesis. According to Stake (1998), instrumental case studies aim to provide insight into an issue or redraw a generalization: the case plays a supportive role, as it facilitates our understanding of something else. The case this thesis covers is the São Paulo City Hall initiative, announced on the 2015 International Anticorruption Day, to publicly disclose a specific dataset aiming at curbing corruption in the city.

Considering the “publicity and accountability conditions” on the ground, the investigation will explore if accountability actors used this specific data for controlling corruption in a megacity in the Global South, focusing on the period between 2016 and 2022, the years when the dataset was fully available in the public domain.

To further explore the relationship between open government data and corruption — beyond the conditions for accountability actors to perform their tasks — this thesis will use elements from the historical institutionalism framework. The goal is to analyze the persistence of corruption even in contexts where open government is a priority, and open data policies have been implemented.

This work will explore the case of São Paulo, Brazil, and a dataset on urban land and real estate property. The city government disclosed this dataset in March 2016 after a public announcement by the city's mayor months earlier in a public event marking international anticorruption day (December 9th). However, in the first quarter of 2022, the dataset was partially removed — an interpretation of the new national privacy law led to the removal of property owners' names from the publicly available IPTU dataset.

The IPTU, an acronym in Portuguese for “urban land and real estate property taxation”, is a database that presents a critical piece of information for corruption control: the name of real estate and urban land title owners. As literature empirically demonstrates (Laboutková, 2018; Žuffová, 2020), having corruption-related datasets publicly available online is not usual. Hence this specific situation in São Paulo, in which a crucial anticorruption dataset is publicly available, provides a rich opportunity for learning about the role of open data on countering corruption.

Moreover, São Paulo is a global city, the largest in South America and the entire Southern Hemisphere, using different metrics (population, GDP etc). This Brazilian metropolis is also a member of megacities coalitions (such as C40, a global platform on climate leadership) and a pioneer member of the Open Government Partnership (OGP), an international multistakeholder platform on transparency, civic participation, and innovation. All of this makes the São Paulo case an exceptional opportunity to learn.

Therefore, learning about the role of this open data intervention to counter corruption in megacities like São Paulo by analyzing the uses of the data adds to the current scientific knowledge, specifically in three ways:

- It contributes to the literature on open data and its role on development, particularly making the case of differentiating the uses of data for anticorruption purposes from other less politically sensitive uses, such as policy analysis.

- This thesis adds to the vast literature on corruption studies, but mainly to the growing corpus of research on local-level corruption, specifically on megacities, those municipalities actively participating in international fora and transnational coalitions and platforms that have become relevant actors in the global arena.
- The thesis also contributes to the current theoretical debate on corruption, adopting an underexplored lens to analyze corruption: historical institutionalism.

Lastly, the thesis will also have implications in policy discussions because it brings a more nuanced and sophisticated analysis of the potentials and limits of open data interventions in countering corruption.

1.1 Theme, Research Problem, and Objectives

In their recent book entitled “Government Transparency: State of the Art and New Perspectives”, Porumbescu et al. (2022) propose the following definition:

“Transparency is the availability of information about an organization or actor allowing external actors to monitor the internal workings or performance of that organization” (2022, p. 10). This “visibility” on the functioning of the institutions must be clear enough for external actors to be able to produce accurate inferences from the available data (Michener & Bersch, 2013). Such inferences will be critical for analysis that will possibly lead to accountability.

Transparency and accountability are close concepts as accountability “broadly denotes the duty of an individual or organisation to answer in some way about how they have conducted their affairs” (Hood & Heald, 2006, p. 989). In that sense, transparency is often viewed as an enabler of accountability.

Jonathan Fox's article “the uncertain link between transparency and accountability” (2007) offers, however, the idea that not all transparency would lead to desirable levels of accountability. The author proposed the concept of “clear” transparency (in opposition to “opaque” transparency) and “hard” accountability (to differentiate it from “soft accountability”). Fox defines “clear transparency” as information-access policies and programmes that reveal reliable information about institutional performance, responsibilities, and spending.

However, reaching “clear transparency” — that is, reliable and useful information on public institutions' activities — is a relevant first challenge, particularly at the subnational level. Coelho et al. (2018), in a paper on proactive transparency policies in Brazilian states and large cities (state capitals and municipalities with more than 400 thousand inhabitants), found what they called “incomplete transparency”: general basic information is available, but more specific

data, about the budget, for instance, were still lacking. The data was collected in 2017, five years after the access to information legislation had come into force in Brazil.

Even when higher levels of transparency are in place, “hard accountability”, which Fox defines as “answerability plus the possibility of sanctions”, is not a given. Information might not be the most crucial aspect in transforming transparency into accountability – that is, universal access to government data may not automatically lead to answerability and possibilities of sanctions. McGee (2019), for instance, argues that accountability challenges are a lot more about power asymmetry than information asymmetry.

In that sense, although the notions of transparency and accountability have for long been linked to the idea of good governance, suggesting that more information combined with sanctions would reduce illicit acts, recent academic debates bring new questions about these assumptions (Doshi & Ranganathan, 2019; Grossi & Pianezzi, 2018; Jancsics, 2019; Lopez, 2017; Relly et al., 2020).

Even though there has been less enthusiasm for the role of information in curbing corruption in recent years, the predominant analytical approach for corruption studies still considers transparency and accountability as critical variables.

According to Vanucci (2015), the most prevalent analytical paradigm to analyse corruption is the “economic paradigm”, to which the principal-agent model is a central pillar. The principal-agent model is commonly used to demonstrate the relationship between actors situated in two fundamental positions: the one who delegates authority, which is the “principal”, and the one other to whom the power is delegated, the “agent”. In democracies, the principals are citizens and voters, while the agents are the elected officials. The principal-agent approach was originally used to study the relationship between the owner of an enterprise (principal) and its managers (agent) when the former cedes control of the company to the latter. The critical point here is to induce the agents' behaviors towards benefitting the principal (the entrepreneur), given that the agents (the managers) have their own interests. In contrast, the principal has little ability to monitor all the agent's actions. In this scenario, information asymmetry is key.

In the principal-agent approach for corruption studies, transparency and accountability are critical elements because the availability and use of government data could reduce information asymmetry between citizens and politicians (respectively principals and agents) thus minimizing the costs of monitoring and sanctioning.

The principal-agent model, however, has been criticized, for instance, by Persson et al. (2013). For the authors, “at least part of an explanation to why anticorruption reforms in countries ridden with rampant corruption have largely failed is that they are based on a

theoretical mischaracterization of the problem of systemic corruption” (2013, p. 450). The authors argue that the existence of “principled-principles”, that is, actors willing to hold corrupt officials accountable, cannot be assumed: “(...) as corruption is the expected behavior in a particular society, we should expect the key instruments to curb corruption in line with the principal–agent anticorruption framework — that is, monitoring devices and punishment regimes — to be largely ineffective since there will simply be no actors that have an incentive to enforce them” (2013, p. 450).

According to Vannucci (2015), there are three main paradigms for analyzing corruption:

- The economic paradigm, which, for being the most prevalent in the academic literature, tends to inform most of the anticorruption interventions in practice.
- The cultural paradigm.
- The neo-institutional approach.

Each paradigm focuses on specific variables and elements. While the economic paradigm is based on the principal-agent model, the cultural paradigm looks at social norms, traditions, and moral values. Lastly, the neo-institutional approach focuses not only on economic incentives or moral values but “also considers mechanisms that allow the internal regulation of social interactions within corrupt networks, and their effects on individuals' beliefs and preferences” (Vanucci, 2015, p. 18).

The economic approach sees corruption as a problem of incentives and opportunities, and the central pillar of this analytical framework is the principal-agent model. For the principal-agent model, information is vital. According to this approach, corruption cases occur mainly because of “hidden information” and “hidden action” problems; that is, the principals (citizens) do not entirely know what agents are doing or not doing.

As corruption is frequently depicted as a result of information asymmetry and the transparency agenda has been consistently promoted, it is unsurprising that many anticorruption interventions are focused on augmenting data availability.

After reviewing the main features of more than 30 international anticorruption initiatives targeting administrative corruption, Villeneuve et al. (2020) confirmed that most of them follow the economic paradigm, focusing on reducing the opportunities and incentives for corruption. Moreover, information-based tools are the most prevalent as transparency is seen as having the potential to reduce monitoring costs while increasing the risks for corruption actors, minimizing the opportunities for corruption.

Even though there have been numerous anticorruption initiatives based on transparency worldwide and, more recently, a buzz around tech-driven initiatives, such as open data, the

literature on open data and corruption is scarce. A systematic literature review, which will be detailed in this thesis, found few articles on the uses and the role of open data to counter corruption. Still, one of the emerging pieces of evidence is that some key datasets to counter corruption, such as land property and company registers, are not often publicly available.

To fill this gap, this thesis will analyze if and how the “principled-principals” — i.e., accountability actors — use a specific dataset seen as key for countering corruption: urban land and real estate property. The “publicity and accountability conditions” framework combined with elements of the historical institutionalism approach is the theoretical lens for analyzing the use of data and its role on corruption control. The central hypothesis is the following: due to the expectation of impunity, historically given, added to the lack of skills and resources to interpret data, accountability actors that could challenge local-level corrupt practices tend not to use the available information, even when a key dataset, such as real estate and urban land ownership, is openly and freely available to be re-used. And this lack of action is present even in large global cities, not only in smaller municipalities.

By responding to these questions, this case study will serve a broader academic purpose, offering insights and more sophisticated frameworks for future analysis on data use for anticorruption purposes, particularly at the subnational level in the Global South.

1.2 The Case of São Paulo, Brazil

On the International Anticorruption Day of 2015, the City Government of São Paulo announced the disclosure of its dataset on urban land taxation as an anticorruption intervention. Two years before, in 2013, São Paulo citizens had witnessed the so-called “Audit Tax Mafia” scandal hitting the headlines. According to multiple reports of that time, municipal tax auditors would take bribes from construction companies and real estate companies to allow these enterprises to illegally obtain municipal licenses.

In March 2016, three months after that ceremony marking the anticorruption day, the dataset on urban land taxation – IPTU on its acronym in Portuguese (Imposto Predial Territorial e Urbano) – was made publicly available in open data format. The IPTU dataset consists of more than 3 million registers, offering comprehensive data on each real estate property in São Paulo and including the names of each property's owners, whether legal or natural person. This information, critical to corruption control, was publicly available until the first quarter of 2022, when the local government understood that, given a newly sanctioned privacy legislation, the names of legal and natural persons should be omitted from the publicly available version of the dataset.

The Table 1 presents the main facts around this case in the period 2016-2022.

Table 1 - Historical summary of the case

Year	Relevant political and administrative events related to open data and anticorruption in Sao Paulo and Brazil
2012	Brazil's national access to information (ATI) law came into force.
2013	<ul style="list-style-type: none"> - Fernando Haddad, elected Mayor in Oct 2012, takes office in January: the creation of the Office of the Comptroller General (CGM-SP) with departmental status – the highest level in the local government organigram – and implementation of the national ATI law at the city level. - “Tax Auditors Mafia” corruption scandal
2014	Creation of Sao Paulo's Open Government Interdepartmental Committee and contacts with international multistakeholder platform Open Government Partnership
2015	<ul style="list-style-type: none"> - Huge anticorruption scandal “Car Wash” in Brazil breaks out. High-level politicians, top executives at state-owned companies, and influential businesspeople go to jail for bribery, corruption, and money laundering — tremendous popular support, anticorruption high on the agenda. - Petrobras (giant oil state-owned enterprise) former director is the 1st convicted in the scope of “Car Wash”. He is found guilty of money laundering via real estate – he purchased a luxury penthouse in Ipanema, Rio de Janeiro, using an opaque Uruguay-based legal person - In Sao Paulo, mayor Haddad announces package of anticorruption initiatives, including full transparency for the IPTU (real estate and urban land property) dataset with the disclosure of names of owners
2016	<ul style="list-style-type: none"> - Sao Paulo City Hall makes IPTU dataset publicly available in an open and machine-readable format. - Sao Paulo officially joins OGP – Open Government Partnership. - Mayor Fernando Haddad runs for re-election and is defeated by Joao Doria
2017	- New government as Doria takes office as mayor; CGM loses departmental status in the city government hierarchy and becomes an area of the Justice Department. Comptroller no longer reports to the mayor.
2018	<ul style="list-style-type: none"> - Brazil's privacy data law comes into force - Fernando Haddad, former Sao Paulo mayor, runs for Brazil's presidency and is defeated by Jair Bolsonaro - Joao Doria resigns from the São Paulo City government and runs for Sao Paulo state governor. He wins. - Bruno Covas, who was elected vice-mayor with Doria at the 2016 elections, becomes the new Mayor. CGM-SP regains departmental status
2019	<ul style="list-style-type: none"> - Media reports on apparent abuses of the Judicial System in the scope of the “Car Wash”. Popular support starts draining - Bolsonaro's government nominates an “anti-Car Wash” prosecutor general
2020	- COVID pandemic. Anticorruption disappears from the public agenda. President Bolsonaro announces that “Car Wash” is over because, in his words, “corruption is over in Brazil”.
2021	- Sao Paulo's access to information interdepartmental committee rules that public disclosure of names of real estate owners violates privacy data law
2022	- Column with names of real estate and urban land owners is deleted from the publicly available version of the IPTU database

As noted above, the period between 2016 and 2022 saw the emergence of privacy legislation and debates in Brazil while popular support and attention toward the anticorruption agenda decreased. In political terms, the six-year period also saw electoral defeats of a previously solid political force: PT, the Workers Party. PT had won the four previous presidential elections (2002, 2006, 2010, and 2014) and the previous municipal election (2012) in São Paulo. However, this political party lost the 2016 and 2020 elections in São Paulo City and the presidential election in 2018.

At the city-level accountability institutions, it is noteworthy that CGM-SP, the Municipal Office of the Comptroller General of São Paulo, was created in 2013, downgraded in 2017, and later, in 2018, it regained its departmental status (i.e., the comptroller reporting directly to the mayor) within the City Government but with another configuration. That is, in these six years, three different institutional models.

Another aspect related to horizontal accountability institutions is that São Paulo City has its own Court of Accounts (TCM-SP), a specialized agency in charge of examining the public accounts of the municipality. TCM-SP is administratively autonomous but is an agency that assists the local government's Legislative Branch. There are only two municipalities in Brazil with their own local Court of Accounts: São Paulo and Rio de Janeiro.

Privacy Legislation

The establishment of privacy protection law affected the transparency of the real estate ownership dataset in São Paulo. The full availability of data on real estate property, transparently informing who owns each property, was in place in São Paulo for six years: from 2016 to 2022. This represented a rich opportunity to analyze the articulation between transparency, accountability, and corruption control. There have been various reports on corruption and money laundering in the urban construction and real estate sector in different parts of the world. In São Paulo, there has been the above-mentioned “Audit Tax Mafia”; also in Brazil, but in Rio de Janeiro, a penthouse in front of the world-famous Ipanema beach purchased by a former government top-official via an opaque legal structure was the proof of the first court conviction within the Brazil’s anticorruption investigation known as “Car Wash” (Brandt et al., 2015).

And there have been reports on corruption and money laundering in the real estate sector from other parts of the world, far beyond Brazil or Latin America. For instance, Hong Kong, where the elder daughter of the Chinese Communist Party’s No. 3 leader allegedly also used

opaque structures to buy luxury real estate, a New York Times investigation shows (Stevenson & Forsythe, 2020). Other reports relate corruption to the real estate sector in locations as diverse as Honduras (Chayes, 2017), Australia, Canada, the United Kingdom and the United States (Martini, 2017).

A European Parliament Research Service (EPRS) report published in 2019 stresses that the share of real estate in criminal assets confiscated, which can be used as an indicator of how much money is laundered through real estate, is estimated at 30% between 2011 and 2013 (Remeur, 2019). Furthermore, the report qualifies the social-economic effects as “enormous”, highlighting that one of the consequences is the distortion of real estate prices, reducing housing affordability (2019, p. 10).

Financialization of the real estate sector around the world, but particularly in megacities and financial centers, has caught attention from scholars and practitioners beyond the governance field: professionals from domains as diverse as urban planning (Farha & UN Human Right Council, 2017) and taxation (Trautvetter, 2020) have alerted on the risks and consequences of high levels of financial flows and low levels of regulation and transparency of this industry.

Using non-transparent companies and trusts or third parties acting as legal property owners to hide the actual owner is not the only option available for criminals. Other techniques are paying cash or via opaque financing schemes.

Additionally, as urban land areas tend to be more lucrative and valued than rural areas, the rapid expansion of cities also leads to corruption cases related to undue influence and bribery payments. The goal of bribe payers is to change regulations to artificially transform rural areas into urban ones, even before specific areas are covered by construction (Borges & Silva, 2019). In that sense, not only real estate ownership data is useful to counter corruption, but also data on ownership of urban land itself.

On the other hand, the literature points out that subnational governments' good governance instruments are scarcer than national or federal governments (Dincer & Johnston, 2020; Erlingsson et al., 2008; Meza & Pérez-Chiqués, 2020; Strach et al., 2019).

Furthermore, cities have grown extensively worldwide, particularly in the Global South, transforming rural countries into urban ones. That makes it urgent for scholars and practitioners to research governance and corruption at the local level.

Given this context, the research question of this thesis is the following:

Have accountability actors been using open data to prevent and repress corruption and money laundering cases involving urban land and real estate in megacities of the Global South? If so, what is the effect of such data use on corruption control?

This broad question is built upon the case of São Paulo’s database on real estate property.

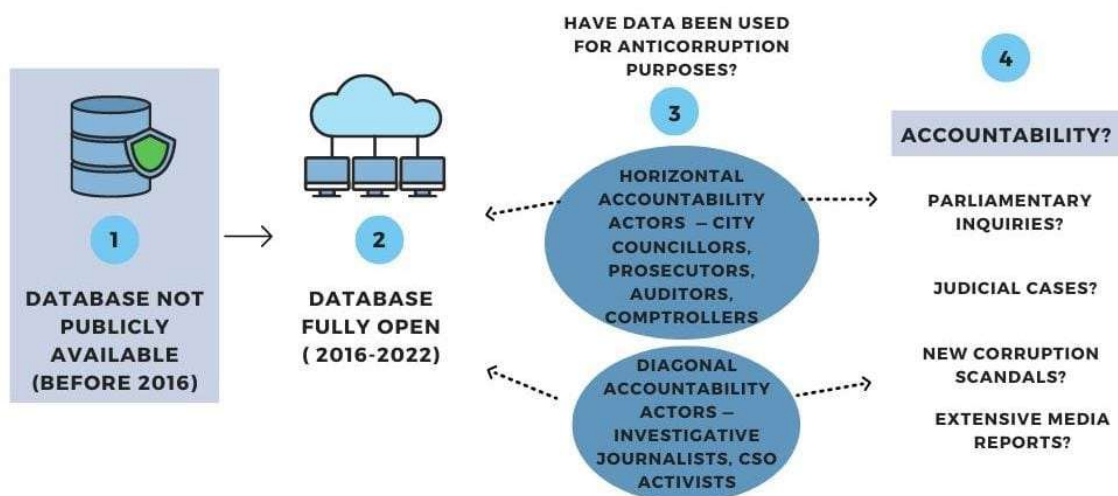
1.3 Diagonal and Horizontal Accountability

Guillermo O’Donnel (1998) introduced the concepts of horizontal and vertical accountability – the former refers to public sector institutions in charge of holding other government agencies into account while the latter refers to the ability of citizens, mainly via vote. Based on O’Donnel, recent studies have suggested that actors in society can activate horizontal accountability. These actors are mainly investigative journalists and civil society organizations, who perform “diagonal” accountability.

An underlying research question is the following: what are the main barriers for data use, when it comes to countering corruption?

The Figure 1 represents the ideal outcome for open data in the case study. It is noteworthy that the role of accountability actors is critical in the process of using data for anticorruption purposes.

Figure 1 – The expected outcome of the disclosure of data in São Paulo



Source: Elaborated by the author of the thesis

The literature review, presented in more detail in the following section, indicates that barely any literature addresses the relationship between open data and corruption. The studies

published on open data and corruption are limited to discussing the processes around transparency and open data policies, mentioning the control of corruption as potentialities of data. Only in recent years have a few empirical studies sought to understand and explain in greater depth the limits of open data in tackling corruption, pointing out accountability failures as one of the central issues.

2 THEORETICAL-EMPIRICAL FOUNDATION

This chapter discusses the three central elements of this thesis: corruption, open data, and accountability.

It starts with a presentation of the current academic debate on corruption conceptualization, presenting scholars' concerns around improving the concept to design policy interventions better. This session continues with a systematic literature review which presents case studies where historical institutionalism has been used to discuss the phenomenon of corruption.

This review helps understand the theories and debates around corruption in recent years for better situating the topic of this thesis – open data as an anticorruption intervention at the subnational level – into this overall discussion.

The next item of this chapter discusses the relationship between transparency and accountability and closes by explaining the concept of accountability actors used in this thesis. It is relevant for the objective of this thesis to frame the discussion on transparency, open data, and accountability and to present the concepts of horizontal accountability actors and diagonal accountability actors. Both categories of actors are key for anticorruption purposes once it lies on them the responsibility for challenging illicit acts and actors.

A brief discussion of open data and development follows to discuss the challenges for data to fulfill the pledges around development goals, particularly regarding data use and its outcomes. In the next item, I discuss open data and corruption, with a systematic literature review showcasing the limited number of studies covering these two elements altogether, specifically in terms of the little knowledge production around the role of open data to countering corruption.

The section closes with a discussion connecting open data, accountability, and corruption while presenting the two theoretical lenses used in the thesis: the “publicity and accountability condition” and historical institutionalism.

2.1 Corruption: a Literature Review

Since the 1990s, when corruption reached the public debate, the most cited definition of corruption is the one that has been coined by the not-for-profit organization Transparency International: the abuse of entrusted power for private gain. However, this definition has been challenged recently, as there has been intensive academic debate about what corruption is and

the conceptual basis that informs the design of anticorruption interventions. This section will present a literature review to demonstrate this current debate, including discussions on anticorruption interventions.

The first step to this literature review was an exploratory reading based on the “Anticorruption Bibliography”, a set of publications that Harvard Law School professor Matthew Stephenson has collected since 2014. In 2021, the bibliography gathered by the researcher became the repository “The (Anti-)Corruption Corpus”, published by the think tank Global Integrity through its Anti-Corruption Evidence Research Programme.

Three analytical approaches emerged from this first exploratory reading, as seen below.

Table 2 – Exploratory literature review on corruption

Analytical approach	Definition	Authors and articles	Main keywords/constructs
Corruption as a social practice	“It is essential to think about corruption in a systemic dimension that combines political morality - presupposed and that establishes the meanings of corruption - with social practice itself, in the dimension of everyday life.” (Filgueiras, 2009, translated from the Portuguese by the author of this thesis)	Torsello & Venard (2016), Filgueiras, (2009), Gonçalves & Andrade (2019), Barr & Serra (2010).	<ul style="list-style-type: none"> • Bribery • Deviant/counterproductive behavior • Ethics – Morality • Social practices • Social norms • Culture
Principle-agent model	“In a democratic society, people are the “principal,” and the government is their “agent.” The agent is supposed to serve the people and work toward achieving their objectives. In reality, however, the agent has its own goals and controls information regarding public processes and resources. This informational asymmetry allows the government to act in its self-interest, possibly to the detriment of people’s welfare.” (Chuah et al., 2020).	Rose-Ackerman, (2013), Klitgaard (1991), Granickas (2014), Lambsdorff (2003), Chuah et al. (2020)	<ul style="list-style-type: none"> • Political economy • Agency relationships • Systems of information • <i>Accountability</i> • Good governance • Transparency • Open Government Data
“Power capture” model	“Although in the short-term corruption may be able to ‘grease the wheels of the economy’, in the long term it negatively affects growth by diverting resources from more productive uses and negatively affects equity by disproportionately benefiting	Doshi & Ranganathan (2019), Mungiu-Pippidi (2013), Cook & Emerson (1978), World Bank (2017),	<ul style="list-style-type: none"> • State capture • Grand corruption • Political corruption • Resources transfer • Predatory behavior • Legal constraints • Normative constraints • Economic Exchange theory

	<p>those in power” (World Bank, 2017) “We thus posit that corruption is not a fixed set of practices, but rather an interpretive rubric that serves to make sense of and distinguish what is ethical or not, what is harmful or not, and what matters or not (and to whom) in the ordinary processes of wealth accumulation and dispossession that define capitalism as we know it” (Doshi & Ranganathan, 2019)</p>	<p>Kauffman & Vicente (2011), McGee (2019), Jancsics (2019), Hellman et al. (2003)</p>	<ul style="list-style-type: none"> • Social Exchange theory • Power and equity • Exchange networks • Collusion
--	--	--	--

Authors who see corruption as a result of social practice focus on moral values and cultural traditions, which would be the key elements shaping individual actions towards corruption. For example, using Durkheimian concepts, Gonçalves and Andrade, in a paper on the corruption scandal labeled “Car Wash” in Brazil, affirm that corruption presents itself as “a pathological social fact according to the Durkheimian conception, through the analysis of four variables: anomie; pathological social fact; social institutions; and individuals”¹¹ (2019, p. 287). According to the authors, anomie is a condition in which “social and moral norms are confused, unclear, or absent”. They add that, in this scenario, “individuals do not adhere to social and moral norms. It consists of a breeding ground for acts of corruption, due to the clear lack of boundaries between the public and the private”.

Similarly advocating in favor of the centrality of society and culture in explaining corrupt behaviors, Barr and Serra (2010) conducted a bribery experiment and concluded that individuals’ propensities to act corruptly might reflect the cultures in which they grew up. In a critical review of the anthropological literature on corruption, Torsello and Venard (2016) conclude that anthropology deliberately avoids a universal definition of corruption — the field prefers ad-hoc explanations using the viewpoint of the observed. The authors argue that although laws specify a sphere of legitimacy, stating what is legal and illegal, the political elite creates these formal rules. Therefore, such rules are not entirely neutral, and, especially in jurisdictions where society is disconnected from the political processes, they may be seen as unfair and illegitimate.

¹¹ Translated (from the original Portuguese) by the author.

The main argument is that because moral values are dependent on the culture and because cultures are plural, moral values are diverse, which makes moral judgments culture-dependent. Instead of being destructive, corruption could be morally acceptable and even socially cohesive in some societies. (Torsello & Venard 2016, p. 48)

Torsello and Venard argue that “people in various cultures will react differently to the prescriptive norms imposed” (2016, p. 50), adding that organizational scholars interested in understanding corruption should research the phenomenon from a local point of view.

Regarding anticorruption intervention, the authors suggest that “[i]n any change program, the necessary changes must be analyzed from the users’ perspective. This is also the case for corruption. As a consequence, a cultural approach could help to design an anticorruption policy” (2016, p. 50).

The cultural approach, however, is not the most prevalent in research about corruption. The most widespread of the three approaches is the one that sees corruption as a result of information asymmetry, using the principle-agent model. Here, the relationship between those exercising power and the governed is crucial. The overall idea is that the ones in power are the agents who presumably will act in favor of the principals’ best interests. The principles here are the people (voters, society) whom those agents (politics and high-level bureaucrats) represent. However, the interests of the agents might not always be coincident with the interests of the principles. Moreover, in public life, James Madison stated in 1788: “If angels were to govern men, neither external nor internal controls on government would be necessary” (Madison, 1961, p. 322).

According to this literature, corruption can be explained by the lack of accountability and the excessive discretion by those who exercise power. The corruption equation, formulated by Robert Klitgaard in the 1990s – one of the founding authors of this paradigm –, is widely known: “Corruption equals monopoly plus discretion minus accountability”. Hence, lowering discretion and augmenting participation and accountability could counter corruption.

Interestingly, literature and practice that emerged in the 1990s, prescribed more openness and transparency, paying less attention to participation and accountability. Indeed, the overall expectation was that more availability of information would naturally and automatically strengthen both civic participation and accountability.

The European Public Sector Information Platform Topic Report 2014/04, entitled “Open Data as a Tool to Fight Corruption,” affirms that “it has been universally accepted that

release and re-use of public sector data improve good governance, effectiveness and citizen participation” (Granickas, 2014, p. 6). The author cites the World Bank, to which open government data is a “catalyst of multiple public policy and development achievements”. In addition, the report argues that “by shedding light on previously hidden aspects of governmental activities, decisions, and expenditures, open government data is expected to strengthen the ability of civil society to monitor the performance of governments” (2014, p. 6).

A more recent World Bank publication, the “World Development Report 2021: Data for Better Lives”, reassures that data can support development in three ways: greater accountability, better policy making and service delivery and increased business opportunities (2021, p. 4).

Pushed by an international call for more open governments, the number of countries with access to information legislation boomed. For example, in 1990, only 13 countries had national legislation on access to public information, whereas in 2014, the world reached a symbolic quantity of 100 countries with access to public information laws. Another symptom of the effectiveness of this call for more transparency in governments worldwide was the creation in 2011 of the Open Government Partnership, a multistakeholder platform proposed by the Obama administration. Actually, on his first day in office as president of the United States, in January 2009, Barack Obama committed to an “unprecedented level of openness in Government” in a “Memorandum on Transparency and Open Government” (Obama, 2009). This memorandum led to the Open Government Directive, published by the US Government in December 2009.

The overall proposition of shedding light on previously hidden actions and aspects of governmental activities, decisions, and expenditure is vital for the principle-agent model explanation of corruption. With more information available, the principles would be able to monitor and control the agents – hence the costs for malfeasances would be higher. Additionally, more civic participation would reduce the opportunities for the self-interest decision-making process and possibly incentivize the creation or the strengthening of formal accountability procedures. According to rational choices, the “abuse of entrusted power for private gain” (Transparency International) could not be the best option in an environment of little information asymmetry, as governance mechanisms would reduce the opportunities for private gains.

However, a third corpus of studies has emerged recently and focuses neither on cultural norms nor formal rules. Instead, these authors look at corruption using the lenses of power and power relations, reflecting on both social practices and organizational controls.

The three analytical approaches mentioned above mirror the three paradigms Vannucci (2015) presented in his corruption conceptualization paper. The economic paradigm, using the principal-agent model, and the cultural paradigm, which focuses on social norms, identically reflect the analytical approaches described here as “corruption as a result of information asymmetry” and “corruption as a result of social norms”.

The third paradigm, labelled above as “power capture model”, is described by Vannucci as the neo-institutionalism approach. In the article, the author suggests that after developing “certain organizational texture and cultural adaptation,” corruption activities count on “informal codes and governance structures” – that is, both social norms and formal rules. According to the author, such codes and structures “provide internal stability and enforcement mechanisms to illegal dealings in specific areas of public activity”. Moreover, such stability and enforcement mechanisms reduce uncertainty among partners “which thus appear more lucrative and less morally censurable”.

Discussing the roles of both social norms and formal rules, Vannucci argue that the “co-evolution of incentives and cultural values” is path dependent. In that sense, “the heritage of corruption in the past produces increasing returns in subsequent periods by providing informal norms, learning of specialized skills and other mechanisms of protection against external intrusion by authorities and internal friction among corrupt actors” (2015, p. 19). Hence, path dependence appears here as a possible root explanation for the persistence of corruption.

Based on the understanding that corruption results from both social norms and formal rules, Vannucci calls for more sophisticated and combined anticorruption strategies. “Only when official rules are complemented by coherent informal institutions, bottom-up initiatives, they tend to produce the expected outcomes and make anticorruption regulation more effective” (Vannucci, 2015).

Jancsics (2019) is another author who has recently reviewed corruption studies. The author argues that it is essential to conceptualize corruption better and define specific types of this phenomenon to tailor the anticorruption strategy.

Based on De Graaf’s previous study (2007) on the causes of corruption, in which the author grouped six theories, Jancsics added a seventh theory: collective action.

The seven theories are summarized in the Table 3:

Table 3 – Theories of corruption

Theory	Main cause/actor	Main feature/driver
Public-choice	Individual	Maximization of private profit
“Bad apple”	Individual	Lack of moral values of individuals influencing other individuals’ behaviors
Organizational culture	Meso-level	Group behavior in a structure influencing individual behavior
Clashing moral values	Societal level	Values and norms of a society influencing individuals
Ethos of public administration	Macro-level	Large-scale public sector reforms undermining the traditional ethos of public administration such as integrity, merit, accountability, responsibility, and longevity of service
Correlation	Macro-level	Quantitative and comparative analysis – income, democracy, voter turnout – influencing or correlating with corruption
Collective action	Individual	Individual perception that all agents are likely to be corrupt

Source: prepared by the author of this thesis combining Jancsics (2019) and De Graaff (2007)

Jancsics notes that the recent collective action theory relates to the public-choice view because it is a rational individual choice. However, collective action theory goes beyond the rational choice approach once it also attempts to explain the absence of action in countering corruption. According to Jancsics, collective action theory argues that

corruption persists because the public agent perceives that all other agents are likely to be corrupt. In a society where corruption is the expected behavior, there will be no actors with incentives to enforce punishment regimes. Therefore no one will be held accountable for the corrupt act. (2019, p. 3)

The lack of “actors with incentives to enforce punishment regimes” is a central element for this thesis as it explores the role of accountability actors.

In an article widely cited², written based on an interview study conducted in two African countries (Kenya and Uganda), Persson et al. (2013, p. 450) argue that “at least part of an explanation to why anticorruption reforms in countries ridden with rampant corruption have largely failed is that they are based on *a theoretical mischaracterization of the problem of systemic corruption*”. The authors add that while most anticorruption interventions are designed considering the principal-agency framework, systemic corruption resembles a collective action problem.

² More than 1,000 citations according to Google Scholar (last seen on 20.December.2022)

Persson, Rothstein and Teorell stress that both Kenya and Uganda have conducted institutional and legal reforms to tackle corruption in the last few decades. They cite the think tank Global Integrity, which ranks the vast majority of Sub-Saharan African countries as having “very strong” anticorruption laws – with “large” or “huge” implementation gaps, though. This, they argue, is evidence that formal rules are not enough to deter corruption.

The main criticism from the authors towards the principal-agent model and the anticorruption interventions based on this framework is the presence of “principled principles” – actors who are ready to hold the powerful to account and are not corrupt themselves – cannot be assumed. Moreover, “to the extent that corruption is the expected behavior”, the benefits of corruption outweigh the costs, at least in the short term, Persson, Rothstein and Teorell sustain.

In another article, the same authors argue that corruption resembles “what is commonly referred to as an assurance game or a multiple-equilibria coordination problem” (Persson et al., 2018, p. 800). In a sense that what action will be taken by one individual largely depend on expectations on how others will act.

Collective action theory does not explain corruption in terms of corrupt actors rationally maximizing their gains. Instead, corruption occurs because all individuals behave in a way that creates room for more corruption: considering the totality of the actors in society, they will either undergo corrupt practices or avoid blowing the whistle. Action against corruption is seen both as costly and ineffective, according to testimonies collected in the African countries by Persson, Rothstein and Teorell. As there is little action against corruption, the risks and costs of corruption tend to be relatively low.

Therefore, one of the critical contributions of the collective action literature is the focus on accountability actors, particularly in explaining why accountability fails. In that sense, collective action theorists criticize the principle-agent approach, which informs most of the anticorruption policy interventions in recent years.

In a study that proposes a classification of anticorruption initiatives, Villeneuve et al. (2020) analyzed more than 30 policy interventions, adding to the understanding (Johnsøn, 2012; Marquette & Peiffer, 2015; Persson et al., 2013) that anticorruption interventions are largely influenced by the economic paradigm. Moreover, the authors found that the information-based implementation tool is the most frequent type of policy tool among anticorruption interventions.

As noted, the most influential approach informing the design of anticorruption interventions assumes that more information will be a central part of the strategy. Furthermore, that expectation brings another assumption: the information-based tools consider that agents – whether in government or in society – will use the information to hold the powerful to account.

However, after the boom of access to information laws and open data portals, noted between the last decade of the 20th century and the first 20 years of the 21st century, there is a relatively broad consensus that laws and tools have not been able to counter corruption. The challenge now is to try to explain why the strategy failed.

In contexts where widespread corruption is in place and has been in place for a long time, information without accountability might be insufficient (Peiffer & Alvarez, 2016; Schuster et al., 2020). The socially spread motto of “everyone has always done that way” as well the capacity of the powerful to largely influence the creation of formal rules favouring impunity might be a better explanation for the persistence of the phenomenon, particularly in contexts such as local governments in the Global South, where corruption seems to be “consolidated” (Meza & Pérez-Chiqués, 2021).

In that sense, the notion that not only formal rules but also informal and social stances might shape corrupt practices make the neo-institutional approach a promising framework for analyzing the persistence of corruption and the seeming failure of anticorruption initiatives.

2.1.1 Systematic Literature Review – Corruption and Historical Institucionalism

Political outcomes, argue James G. March and Johan P. Olsen, are “a function of three primary factors: the distribution of preferences (interests) among political actors, the distribution of resources (powers), and the constraints imposed by the rules of the game (constitutions)” (1983, p. 739). None of these factors are entirely exogenous to institutions, the authors argue. These remarks are found in the influential article “The New Institutionalism: Organizational Factors in Political Life”.

In the paper, March and Olsen discuss the revival, at that moment (early 1980’s), of the scholars’ attention to institutions, after a period in which theories of politics had been more focused on portraying political life as a reflection of society, history, individual behaviors – that is, external, exogenous to institutions. The authors then pointed out that the “new institutionalisms” were correct in bringing institutions back in (Skocpol et al., 1985), mainly because “[e]mpirical observations seem to indicate that processes internal to political institutions, although possibly triggered by external events, affect the flow of history” (1983, p. 739).

In that sense, this thesis assumes that the neo-institutionalism is indeed an adequate approach to analyze corruption, understanding this phenomenon as a result of both processes,

rules and norms endogenous to institutions added to exogenous dynamics, including cultural and societal norms and practices.

The choice of neo-institutionalism presents the challenge of specifying what type of neo-institutionalism and what elements of such approach are the most appropriate.

In a seminal paper first published in 1996, Peter Hall and Rosemary Taylor describe three analytical approaches calling themselves “new institutionalism”: historical institutionalism, rational choice institutionalism, and sociological institutionalism. According to the authors, “(...) these approaches developed in reaction to the behavioral perspectives that were influential during the 1960s and 1970s and all seek to elucidate the role that institutions play in the determination of social and political outcomes” (1996, p. 5). Each analytical approach, however, “(...) paint quite different pictures of the political world” (1996, p. 5).

According to Hall and Taylor, sociological institutionalisms argue that formal rules and practices should be seen as culturally specific: all forms of bureaucracy can and should be explained in cultural terms. On the other hand, rational choice institutionalism is largely influenced by the notions of cost and benefit – property rights, rent-seeking, transaction costs are essential elements in the operation and development of institutions. The authors also mention the principal-agent model as influential in rational choice institutionalism.

Three analytical approaches to explaining corrupt practices mentioned earlier can be seen vis-à-vis the three new institutionalisms labelled and explained by Hall and Taylor. What emerges from this parallelism is that the studies using cultural lenses to explain corruption apparently utilize approaches related to sociological institutionalism. In contrast, the economic paradigm used by many scholars studying corruption seems to be associated with rational choice institutionalism.

The other form of new institutionalism, historical institutionalism, is the one that will be used in this thesis. Elements of historical institutionalism will be mobilized so as illuminate the perceptions of corruption persistence and verify if such perceptions impact accountability actors’ choices in terms of calculating the costs they bear for countering corruption.

A systematic literature review has been undertaken to form a baseline from which the current analysis starts. In February 2022, a search on three academic databases – Scopus, Web of Science, and EBSCO – was executed to identify studies using historical institutionalism to analyze corruption. The search, using “historical institutionalism” and “corruption” in the “title”, “abstract” and “keywords” fields, found only nine documents: one conference paper, whose author two years later launched a book on the same topic, and eight articles. One of

which mentions corruption just laterally; hence it was sidestepped. The book and the seven articles are analyzed below.

The least recent study found is a conference paper presented in 2003 at the American Political Science Association Annual Meeting. Two years later, the author, Rasma Karklins, launched a book entitled “The System Made Me Do It — Corruption in Post-Communist Societies”. In a chapter about political legacy, Karklins highlights a path-dependent aspect: the paths taken in post-communist countries in the early stages of the transition from communism affected the choices available later. The transition from communism is a typical moment that Pierson defines as a “critical juncture”: a situation in which the reproduction of a system is disrupted (Pierson, 1997, p. 9). According to Karklins, the elite who profited the most in the first stages of privatization became the primary opponents of further reform. “The informal networks of power elites and regular citizens that characterized Soviet-type regimes gave left an imprint on post-communist societies” (2005, p. 75). The author concludes that “established structures and procedures retain influent both through inertia and as a safe net in confusing times” (2005, p. 89).

The second least recent study was published in 2012. Its author, Algirdas Acus³, explains that the article's purpose is to answer the following question: “why is corruption so resistant despite a convincing progress in adopting good governance institutions worldwide?”. Using two case studies, Argentina and Chile, in a comparative perspective, the author combines game theory and historical institutionalism in his analysis to assess if corruption and its contrast – good governance – are both stable equilibria. The author argues that different paths led to varying types of modernization in Argentina and Chile. Modernization, stresses Acus, “is a kind of explosion that upsets the old order”, as the development of industry, cities, and education profoundly changed societies' economic, social and political structure. The author argues that in Chile, modernization was gradual and has covered heavy industry-intensive regions, favoring the development of both right-wing and left-wing parties. This has created the conditions for genuine party competition to take hold — all the most influential groups in society could compete, hence reducing the incentives for corruption. In Argentina, on the other hand, modernization was fast and included other groups. This type of modernization, according to Acus, prevented both the right and the left from creating their own parties. As a result, the most influential groups in Argentinian society remained unrepresented, which encouraged corruption.

³ I have to thank Algirdas Acus for generously sending me an English version of his article, originally published in Lithuanian.

The author argues he confirmed the hypothesis of “stable equilibria”, adding that both good governance and corruption “are maintained through the structure of competition – interaction among individuals”. Acus states that “institutional design or cultural-moral norms are overestimated” and “predominant explanations of corruption, namely the principal-agent model and the cultural-moral argument, do not properly address the question of individual motivation to act partially or impartially”; therefore, continues the author, “they lose the explanatory power”. Here again, neither institutions nor culture explain the persistence of corruption. Instead, historically inherited incentives and disincentives present in a specific context explain the phenomenon's persistence.

The other seven articles found during this systematic literature review are very recent, less than ten years old: they have all been published either in the second half of the last decade or in the early 2020s. This is evidence that historical institutionalism as the main lens to analyze corruption is a recent trend.

In a 2016 study, Kim Moloney and Hyo-Youn Chu used historical institutionalism to explain the apparent implausible coexistence of high public service motivation (PSM) and an unfavorable ethical climate in Jamaica. The authors found intriguing results in a survey with 157 public sector management and governance students, of which 66% were civil servants at the Government of Jamaica. Along with high levels of PSM, survey responses showed high levels of perception of civil servants (both junior staff and senior managers) and politicians abusing their power, acting unethically, and serving their own interests before serving the public interest. “Distrust interacts with weak monitoring institutions and an ‘informer culture’ to create an ethical climate unable to encourage PSM value enactment” (2016, p. 449), writes Moloney and Chu.

If the civil service (and the society in which it rests) does not uphold meritocracy and safe whistle-blowing or if the system punishes those who attempt to notify others of abuse or does not prosecute and convict those who commit crimes, then substantive PSM values are disassociated from their instrumental impact ... Perceptions of unprosecuted mismanagement, misbehavior, or corruption among civil servants and politicians may discourage intrinsically motivated respondents to be proactive in the presence of abuse. (Maloney & Chu, 2016, p. 449)

The authors argue that historical institutionalism contains calculus and cultural components.

By choosing not to report a misbehaving colleague, civil servants may employ a calculus approach as not reporting might preserve one's job given other institutional weaknesses. The cultural component interacts with the calculus component by suggesting 'good' Jamaicans will not inform. Each approach explains what at first appears contradictory: high PSM within a country known to suffer from corruption. Institutional weaknesses and an 'informer culture' may prohibit the active implementation of one's passive PSM values. (Maloney & Chu, 2016, p. 438)

In a 2017 study on Greece during the 1990s and 2000s, a period in which the country was facing fiscal and economic crises, Aris Trantidis and Vasiliki Tsagkroni used historical institutionalism lenses to analyze how political actors continued to extract private gains from the state even in a context of diminishing fiscal conditions. The authors found that the political actors adjusted in the forms of state capture. According to Trantidis and Tsagkroni, "extraction *through* the state and extraction of resources *from* the state are alternative ways by which political actors can accommodate the rent-extracting demands of their clients" (2017, p. 276, italics in the original). This hybrid form of state capture, which the authors label "client corruption", requires a legal framework allowing form considerable administrative discretion and a framework of impunity "designed to enable corrupt activities".

According to Trantidis and Tsagkroni, a historical institutionalist approach helped

unearth nuanced dynamics of state capture: how types and modes of selective benefits extraction are reconfigured in alignment with the dynamics of the economy and the terms of political competition, and how they relate to one another. The form of change may vary, depending on the adaptive responses of the key actors involved. (2017, p. 268)

Ina Kubbe authors a paper published in 2018 that investigates corruption's influencing factors in Europe from 1995–2013. Drawing on sociological and historical institutionalism and applying panel and multilevel analysis, the author found that "corruption exists, persists, and varies significantly across cultures in Europe" (2018, p. 233). The study, however, finds certain countries' contextual conditions that influence the extent of corruption over time and across European countries, such as the economic development, international integration (EU-membership), women's percentage in parliaments, Protestantism, the degree and duration of democracy as well as a post-communist past. Regarding historical factors, Kubbe notes

significant relationships between the durability of democratic systems and a country's communist past, and the extent of corruption. According to the author, the finding “implies that democratic structures not only decrease levels of corruption, but that this effect is also strengthened by the duration of democratic principles. In other words, the longer a democracy lasts, the less corrupt it is” and adds that the relationship between the duration of democracy and the extent of corruption is “even stronger than the relationship with a country's degree of democracy”. History shapes institutions more decisively than specific measures aimed at increasing the level of democracy. Interestingly, Nordic countries, which possess the longest-lasting transparency laws in the world – Sweden's legislation dates to the 18th century, for instance –, are frequently seen as some of the least corrupt in the world, according to Transparency International's Corruption Perception Index.

The findings of Kubbe's article also build on the arguments from the “corruption as a collective action problem” as the author explains.

It seems that people have greater expectations and a higher estimated probability that, for instance, a given public official will engage in corrupt acts in societies with high levels of corruption. These results also clearly demonstrate the cultural transmission of corruption, which implies that individuals from societies in which corrupt transactions are quite common are more likely to engage in corruption and expect others to engage in it as well. (Kubbe, 2018, p. 232)

Two papers on corruption using HI lenses were published in 2019: one investigating why Mexican firms and individuals do not use offshore finance to plan taxes, avoid regulations, or launder money, and one on transparency and campaign finance in Taiwan.

Author of “All exclusive: the politics of offshore finance in Mexico” (2019), Andrea Binder finds that Mexican corporations and wealthy individuals make limited use of offshore finance. According to the author, this occurs due to the historically given concentration of political and economic power, which leads to the onshore economy offering similar rents for economic elites as offshoring. According to the author,

the relationship between the Mexican government, its financiers and taxpayers shaped banking services and tax policies such, that Mexican corporations and wealthy individuals have limited incentives to go offshore. The handful of large, internationally active Mexican economic actors can use the mostly foreign-owned banking system to

access international financial markets without having to go offshore. Tax evaders and money launderers, on the other hand, can hide in the onshore informal economy. That is, the oligopolistic structure of the Mexican economy in combination with its large informal sector offers similar rents for capital as offshore financial services, while demanding less legal and financial sophistication. (2019, p. 315)

A 2019 paper by Po Liang Chen, on transparency and campaign finance legislation in Taiwan, adopted a historical institutionalism approach to analyze the evolution of electoral clientelism and its interaction with campaign finance law before the enactment of the legislation in four eras: the Japanese colonial rule era (1935-1945), the transition from the Japanese colonial rule era to the Republic of China era (1945-1949), the authoritarian era (1949-1991), and the democratic era (1991-2004).

Chen argues that “fair elections did not produce clean government in Taiwan” and concludes that throughout Taiwanese history, authoritarian governments “monopolized economics and financial sectors... [and] used its economic advantage to build up a system of electoral clientelism. This system allowed the government to manipulate local factions through the exchange of material interests or privileges in return for the mobilization of voters” (2019, p. 41).

A 2020 paper examined changes in Italian anticorruption agencies’ mandates and power over time. Drawing from insights on historical institutionalism, this article, by Di Mascio et al. (2020), undertakes a qualitative longitudinal analysis of organizational change regarding anticorruption agencies (ACAs). The authors argue that their investigation produced three main findings. Firstly,

the cyclical fluctuation displayed by the historical trajectory of agencies implies that delegation has been primarily affected by domestic factors... the extent of delegation corresponds to the balance struck by governments, between expected credibility gains as a principal and expected credibility losses as a target, in a context punctuated by domestic crises and scandals that acted as triggers for reforms. (Di Mascio et al., 2020, p. 388)

Secondly, ACAs can be reformed according to a process of layering, whereby organizational evolution will eventually alter the logic of the regulatory framework. This process is determined by the presence of many veto players and low discretion in the

implementation of the regulatory framework. Third, that political leaders are indeed particularly proactive about upholding the formal independence of regulators, which relates to the symbolic properties of delegation, while agency leaders prefer to enjoy more de facto independence. Citing Mahoney (2000), the authors write that evolution can be understood “as a ‘reactive sequence’— that is, a chain of events linked through reactions and counter-reactions” (2020, p. 369).

The most recent article using historical institutionalism lenses is a 2021 paper on crime and security in the Mexican state of Michoacán. Authors Aguirre-Ochoa and Gómez (2021) argue that the historical change in Mexico after the year 2000, with the end of decades of one political party (PRI) ruling the country, was not enough to improve policies against crime and violence in Michoacán. The article, stress the authors, represents a continuum of a 15-year research project, and during the period, primary data from municipal and state-level civil servants as well as journalists and representatives of society at large have been collected. According to Aguirre-Ochoa and Gómez, the Mexican presidents elected after the PRI-era, even when demonstrating willingness to confront the situation of violence, “had to adjust their policies to preexisting interests and power groups, limiting the capacity of real change demanded by society” (2021, p. 47). In the period, criminal groups “forged complex networks and alliances with political and economic groups throughout the country, which the new governments of the post-PRI stage were unable to break” (2021, p. 47).

In the next subsection I explore features that this thesis will use based on the literature review about historical institutionalism and corruption.

HI's distinctive features

Hall and Taylor argue that four features of historical institutionalism are distinctive from the other forms on new institutionalism. Firstly, historical institutionalism sees the relationship between institutions and individual behaviors in broader terms —individuals may use the “calculus approach” (“institutions affect behavior primarily by providing actors with greater or lesser degrees of certainty about the present and future behavior of other actors”), the “cultural approach” (as institutions provide “moral or cognitive templates”, the choice of a course of action depends on the interpretation of a situation rather than on purely instrumental calculation) or both. Second, they emphasize the asymmetries of power and associate it with the operation and development of institutions. Third, they emphasize path dependence and

unintended consequences in the development of the institutions. Lastly, they integrate institutional analysis with the contribution of other factors, such as ideas.

As seen in the Table 4, power asymmetry and path dependence are the most frequent elements found in the nine articles analyzed.

Table 4 – Systematic literature review on corruption and HI – summary

Article	Results	Main element of the Historical Institutionalism approach (Hall and Taylor, 1996)
Acus (2012)	“It shows that general expectations are driven by the first game of Deer Hunt that was once played: modernisation. Using a process tracing approach, the modernisations of Argentina and Chile are presented, and it is shown how the former has remained in a deadlock of corruption while the latter has entered an equilibrium of good governance. Different paths led to different modernisations.”	Path dependence and unintended consequences in the development of the institutions
Karklins (2005)	“The informal networks of power elites and regular citizens that characterized Soviet-type regimes gave left an imprint on post-communist societies” (2005, p. 75). The author concludes that “established structures and procedures retain influent both through inertia and as a safe net in confusing times” (2005, p. 89).	Path dependence and unintended consequences in the development of the institutions
Moloney & Chu (2016)	“By choosing to not report a misbehaving colleague, civil servants may employ a calculus approach as not reporting might preserve one’s job given other institutional weaknesses. The cultural component interacts with the calculus component by suggesting ‘good’ Jamaicans will not inform. Each approach explains what at first appears contradictory: high PSM within a country known to suffer from corruption. Institutional weaknesses and an ‘informer culture’ may prohibit the active implementation of one’s passive PSM values.”	Relationship between institutions and individual behaviors seen in broader terms (“calculus approach” and “cultural approach”)
Trantidis & Tsagrioni (2017)	“how types and modes of selective benefits extraction are reconfigured in alignment with the dynamics of the economy and the terms of political competition, and how they relate to one another. The form of change may vary, depending on the adaptive responses of the key actors involved.”	Asymmetries of power associated with the operation and development of institutions Relationship between institutions and individual behaviors

Kubbe (2018)	“(…) democratic structures not only decrease levels of corruption, but that this effect is also strengthened by the duration of democratic principles. In other words, the longer a democracy lasts, the less corrupt it is”	Path dependence and unintended consequences in the development of the institutions
Binder (2019)	“the relationship between the Mexican government, its financiers and taxpayers shaped banking services and tax policies such, that Mexican corporations and wealthy individuals have limited incentives to go offshore...the oligopolistic structure of the Mexican economy in combination with its large informal sector offers similar rents for capital as offshore financial services, while demanding less legal and financial sophistication”.	Asymmetries of power associated with the operation and development of institutions
Chen (2019)	“throughout Taiwanese history, authoritarian governments “monopolized economics and financial sectors... [and] used its economic advantage to build up a system of electoral clientelism. This system allowed the government to manipulate local factions through the exchange of material interests or privileges in return for the mobilization of voters”	Asymmetries of power associated with the operation and development of institutions Path dependence
Di Mascio et al. (2020)	“(…) the cyclical fluctuation displayed by the historical trajectory of agencies implies that delegation has been primarily affected by domestic factors... the extent of delegation corresponds to the balance struck by governments, between expected credibility gains as a principal and expected credibility losses as a target, in a context punctuated by domestic crises and scandals that acted as triggers for reforms”.	Relationship between institutions and individual behaviors seen in broader terms (“calculus approach” and “cultural approach”)
Aguirre-Ochoa and Gómez (2021)	“the Mexican presidents elected after the PRI-era, even when demonstrating willingness to confront the situation of violence, “had to adjust their policies to preexisting interests and power groups, limiting the capacity of real change demanded by society” (2021, p. 47). In the period, criminal groups “forged complex networks and alliances with political and economic groups throughout the country, which the new governments of the post-PRI stage were unable to break”	Asymmetries of power associated with the operation and development of institutions

Powerful actors seem to be better positioned to influence and shape policies and institutions across time, while less powerful actors are unable to promote desired change given the strength of the corrupt networks and alliances. This could explain the persistence of

corruption even when important changes occur, and windows of opportunity are open (“critical junctures”). Critical junctures occurred multiple times in different contexts, such as the aftermath of a severe fiscal crisis (Greece), the inauguration of new governments after decades of one-party leading the country (Mexico), or after different countries colonizing a region (Taiwan).

Path dependence may explain both good governance — Chile (Acus, 2012) and some countries in Europe (Kubbe, 2018) — and corruption, as in the case of Argentina (Acus, 2012) and post-Communist countries in Europe (Karklins, 2005)

For these reasons, the elements of power asymmetry and path dependence derived from historical institutionalism will be used as lenses for the thesis.

Corruption consolidation framework in local governments

Elements from the “corruption consolidation framework in local governments”, proposed by Meza and Pérez-Chiques (2021), reinforces the relevance of using historical institutionalism as an analytical framework.

The authors suggest a four-dimensional framework to study how corruption is consolidated in local governments. In line with the arguments raised by Bo Rothstein and other “corruption as a collective action problem” theorists, Meza and Pérez-Chiques argue that scholars have been long neglecting a better conceptualization of corruption and these conceptual and theoretical flaws have led to the creation of anticorruption policies that tend to be overly reliant on formal institutions and have proven ineffective. Based on mixed methods, including dozens of in-depth interviews in two Mexican cities, their grounded constructed framework focuses on how networks may be shaped and organized to perform corrupt practices, and how opacity and weak checks and balances may grant impunity. In line with the results shown in the papers selected for the literature review on historical institutionalism and corruption, Meza and Pérez-Chiques’s results show that corrupt practices can be malleable and resilient, with the ability to circumvent formal anticorruption strategies.

These are the four elements of their framework:

- **Networks.** The shape of corruption networks depends on the features of the process and its environment.
- **Organization.** The organization nurtures a set of mechanisms to induce the functionality of networks or to inhibit whistleblowers that may hinder the operations of the corruption schemes.

- **Opacity.** A necessary condition for a corrupt network to act with impunity is to perform under high levels of opacity.
- **Checks & Balances(C&B).** A necessary condition for a corrupt network to act with impunity is to face weak or controllable C&B.

As São Paulo case study refers to a context where the levels of transparency are supposedly high, I will leave the opacity element out, but consider the other three.

Therefore, the corruption persistence framework will rely on elements of power imbalance and path dependence shaping institutions across the years. Personal and professional connections (networks) built over the years and organizational mechanisms constantly adapted allow corrupt practices, while weak or controllable checks and balances ensure impunity.

2.2 Transparency and Accountability – “Transparency Dynamics”, Targets and “More Complex Accountability”

Jonathan Fox (2007) suggests that there are two types of transparency – opaque transparency and clear transparency – and, further, two types of accountabilities: soft and hard accountability.

Opaque transparency involves the dissemination of information that, in practice, does not reveal how institutions function. On the other hand, clear transparency concerns programs and practices of access to information that disclose valuable and reliable information about institutional performance. It is this kind of transparency that could eventually lead, but not necessarily, to hard accountability. Clear transparency, however, is no guarantee of hard accountability since hard accountability requires the intervention of other state actors.

As Meijier (2013) states, the dynamics of transparency involve:

- The institutional relationship between government and society.
- The nature of information exchange (ease, accessibility, timeliness).
- The domains of public administration targeted by transparency policies.

Hence, for the desirable products of transparency to be observed, it is necessary to consider the spaces of exchange between external and internal agents to the public

administration and the relevance of the actors targeted by transparency mechanisms. These dynamics proposed by Meijer are well aligned with the neo-institutionalism framework.

Considering these dynamics, each government agency, each public policy, and each open database may bring different and unique implications in promoting accountability in practice. For this thesis, one element of such dynamics is particularly relevant: the domains of public administration targeted by transparency policies. When discussing transparency, accountability, and anticorruption, the target actors of greater transparency and accountability are relevant to whether or not the expected results are achieved. It is so because, depending on the resources and power of these target actors, they can protect themselves from pressures and from being held to account. Accountability may be seen as a two-dimension concept, which encompasses both answerability – the obligation to respond, to be accountable – and enforcement – “the ability of agencies to impose sanctions and loss of power for those who violate public duties” (Pinho & Sacramento, 2009, p. 8).

In that sense, a crucial question one may formulate is “what is the deterrent capacity of these agencies?”. In other words, how efficient are their “teeth” – and, conversely, “to what extent do mandate holders and their networks of influence concentrate enough power to protect themselves from enforcement actors and demand answerability from them?”.

"More complex" accountability

In Schedler's (2008, p. 7, my translation) definition, accountability refers to “the obligation to open oneself up to the public, obligation to explain and justify one's actions, and subordination to the possibility of sanctions”. O'Donnell (1998) differentiates between horizontal accountability, in which one public body is accountable to another, and vertical accountability, in which the state is accountable to society. Such differentiation already inserts the centrality of the dialectical, dual process, in which an essential element of the concept is that accountability is necessarily directed to something or someone.

Bovens et al. (2014) also highlight the issue of the actors involved in the accountability process. For the authors, “in the strict sense, accountability refers to the requirement that an actor, whether an individual or an organization, be accountable for its actions to an independent authority”. The first question to face about accountability would be therefore to ascertain who those actors are and what kind of action is being considered, “particularly in terms of controlling inappropriate actions”.

Schedler (2008, p. 25) argues that accountability is a “modest concept [that] admits, from the outset, that politics is a human enterprise, which as such is characterized by elements of freedom and indeterminacy”. For the author, accountability must rely on carefully constructed rules, but it must not suffocate the exercise of power in a regulatory straitjacket.

Regarding control bodies and the classic problem embedded in the question “who controls the controllers?”, Schedler (2008, p. 20) advocates a format of a “recursive network that maintains intransitive accountability relationships [in which] A is accountable to B, who is accountable to C, who is accountable to A”.

Citing Collier and Mahon (1993), Schedler (2008, p. 24) also argues that accountability is a “radial” rather than a “classical” concept: “Classical concepts are defined by a hard, invariable core of basic characteristics. Radial concepts, on the other hand, rather than sharing a common essence, share a certain 'family resemblance' (Wittgenstein)”.

A radial concept, however, should not be devoid of clarity, states Schedler (2008), for whom conceptual clarity may bring practical consequences.

By way of conclusion, the author cites some important requirements for the understanding of the term and its concrete consequences. accountability should include the following aspects:

- Be mandatory in nature - to prevent bureaucrats and political actors from presenting accountability as an act of generosity
- Be based on the three pillars (information, justification and accountability) - to avoid “politicians selling us quick platters of confusing information as if it were the full menu of accountability”;
- Maintain a “relative modesty” - so that we don't confuse “accountability with more rigid and expensive techniques of political control”;
- Have a public character - so that we don't “settle for behind-the-scenes control exercises”;
- Recognize its multifaceted nature - to avoid “resigning ourselves to the failure of one or another specific form of accountability”.
- To pay attention to the complementarity of vertical and horizontal controls – in order not to forget and not to diminish the relevance of either of the two “fundamental columns of democratic accountability”.
- Adopt intransitive recursivity - to avoid that “certain islands and summits” escape the requirement of accountability.

Recalling that agent-principal accountability theory was originally developed to explain market relations (and not relations between states and societies), McGee (2020) states that this approach interprets failures in accountability as failures of “visible and recognizable actors”, while other approaches on accountability, “more complex”, according to the author, go beyond managers and public agencies taken individually and also consider “the biases in public institutions and public services that are 'captured' by elites who wish to preserve their own privileges”.

Even more complex, McGee continues, are versions that locate accountability problems in the very systems and structures of society, politics, and economics, and point to the need for a permanent engagement in “accountability politics” to neutralize, replace, and ultimately transform entire governance behaviors, attitudes, and cultures. According to the author, in the absence of accountability, people who are relatively powerless just ask for answers - along the lines of Jonathan Fox's “soft accountability”. At the other extreme, external actors would be able to actually produce answers in the form of sanctions and compensations that exhort powerful actors to play their roles properly - a kind of Fox’s “hard accountability”.

Recounting concrete experiences in the field, working with grassroots organizations in Latin America (Colombia), McGee claims to have had contact with projects that started with observable accountability problems and, instead of being guided by a more careful analysis of the concrete case, ended up being guided by a perspective that sought generic commitments to produce more transparency in the expectation of strengthening accountability (“in precariously defined terms”, according to the author).

Thus, based on the characteristics proposed by Schedler and on McGee's considerations about “more or less complex” theories of accountability, and considering the new perspectives on the phenomenon of corruption – historical institutional approaches, particularly power imbalance and path dependence rather than information asymmetry –, the challenge of holding the powerful accountable might mean not only being able to observe the behavior of political agents or bureaucrats and demanding government transparency in a generic way.

From this “beyond transparency” perspective, it is also necessary to understand the behavior of, for example, the private sector, especially in its relationship with senior government officials and with political, judicial, electoral, and legislative institutions. In this sense, instead of addressing transparency in a generic way, there is now a demand for the opening of specific databases, a step that would be necessary for more focused accountability purposes.

In this context, the debates and demands around transparency and openness of data change significantly, along with the challenges of opening up these types of datasets and making them publicly available or allowing crucial datasets to be shared between different public law enforcement institutions. Moreover, depending on the type of data and what it can reveal, the challenges of accountability, that is, the challenges of holding well-positioned actors in power structures accountable for possible misconduct or criminality, seem to increase.

As accountability processes may considerably change depending on the type of actors to be held accountable, it should be interesting to understand the emerging concept of “grand corruption” as a differentiation from “petty corruption”. The latter essentially describes bribe payments to street-level bureaucrats, usually in order to more quickly access some public service, while the latter refers to high-level corruption involving powerful actors in both government and the private sector.

Ang (2020), for example, suggests that each type of corruption has a motivation: in “grand corruption”, payments are intended to “gain access” to decision-makers and decision processes, while in “petty corruption” the goal is to “speed up” (paperwork and bureaucratic processes). The author then classifies the two types of bribery as “access money” and “speed money”. The latter case means “low-value bribes that businesses or citizens pay to bureaucrats to get around obstacles or speed things up”, while “access money” includes high-value rewards offered by business actors to politicians or top administration bureaucrats not only for speeding up decisions, but also for exclusive access and valuable privileges.

Another author who examines the phenomenon of corruption more closely is Jancsics (2019), who, based on two critical dimensions – “form of transfer” and “principal beneficiary” – proposes four categories of corruption: market corruption; social bribery, corrupt organizations, and state capture. The first two categories relate to the phenomenon of “petty corruption” (one form makes use of more lasting personal ties, while the other relates to immediate impersonal exchange), while “corrupt organizations” and “state capture” are categories more related to “grand corruption”.

Further evidence of the growing interest in the topic of “grand corruption”, a phenomenon involving actors endowed with great economic and financial power, is that in 2017 the signatory countries of the United Nations Convention against Corruption (UNCAC) adopted a resolution entitled “Preventing and combating corruption in all its forms more effectively, including, inter alia, when it involves vast amounts of assets, based on a comprehensive and multidisciplinary approach, in accordance with the United Nations Convention against

Corruption”⁴. From this resolution, two meetings with hundreds of experts were held in 2018 and 2019 to discuss exactly how to control grand corruption, or "corruption involving vast amounts of assets," in diplomatic parlance. I attended both meetings (“Global Expert Group Meeting on Corruption Involving Vast Quantities of Assets”), in Lima, Peru, in December 2018, and in Oslo, Norway, in June 2019, and could directly observe discussions about focused transparency and sector-level corruption.

In the outcome document⁵ from the Oslo meeting in 2019, experts recommend governments to disclose specific datasets that could potentially be useful to counter “grand corruption”, including data on real estate ownership.

In sum, apparently not all transparency and accountability are the same. Depending on the “target”, there seems to be more or less transparency and more or less accountability. What makes the relationship between transparency and accountability more complex than previously stated by scholars and practitioners.

Accountability actors

In a seminal paper on accountability, “Horizontal Accountability in New Democracies”, published in the 1990’s, Guillermo O’Donnell proposes the differentiation between vertical accountability and horizontal accountability. Considering the Robert A. Dahl’s “polyarchy” perspective (free and fair political competition), the author analyzed in his paper the characteristics of accountability in the new political democracies established around the world, particularly around the end of 1980’s and the beginning of 1990’s, after the end of dictatorships.

O’Donnell suggests that vertical accountability is composed by free and fair elections along with freedoms of speech, the press, and association. These aspects “permit citizens to voice social demands to public officials (elected or not) and to denounce these same officials for wrongful acts that they may commit” (1998, p. 112).

Horizontal accountability, on the other hand, is comprised of government institutions whose primary responsibility is to hold elected public officials into account.

Recently authors have suggested a third type of accountability: diagonal accountability (Bernhard et al., 2017; Lührmann et al., 2020; Mechkova et al., 2019), which operates between

⁴ Resolution 7/2 . See <https://www.unodc.org/unodc/en/corruption/COSP/session7-resolutions.html> (last visited on 21 Dec 2022)

⁵ See https://www.unodc.org/documents/corruption/meetings/OsloEGM2019/Oslo_Outcome_Statement_on_Corruption_involving_Vast_Quantities_of_Assets_-_FINAL_VERSION.pdf (last visited on 21 Dec 2022)

the horizontal and the vertical dimensions. Diagonal accountability “reflects the contribution of non-state actors to accountability. Civil society organizations, an independent media, and engaged citizens can use a broad range of actions to provide and amplify information about the government, thereby holding it accountable” (Lührmann et al., 2020, p. 813).

According to Fox (2007, p. 669): “(...) dealing with hard accountability would involve going beyond the limits of transparency and dealing with both the nature of the governing regime in place and the ability of civil society to encourage accountability institutions to do their job”.

In that sense, this thesis will consider accountability actors those actors involved both the horizontal and in the diagonal accountability dimensions: prosecutors, auditors, councillors, comptrollers (horizontal accountability) as well as investigative journalists and civil society organizations (diagonal accountability).

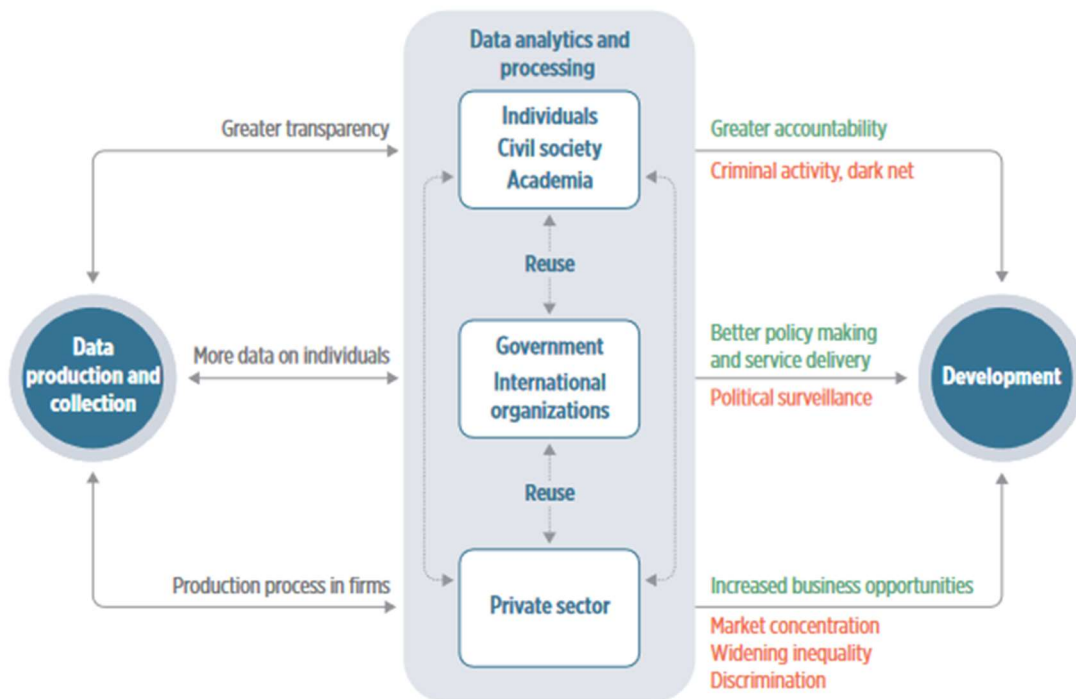
2.3 Open Data for Development – Not All Data Are The Same

Data for development is seen as having transformative positive potentials, in social and economic lives of societies and in terms of governance (Braunschweig et al., 2012; Kitchin, 2021; Taylor & Schroeder, 2015). This subsection explores the challenges in disclosing data for anticorruption purposes in general and, specifically, the potential relevance of ownership data disclosure and analytical use for countering corruption and money laundering at the real estate sector.

Data has become a central issue for multilateral development agencies. The 2021 version of the World Bank annual publication “World Development Report” is entirely dedicated to the transformational power of data. For this reason, the study is entitled “Data for Better Lives”. According to the authors, there are potential gains and risks when it comes to data. The Figure 2 can be found at page 4 of the report.

Figure 2 – Open data for development – the World Bank’s approach

Figure O.1 How data can support development: A theory of change



Source: WDR 2021 team.

Note: Positive impacts are shown in green; negative impacts are shown in red.

As noted, while recognizing possible negative impacts of data, the report assumes that data may bring three major development outcomes: **greater accountability, better policymaking and service delivery and increased business opportunities.**

It is also assumed that individuals, civil society, academia, governments, international organizations and the private sector are all equally able to analyze and process data.

A variety of sectors in society can produce data; however, the public sector makes and stores the most extensive amounts of data; additionally, governments have legal incentives to share data, in contrast to the private sector.

Open government data can basically be defined as non-private, non-confidential data, produced with state resources and made publicly available without restriction on use or distribution (Janssen et al., 2012, p. 258). They can also be defined as “the publication and dissemination of public sector information on the web, shared in a raw and open format, logically comprehensible, so as to allow its reuse in digital applications developed by society” (Cunha et al., 2015, p. 14).

Technology has increased the possibilities of processing and analyzing publicly available data. Hence, since the beginning of this 21st century there has been intense debate

and great enthusiasm regarding the transformations that open data would bring. In recent years, though, starting around 2014, the agenda is in a more realistic phase, in which there is less enthusiasm and more feasible goals, and ambitions are set (Matheus & Janssen, 2020).

In that sense, Albuquerque et al. (2021) propose a new critical approach to the transformative potential of data, suggesting a research agenda that analyzes three aspects, namely, the generation, circulation, and use of data, and further, within each of these pathways, the role played by data could be modulated by “data functions”: referential, metalinguistic, factual, referential, metalingual, factual, conative, expressive (also called emotive), and poetic.

The case study pertaining to this thesis assesses the use of data, focusing, however, on a single function, the referential – in which data offer “indicial access to contextual elements to support transformations” (Albuquerque et al., 2021, p. 6).

Context is relevant to inform data availability, data use and its effect. This case study analyses the uses and barriers for use of a dataset on real estate and urban land ownership. This is a relevant dataset to countering corruption because it offers data about asset ownership, and illicit enrichment is associated with corrupt practices.

To illustrate the emerging calls for the disclose of specific datasets about asset ownership, rather than a vague call for more transparency in general terms, it is interesting to present the case about another asset-related dataset relevant to countering corruption: data on beneficial owners of legal structures.

It is also interesting to discuss the transparency of legal persons in the context of this thesis due to the relationship between the real estate industry and money laundering.

Beneficial ownership transparency

In the context of the international debate among “practitioners” on how to tackle corruption, particularly “grand corruption”, and in line with the referential function of data (Albuquerque et al., 2021) and the “complexification” of accountability (McGee, 2019), one agenda that has recently become relevant is beneficial owner transparency; that is, the disclosure of data on the individuals who control legal entities. A search in the “knowledge hub” section of the NGO Transparency International website⁶ shows that the term “beneficial ownership”, i.e., “ultimate beneficiary”, was almost absent until 2015 (only 13 records, two from 2013; six from 2014; and five from 2015). Between 2016 and 2020 one can find 51

⁶ See <https://knowledgehub.transparency.org/> (last visited on 22 Dec 2022)

records, an annual average of ten. In 2020 alone, there were 15 reports and studies mentioning the term “beneficial ownership”.

The topic of beneficial owner transparency has recently entered the legal arena in some countries. On the last day of 2020, for example, the United States Congress passed the so-called Corporate Transparency Act, which obliges legal entities to inform authorities who the natural persons are that control that legal entity, recording the following data: name, address, date of birth, and a personal document number. Under the law, it is federal and local authorities that can access this data⁷.

The European Union (EU) has taken even faster steps toward complete transparency of the final beneficiary. According to two European anti-money laundering directives (4th directive, 2015, and 5th directive, 2018), as of January 2020 all EU countries should have public registries, i.e., publicly available ultimate beneficiary data (Van der Merwe & Transparency International, 2020).

The premise behind the interest in this particular data set is that many legal entities exist to hide illicit money. And indeed, there is growing evidence that legal structures based in jurisdictions considered “tax havens” are used for this specific purpose, as demonstrated in the scandals uncovered by the Panama Papers and other leaked documents and journalistic reports. Even before the media scandals, a World Bank study found that companies registered in jurisdictions regarded as tax havens were used to hide the proceeds of corruption in 128 out of 150 cases of “grand corruption” analyzed (De Willebois et al., 2011).

The assumptions around the transparency of the ultimate beneficiary align with McGee's call regarding the apparent insufficiency of understanding accountability within the confines of “visible and recognized actors”. In this sense, actors and processes that show the weight and impact of economic and financial power in society – particularly those that seem to be implicated in the phenomenon of “grand corruption” – should also be more open to public scrutiny.

Indicative of the challenges around corruption-related data is a court decision made public in November 2022. According to the Court of Justice of the European Union, “the provision whereby the information on the beneficial ownership of companies incorporated within the territory of the Member States is accessible in all cases to any member of the general

⁷ See “Landmark Bill Ending Anonymous U.S. Companies Is Enacted”, press release by the organization FACT Coalition: <https://thefactcoalition.org/landmark-bill-ending-anonymous-u-s-companies-is-enacted/> (last visited on 21 Jan 2022)

public is invalid”⁸. The court decision refers to a 2018 European Union Directive established to prevent the use of the financial system for money laundering or terrorist financing.

This case seems to demonstrate the relevance of considering external actors and interests beyond the public sector in shaping transparency and accountability policies. It is also evidence of the centrality of interactions between society and state in given contexts to explain transparency policies, as Meijier (2013) argues.

Legal structures and the real estate sector

One of the main arguments in favor of more transparency on who controls legal persons refers to the purchase of real estate properties. Evidence gathered from leaked cases published in reliable press⁹ and technical studies on “grand corruption” and the real estate sector¹⁰ present this phenomenon. In 2022, FATF, an intergovernmental organization, published a guidance¹¹ for the real estate sector. In the document, the organization states that real estate “attracts criminals who use real estate in their illicit activities or to launder their criminal profits” and added that, according to their assessments, “the real estate sector often has poor understanding of these risks and regularly fails to mitigate them”.

Although the use of companies to acquire real estate is not illegal per se, there are several documented cases in anticorruption investigations in which the judiciary has ruled that purchases of real estate, usually of a high standard, constitute money laundering obtained through acts of corruption. The first judicial conviction in the scope of the so-called “Lava Jato Operation”, an impactful corruption scandal in Brazil, is an example of this. A former top-executive of the Brazilian state-owned oil giant Petrobras was convicted in the first instance for having purchased a penthouse in Ipanema, Rio de Janeiro, through a company based in Uruguay. The company had no activity in Brazil, no offices or employees. And the only

⁸See “Anti-money-laundering directive: the provision whereby the information on the beneficial ownership of companies incorporated within the territory of the Member States is accessible in all cases to any member of the general public is invalid”, press release 188/22 (<https://curia.europa.eu/jcms/upload/docs/application/pdf/2022-11/cp220188en.pdf>) Last visited on 18 Jan 2023)

⁹See “Secret real estate purchases are a driving force behind the offshore economy” (<https://www.icij.org/investigations/pandora-papers/secret-real-estate-purchases-are-a-driving-force-behind-the-offshore-economy/>) 3 Nov 2021

¹⁰See for example Kumar, L., & de Bel, K. (2021, August). Acres of Money Laundering: Why U.S. Real Estate is a Kleptocrat’s Dream. Global Financial Integrity. (<https://gfindegrity.org/report/acres-of-money-laundering-why-u-s-real-estate-is-a-kleptocrats-dream>)

¹¹See “Risk-based Approach Guidance for the Real Estate Sector” (<https://www.fatf-gafi.org/en/publications/Fatfrecommendations/Guidance-rba-real-estate-sector.html>) Last visited on 18 Jan 2023). The Financial Action Task Force (FATF) is an inter-governmental body that sets international standards that aim to prevent global money laundering and terrorist financing

operation carried out in the country was exactly the purchase of the property in Rio de Janeiro (Brandt et al., 2015).

A study I coordinated, published in 2017¹² based exactly on the IPTU dataset made available a year earlier, pointed to the existence, in the area known in São Paulo as “expanded downtown” (the city's historic center plus middle and upper-middle-class districts near the most central regions of the city), of more than 3,4 thousand properties belonging to 236 companies formed in São Paulo. As a champion in open government policies, São Paulo’s company register agency offers the possibility of accessing data on who are the founders, partners and the controllers of any given legal person established in São Paulo state. Even with such data, it was not possible to determine who were the natural persons behind those 236 companies, because they were controlled by other legal persons, corporate arrangements based in non-transparent jurisdictions, among them the British Virgin Islands, the US state of Delaware, and Uruguay.

This anecdotal evidence demonstrates the complexity behind open data for anticorruption purposes: even when crucial datasets are made publicly available — real estate property and company register —, layers of opacity may exist, making it difficult for accountability actors to investigate and sanction corruption or money laundering even when they can access information about persons (natural or legal) who directly own properties.

In that sense, this thesis assumes that not all data are the same, in terms of generation, circulation and use (Albuquerque et al., 2021). The dynamics around data availability and data reuse seems not to be identical considering the three development goals that data can contribute for, according to the World Development Report 2021 – Data for Better Lives: “greater accountability”; “better policy making and service delivery”; and “increased business opportunities”. Apparently, data availability and data use for policy analyses or business purposes can be differentiated from the dynamics around data use for anticorruption purposes.

2.4 Open Data and Corruption – a Systematic Literature Review

The first immediate insight from the systematic review is that the literature on open data and corruption is extremely limited. There is extensive literature on corruption on the one hand, and extensive literature on open data on the other. However, the convergence between the two

¹² See “Does corruption live next door? Shell companies and the real estate sector in the largest city in the Southern Hemisphere” (<https://knowledgehub.transparency.org/product/s%C3%A3o-paulo-does-corruption-live-next-door-shell-companies-and-the-real-estate-sector-in-the-largest-city-in-the-southern-hemisphere>)

is minimal. The term “transparency” is much more commonly found in the literature on corruption; however, the term “open data” is still incipient in that literature.

The second finding of the literature review relates to the accountability debate, as it shows that the most commonly used analytical lens in this nascent literature is still the “principal-agent” one, according to which the principal is society, and the agent is the state. The principal-agent theory is often made explicit in the texts, according to which greater data openness would lead almost naturally and necessarily to greater accountability of state actors and institutions. This indicates the prevalent approach of a “more simplified”, “less complex” version of accountability.

Therefore, there is a gap in the literature regarding more in-depth analyses of the possible impact of open data on controlling corruption. There are practically no sophisticated analyses about the potential of open data for anticorruption policies or about the accountability dynamics that this greater transparency could strengthen. There are also no studies on the relationship between power asymmetry and greater transparency and accountability.

Procedures for the systematic literature review

The review of the literature on open data and corruption sought to analyze how authors approach each of the two concepts/phenomena and how they relate them, and also to assess whether and how the theme of accountability connects the discussions on open data and corruption.

I performed the mapping following the flow proposed by the PRISMA approach for systematic literature reviews (Moher et al., 2009). The search was performed in the Web of Science, Scopus and Ebsco in June 2021.

I searched for the term “open data” or “open government data” in a first moment; for the word “corruption” in a second moment; and then I combined the two searches. I did not put any filter or limitation - neither of languages, nor of types of documents, nor of time.

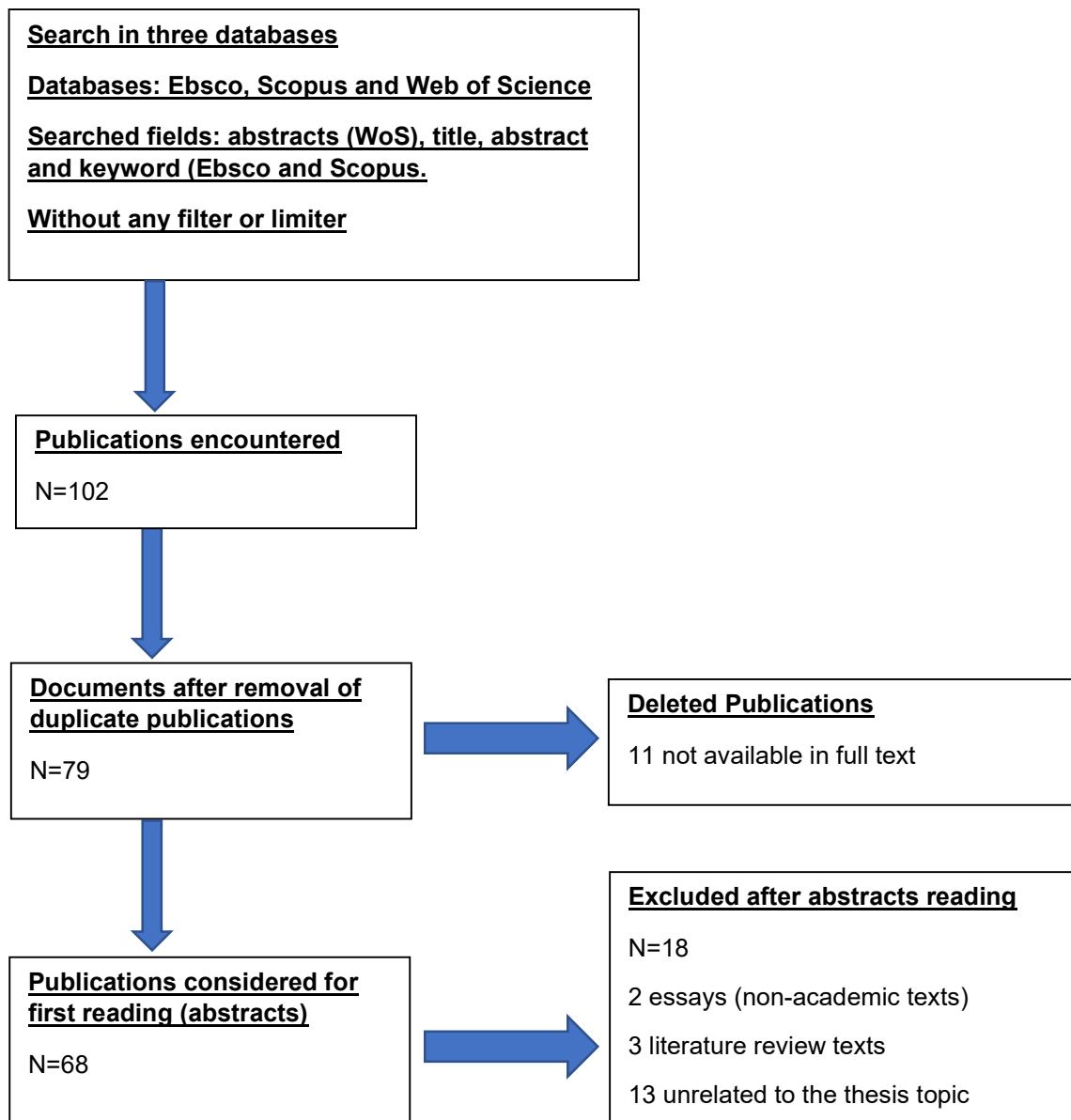
From the outset, one can see that there is a profusion of works dealing with open data and corruption, separately. As seen below, one will find thousands of publications — books, papers, presentations — using the expression “open data” and dozens of thousands of documents on “corruption”.

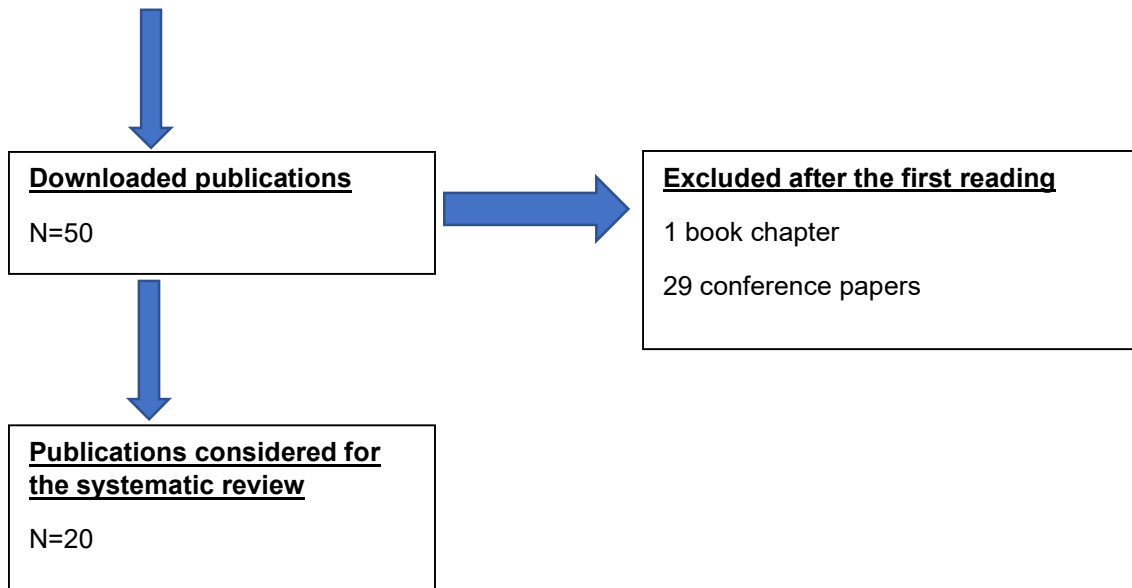
However, there are far fewer publications that mention the two phenomena together (given the specificities of the tools, the search fields are slightly different between the databases).

Table 5 – Number of publications meeting the search criteria – open data & corruption

Research Databases	Field(s)	“Open (Government) Data” (# documents)	Corruption (# documents)	“Open Data” & Corruption (#)
Web of Science	Abstract	5.274	19.564	24
Ebsco	Title, abstract and keyword	3.227	38.851	12
Scopus	Title, abstract and keyword	12.929	34.011	66
<i>Totals</i>		<i>21.486</i>	<i>101.539</i>	<i>102</i>

Among the 102 publications found in the sum of the three databases, there were duplicates, which were detected by the similarity in titles and authors. After eliminating the duplicates, 79 remained.

Figure 3 – Diagram of the systematic literature review on “open data” and “corruption”, according to the PRISMA approach



I could not find eleven of the 79 initially identified publications – not even using platforms such as Google Scholar and Research Gate.

From the remaining set of 68 documents, I removed 18 due to the following reasons:

- Two texts were essays published in non-academic journals (“Chronicle of Philanthropy” and “GIM - Global Magazine for Geomatics”).
- Three texts dealt with literature review (Bachtiar et al., 2020; Herala et al., 2016; Matheus & Janssen, 2020). They have not used in other parts of the thesis, but have been considered in the discussions about corruption because they do not develop this item further.
- Thirteen that are unrelated to the thesis topics:
 - Eight deal with a third subject, unrelated to open data and corruption, only making passing mention of the term “corruption” and the expression “open data” (Bader, 2012; Coates, 2019; Chornous & Iarmolenko, 2019; Jakóbič, 2020; Kakileti et al., 2020; Nowak, 2017; Nowak, 2018; Vijayakumar & Bharathi, 2021)
 - Three deal with corruption or related topics but do not address “open data” - these texts only appeared in the search results because these publications are open access — therefore they carry among their keywords the phrase “open data” (Köbis et al., 2017; Lin et al., 2018; Vitvitskyi et al., 2020).
 - One deals with corruption, with a brief mention of the expression “open data”, without developing the relationship between them and without even mentioning the term “accountability” (Gross et al., 2018);

- One deals with open data and only briefly mentions the word corruption without mention of the word "accountability" (Ridgway, 2016).

Systematic literature review

After excluding 29 papers presented at conferences and one book chapter, the 20 articles considered for the systematic review remained. These articles are scattered in 17 different journals:

Table 6 – Journals that published the 20 selected articles

Journal	# articles
Government Information Quarterly (GIQ)	3
International Review of Administrative Sciences (IRAS)	2
Asian Social Science; Innovation: the European Journal of Social Sciences; International Journal of Electronic Governance; International Journal of Electronic Government Research; International Journal of Public Sector Management; International Political Science Review; Investment Management and Financial Innovation; Journal of Enterprise Business Management; Profesional de la Informacion; Public Administration and Information Technology; Revista Latinoamericana de Investigación en Organizaciones, Ambiente y Sociedad; Revista Trilogía; Telematics and Informatics; Tripodos; World Customs Journal	1 each

Systemic approach - data, accountability and control of corruption

In a book first released in 1967, the Italian thinker Umberto Eco defined as “apocalyptic” those who saw the then-emerging phenomenon of mass media as platforms for manipulation and authoritarianism. On the other hand, the “integrated” ones were on the opposite extreme: they saw a way for social improvement in the then-emerging phenomenon with the renewal of cultures and practices (Eco, 2011).

In discussions of digital technologies, one could say that there are also the optimistic authors and the ones who are not so positive. One does not find apocalyptic analyses in the literature on open data, but rather, a more realistic, less romanticized view of the positive impacts of open data. Still, on the other hand, it is possible to observe the “integrated” authors, for whom open data will contribute to social improvements, including fighting corruption.

In the “integrated” group are articles that deal more specifically with open data, e-government, transparency, open government, or digital government without deepening the relationship of these mechanisms with accountability or the control of corruption. These two phenomena – accountability and corruption – are mentioned superficially. Greater

accountability and corruption control are seen as likely positive consequences of open government or open data policies.

Table 7 – Articles on open data and corruption arguing for likely positive outcomes of open government policies

“Integrated”		
Article	Authors	Excerpt
Preliminary insights from the Philippine Bureau of Customs imports database	Mendoza & Ko (2015)	“Open data and greater transparency in public sector operations and transactions can serve as powerful levers for supporting and sustaining reform” (p. 88).
Motivations for open data adoption: An institutional theory perspective	Altayar (2018)	“OGD enhances transparency and <i>Accountability</i> , facilitates access to government data, supports innovation, improves government services, operational benefits and encourages participation” (p. 10).
Ensuring transparency of key public finance authorities	Bukhtiarova et al. (2019)	“Transparency and openness of public finances are one of the tools of combating corruption, since it provides users with reliable and timely information, contributes to the formation of a responsible government and prevents abuse by public finance authorities”
La transparencia de los partidos políticos como estrategia electoral. Una evaluación de sus promesas y sus páginas web	Diéz Garrido et al. (2019)	“Transparency has become one of the most valued aspirations that politicians want to demonstrate to the public, as it is a sign of legitimacy, evolution, and the fight against corruption” (p. 83).
Estrategia de gobierno digital para la construcción de Estados más transparentes y proactivos	Toro-García et al. (2020)	“un Gobierno transparente es aquel que publica oportunamente datos abiertos de las actividades de sus diferentes organismos, bien sea para dar respuesta a los requerimientos de la ciudadanía o bien por iniciativa propia. Con esto se busca no solo poner un alto a la corrupción, sino también garantizar el derecho de los ciudadanos a conocer lo que los gobiernos hacen con sus impuestos” (p. 72).
Open government data portal usability: A user-centred usability analysis of 41 open government data portals	Nikiforova & McBride (2021)	“Open government data, as a phenomenon, may be considered an important and influential innovation that has the potential to drive the creation of public value via enabling the prevention of corruption, increase in <i>accountability</i> and transparency, and driving the co-creation of new and innovative services” (p. 12).
Participación, colaboración, transparencia y datos abiertos.	Acosta & Plata Pineda (2016)	“Es indispensable destacar que las iniciativas de Datos Abiertos constituyen un escenario propicio para el desarrollo

Tras la definición de Gobierno Abierto.		eficiente de las políticas de Gobierno Abierto, ya que la información en bruto proveída (o lo que es lo mismo, el conjunto de empresas, emprendimiento, academia, el activismo social, la ciudadanía, etc.) son la materia prima para la generación de nuevos servicios y aplicaciones de valor agregado” (p. 220).
Portales de datos abiertos. Metodología de análisis y aplicación a municipios españoles	Royo-Montañés & Benítez-Gómez (2019)	“(…) se analiza la información económico-financiera que ofrecen los municipios en sus portales de datos abiertos, por ser fundamental para la rendición de cuentas, la lucha contra la corrupción y la evaluación de la sostenibilidad y eficacia de la prestación de servicios públicos” (p. 5).
Public value of e-government services through emerging technologies	Valle-Cruz (2019)	“The e-government has the potential to improve transparency and efficiency, as well as to reduce corruption and bureaucratic bad practices, as well as to boost citizen participation and improve government to citizen interaction” (p. 538).
Empowering communities and Improving Public Services Through Open Data: South African Local Government Perspective	Bvuma & Joseph (2019)	“Open data plays an important role in local government and benefits communities and has potential to increase the key principles of transparency, participation, accountability” (p. 157).
Evidence of an Open Government Data Portal Impact on the Public Sphere	Lourenço (2016)	“Evidence was found of data disclosed through a portal actually being used for a wide range of purposes; (..). These purposes include monitoring resource misuse and public spending overview analysis, to discuss (question and explain) specific contracts and spending options, to hold public agents accountable in what concerns the tendering process, the outputs and the outcomes of governmental actions, to prevent and denounce possible cases of corruption or nepotism, and to support arguments within broader policy discussions occurring in the public sphere”.
Role of E-Government and Open Data to Enhance Transparency Practices	Sriyakul & Jermstiparsert (2020)	“(…) e-government adoption and open data have significant impacts on transparency”

Some “integrated” articles analyzed during the literature review shed light on issues of interest to the present thesis. For example, by analyzing a portal containing data on public procurement in Portugal and concluding that the information there is indeed used for accountability purposes, Lourenço (2016) contributes to the reflection on the use of data and

the relevance of specific datasets. In the same vein, Mendoza and Ko (2015) argue that the public availability of data on imported items in the Philippines can help detect possible fraud. Also, Royo-Montañés and Benítez-Gómez (2019) conclude, their analysis of data openness by municipal governments in Spain, that Spanish cities are more interested in publishing data for business use, for innovation purposes, than for accountability purposes. These findings present insights about generation (types of data specifics), circulation, and use of open data for anticorruption purposes.

In presenting a systematic review of the literature on open data, Matheus and Janssen (2020) identify the period between 2011 and 2013 as a hype phase, a phase of enthusiasm about the potential of open data. Around 2014, however, a new stage began: a “realistic phase”, in which more feasible ambitions for open data are established.

Since Umberto Eco's term “apocalyptic” would not fit accurately, I will label the second group of articles identified in this literature review on open data and corruption as “realistic”.

In this group are articles that understand open data to be relevant but suggest that it takes more than opening data to achieve benefits for society, such as better control of corruption.

Table 8 – Articles on open data and corruption establishing feasible ambitious for open government policies

“Realistic”		
Article	Authors	Excerpt
The paradox of the right to information law in Africa	Adu (2018)	“It is significant to note that making information available will not prevent corruption if the conditions for publicity and <i>accountability</i> are weak” (p. 5).
Transparency-enabling information systems: trust relations and privacy concerns in open governance	Gritzalis et al. (2019)	“Although the capacity of IT for enabling and intensifying transparency is evident, it is challenging to design, implement, diffuse and institutionalise such systems (...) The adoption of open government systems is associated with substantial changes, such as the alteration of work processes in public organisations, the requirement for interoperability with already existing systems, the debasement of the pre-existing power balance among stakeholders, the handling and overriding of users’ resistance, and many others”.
Social-Communicative Innovations in Anti-Corruption Activities (Regional Aspect)	Leontieva et al. (2015)	“(...) optimization of the anticorruption system of public communication requires developing complex social innovation technologies integrating information communication, media, advertising, reflexive motivational, and PR approaches” (p. 392).
Evaluating the role of online data availability: The case	Murillo (2015)	“Our results lead us to conclude that the evaluated government portals, as tools for providing economic and institutional transparency in relevant

of economic and institutional transparency in sixteen Latin American nations		areas where corruption occurs in these nations, are not living up to expectations as they fail to leverage the publicity condition (i.e. the lack of education, media reach, ICT illiteracy, and illiteracy)".
E-government as an anticorruption tool: panel data analysis across countries	Park & Kim (2020)	"(...) open government as one type of e-government is dependent on the rule of law to reduce corruption" (p. 12).
From information to predictability: transparency on the path to democratic governance. The case of Romania	Schnell (2018)	"This article suggests that transparency advocates should pay more attention to issues such as increasing citizen participation, opening up decision-making rather than just data, strengthening the rule of law, and fighting corruption".
Do FOI laws and open government data deliver as anticorruption policies? Evidence from a cross-country study	Žuffová (2020)	"The results add to the previously accumulated knowledge in the field and demonstrate that access to free and pluralistic media and internet might be crucial for any transparency measure to be effective also as an anticorruption measure. It might not be enough for the information and data to be out in the public domain; the information needs to be acted upon" (p. 11).
Open government partnership: unutilized potential in post-communist EU members? (Case of the Czech Republic).	Laboutková (2018)	"(...) open government might mean transparency about government if the effort will focus rather on political objectives (like a matter of proactive publication of information, equal access to information and documents for all citizens, or the obligation of public bodies to consult with citizens and other stakeholders before a decision is made) not technological ones (like a type of medium, format, etc.) and rather on motivations (like a desire to rebuild citizens' trust in government) not stimulations (like pressures from external bodies to improve governance)" (p. 370).

One can see, therefore, that the more "realistic" authors conclude that transparency is not enough to control corruption, emphasizing either issues related to social control (freedom of the press and universalized internet) or issues related to institutional control ("rule of law" and judicial independence), or both types of accountability.

Although this second group of studies establishes the relationship between open data, accountability and corruption more clearly, it does not delve into the concepts of accountability and corruption.

In the discussion on accountability, the texts use the more classical concept, based on the principal-agent theory, in some cases explicitly mentioning this approach (Murillo, 2015; Žuffová, 2020) and, in others, the classical "less complex" concept of accountability is evident, but there is no mention. The analytical framework "publicity and accountability conditions" is

used in the two articles that explicitly mention principal-agent theory. Here, the “publicity condition” refers to moments when information reaches the principal, and he/she is able to process the information and thus acquire knowledge. The “accountability condition” is represented by fair and free elections and other tools that allow the principal to control the agent.

The discussion of types of corruption and any implications or distinctions concerning transparency and accountability processes are almost non-existent in the selected articles – only three articles touch on the topic of grand corruption. Leontieva et al. (2015) mention the relationship between corruption and power when discussing the corruption formula created by Robert Klitgaard (1998): Corruption = monopoly + discretion - accountability and highlight the monopoly of power. Žuffová (2020) mentions the difficulty of opening specific databases, such as those of contracts, company records, and land holdings, which are “extremely important for monitoring the use of public resources and exposing corrupt behavior” (Žuffová, 2020, p. 6). Finally, Laboutková (2018) highlights the difficulty countries have in regulating lobbying and providing transparency of lobbyist records.

None of these papers, however, discusses in more depth the role of actors potentially involved in corruption cases in the shaping of institutions nor do they discuss the ability of these actors to escape accountability obligations.

Table 9 – Summary of the literature review on Open Data and Corruption

	Potentials/Challenges
“Integrated”	<ul style="list-style-type: none"> • Impact on transparency • Monitor of resource misuse • Overview analysis • Hold public agents accountable • Prevent and denounce possible cases of corruption or nepotism • Support of arguments within broader policy discussions • Effectiveness of public service delivery. • Support innovation • Sign of legitimacy, evolution
“Realistic”	<ul style="list-style-type: none"> • Information needs to be acted upon (not only available in public domain) • Need of free and pluralistic media, universal access to internet, ICT literacy • Need of rule of law in the context • Need of strong publicity and accountability conditions • Additional attention to civic participation and decision-making processes • Design, implementation, diffusion and institutionalization of IT systems

It can be said, therefore, that this emerging literature on open data and corruption is beginning to deepen by looking at more specific databases and employing accountability concepts to help explain the practical limitations of the role of open data on controlling corruption. However, the analytical lens used to describe accountability, when used, is the classic, “less complex” one that locates, as McGee points out, accountability failures as the only faults of public officials, political leadership, or state agencies. Furthermore, none of the articles analyzed deals with grand corruption.

2.5 Open Data, Accountability and Corruption: “The Publicity and Accountability Conditions” and “Corruption Persistence”

When discussing the incentives and barriers for anticorruption activities focusing on accountability actors, context and perceptions should matter, as the literature on corruption as a collective action suggests. Moreover, academic papers on corruption using historical institutionalism lenses also seem to indicate that corruption persists even when “critical junctures” and windows of opportunity appear, as seen in case studies about Greece and Mexico mentioned in the literature review above (Aguirre-Ochoa & Gómez, 2021; Trantidis & Tsagkroni, 2017). In that sense, Transparency International’s most well-known indicator, which measures **perception of corruption in the public sector** — therefore it does not say anything about corruption *per se* —, seems to be a relevant indicator in the context of this thesis.

Corruption in the public sector has been perceived as high for many years in countries of the so-called Global South. Among the G20 countries neither located in Europe nor in North America (and excluding Australia, Japan and South Korea), this is the scenario (the lower the score, the more corrupt a country is perceived):

Table 10 – Transparency International’s CPI (Corruption Perception Index) for the G20 countries of the “Global South”

Country	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Saudi Arabia	44	46	49	52	46	49	49	53	53	53
China	39	40	36	37	40	41	39	41	42	45
South Africa	43	42	44	44	45	43	43	44	44	44
India	36	36	38	38	40	40	41	41	40	40
Argentina	35	34	34	32	36	39	40	45	42	38
Brazil	43	42	43	38	40	37	35	35	38	38
Indonesia	32	32	34	36	37	37	38	40	37	38
Turkey	49	50	45	42	41	40	41	39	40	38
Mexico	34	34	35	31	30	29	28	29	31	31
Russia	28	28	27	29	29	29	28	28	30	29

As noted in the table above, in nine years the perception of corruption has not changed significantly in most of these countries. During this nine-year period, the three Latin American countries on the list saw public sector reforms toward more transparency. In Brazil, the access to information law came into force in 2012, including the obligation for governments to publish information online in open, machine-readable¹³, formats. In Argentina, similar legislation came into force¹⁴ in 2016, and a government agency dedicated to assuring transparency was created. In Mexico, constitutional and legal reforms¹⁵ took place in 2015 and 2016, including the consolidation of a national transparency system and strengthening the national transparency agency.

Apparently, thus, institutional reforms pushing for more transparency have had a negligible effect on the overall perception of corruption in the public sector in these countries, including Brazil.

In that sense, the “publicity and accountability condition” (Lindstedt & Naurin, 2010), is a valuable approach. As the authors argue, in the principal-agent perspective, transparency might not be enough to curb corruption because the information made available needs to a) reach the public, who in turn need to b) be able to process such information and c) be enabled to act on the acquired knowledge.

For this thesis, though, I’ll use a more nuanced approach, adding specific elements to the concept of “publicity and accountability condition”.

As pro-transparency institutional and legal reforms are apparently failing to address corruption in Global South and the phenomenon of corruption seems to be even more challenging at the municipal level I will focus not on the “public”, generally considered, but

¹³Article 8th of Brazil’s access to information law (https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm) states that “It is the duty of public agencies and entities to promote, independently of requests, the disclosure (...) of information of collective or general interest that they produce or hold in custody. (...)

§ 2 In order to comply with the provisions in the caput, public agencies and entities must use all legitimate means and instruments at their disposal, with disclosure on official sites on the World Wide Web (Internet) being mandatory.

§ 3 The sites referred to in § 2 must, in the form of a regulation, meet the following requirements, among others:

- I - contain a content search tool that allows access to information in an objective, transparent and straightforward manner, and in easy-to-understand language;
- II - allow the recording of reports in several electronic formats, including open and non-proprietary ones, such as spreadsheets and text, in order to facilitate the analysis of the information;
- III - enable automated access by external systems in open, structured, and machine-readable formats;
- IV - disclose in detail the formats used for structuring the information;

¹⁴The Argentine Law 27.275 created AAIP, an acronym in Spanish for Agency for Access to Public Information. See <https://www.argentina.gob.ar/justicia/derechofacil/leysimple/acceso-la-informacion-publica>

¹⁵Mexico’s agency is INA – National Institute for Access to Information; before the 2015 reform, the agency had only mandate over the federal level of government and it was called IFAI - Federal Institute for Access to Information. See <https://www.gob.mx/sfp/acciones-y-programas/reforma-en-materia-de-transparencia-64474>

instead on those in charge of institutional and diagonal accountability — institutional accountability actors (such as prosecutors, auditors, and comptrollers), and diagonal accountability actors (investigative journalists and anticorruption activists). In addition, following Hulstijn et al. (2017) arguments, accountability actors operating with the “publicity and accountability conditions” should be those who are experienced in participating in a “forum of actors”.

Hence, the “public” in the classical concept is replaced by “community of practice of accountability”.

The expression "community of practice" became well-known following a 1991 publication by scholars Etienne Wenger and Jean Lave, who defined it as a set of relations among persons, activity, and world, over time and in association with other tangential and overlapping communities of practice. A community of practice is an intrinsic condition for the existence of knowledge, not least because it provides the interpretive support necessary for making sense of its heritage (Lave & Wenger, 1991).

However, as the principal-agent framework is connected with the economic paradigm for analyzing corruption, and this thesis will use the neo-institutional approach instead, the adapted version of the “publicity and accountability condition” is not the only theoretical lens used here.

In addition, I will test the concept of “corruption persistence”, framed from the historical institutionalism approach, especially the elements of power imbalance and path dependence. According to the collective action approach to corruption, the main deficiency of the information asymmetry view is that it assumes that whenever there is information available, accountability actors will automatically take act and hold powerful to account.

The Table 11 then presents the perspectives and concepts that will serve to illuminate the core elements of this case study.

Table 11 – Theoretical perspectives used in the thesis

Concept/theoretical approach	Definition	Core elements
<i>Publicity and accountability condition</i>	Relevant and specific datasets reach communities of accountability practices, and these actors have the technical skills to interpret data and are free to challenge the authorities (Adapted from Lindstedt & Naurin, 2010)	Open data, data use and “more complex” accountability
<i>Corruption persistence</i>	Power imbalance and path dependence shape institutions and society across history, protecting corrupt actors from being properly held to account. Personal and professional connections (networks) built over the years and organizational mechanisms constantly adapted allow corrupt practices, while weak or controllable checks and balances ensure impunity (Hall & Taylor, 1996; Meza & Pérez-Chiques, 2020; Persson et al., 2013).	Corruption as a result of power asymmetry and path dependence

These are the theoretical lenses that guide my research. I would not say, however, that this is a theory-based approach; instead, I will use a mixed approach, combining what emerges from data collection with existing concepts and theories. I will use an adaptation of the so-called “Gioia method” (Langley & Addallah, 2011).

3 METHODOLOGY

This section presents the methodology used in the thesis. I start with a brief epistemological positioning, followed by the research design. Next come the exposure of my choices for data collection and analysis. This section is concluded with a discussion of the validation criteria.

3.1 Epistemological Positioning

This thesis holds an interpretivist position, assuming that realities are socially constructed (Cunliffe, 2011; Watson, 1995). In applied social sciences, the object of study is often a social phenomenon, the functioning of a political institution built by political actors or policy interventions designed by human beings with preferences, views, and ideologies. In that sense, this study rejects the positivist paradigm.

This thesis investigates open government data as an anticorruption intervention in megacities. Specific actors in society are expected to use information made publicly available via transparency mechanisms aimed to holding authorities to account. The events and human actions analyzed in the thesis go far beyond rationalities and may be perceived differently by the diversity of actors. Therefore, it is reasonable to state that this reality is socially constructed.¹⁶

3.2 Research Design

According to Stake (1998), there are three types of case studies:

¹⁶At this point, it is also crucial to expose my position before this São Paulo case study. I worked as a top civil servant at the Office of the Comptroller General of the São Paulo City Hall. For two years (from 2013 to 2015), I led the local government's implementation of access to information and open data policies. As the Head of the Integrity Department, I was directly involved in implementing access to information procedures and establishing open data portals. Before that, I was an investigative data-driven journalist and a civil society anticorruption activist. After serving the City Government, I worked as a senior consultant for Transparency International – Brazil as well as advisor for transparency, social accountability and anticorruption projects for United Nations agencies (UNODC, UNDP) and the World Bank. In that sense, both the institutional and social accountability ecosystems in Latin America, Brazil and Sao Paulo are familiar to me.

While working at the City Government, I learned that one of the critical datasets journalists and activists in São Paulo wanted to access was the IPTU database. Also, as a civil servant, I was aware that the local Finance Department refused to allow access to this database even within the government. Other departments of the City Hall could not access this information before the national access to information law and before the establishment of transparency policies in the local government.

In that sense, I was aware of the enthusiasm around this dataset when it was eventually disclosed in March 2016.

- Intrinsic: case studies conducted primarily because the researcher wants better understanding of a particular case, because of its richness, but not necessarily with the intention of theory building.
- Instrumental: case studies conducted with the aim to provide insight into an issue or to redraw a generalization. The case plays a supportive role, as it facilitates our understanding of something else. Researchers will still look at it in-depth – however they are in search of something external to the case.
- Multiple (or collective) case studies: instrumental studies extended to several cases. In collective case studies, there is even less interest in one single case.

This thesis is an instrumental case study. I wanted to investigate the São Paulo case, assuming that the findings of this study will bring insight into a more nuanced understanding of the role of open data policy to counter corruption, particularly in megacities of the Global South.

In that sense, I analyzed the role of the disclosure of the real estate and urban land property in São Paulo. The complete database, including the full name of each owner of properties in São Paulo City, was publicly available, cost-free, on Internet, for six years: from early 2016 to early 2022. The column with the names of natural and legal persons was deleted from the public version of the dataset following the enactment of the data privacy legislation.

Assuming that the expected outcome of this open data initiative depends on the uses and the practical consequences of such data use, I interviewed accountability actors, both from diagonal accountability and institutional. Additionally, I sought publicly available documents published by accountability institutions, such as media reports and or audit reports. Finally, I also counted on my direct observational point, once I have been an active participant in the integrity ecosystem in São Paulo and Brazil for 15 years. In this period, I had multiple professional roles: anticorruption civil society organizations manager and senior consultant, high-level civil servant at São Paulo City Hall, researcher on transparency and anticorruption.

An analysis of São Paulo City's decision to disclose a critical anticorruption dataset can bring valuable insights into the relationship between open data and corruption control at the local level in the Global South context. Political and administrative costs, legal challenges, and the fact that the case occurred during a period in which anticorruption was high in the public agenda — all these events add layers of interest to the São Paulo case.

3.3 Data Collection

Data for case studies can be collected from a variety of sources and modes, such as documents, interviews, and observation. For this case study, I used semi-structured interviews, document analysis and my own observational point of view. The interviews with experienced accountability actors were crucial for understanding the case.

Interviews

For this in-depth case study, I have established a number of 20 interviewees, from both social and institutional accountability organizations. According to Strauss and Corbin (1990, p. 176), theoretical sampling is “sampling on the basis of concepts that have proven theoretical relevance to the evolving theory”. For Bowen (2008), an “appropriate” sample comprises participants who best represent or have knowledge of the research topic. In this sampling strategy, the researcher does not seek 'generalizability' or 'representativeness.' Therefore, the focus is on sampling adequacy instead of sample size.

Following the previously outlined theoretical framework and the importance of fora of actors or “communities of practices”, I identified these communities. To more precisely prepare an adequate sample, I also selected institutions operating from and based in São Paulo. Combining both the relevant fora and the selected institutions – better serves the objective of choosing names of accountability professionals.

Following direct experience, based on my previous academic and professional background in São Paulo, and the main elements of the thesis — open data, accountability, and corruption –, I identified four relevant communities of practice and four institutions that meet the criteria. They are recognized as communities and organizations that promote data access, data use, accountability practices, and the anticorruption agenda. Although not a participant in any of the four communities identified, the São Paulo City Council is on the list because of the oversight function of the Legislative power. These are the eight communities and organizations in alphabetical order:

- **Abraji** - Portuguese acronym for *Brazilian Association of Investigative Journalism*. The objectives of the association are threefold: i) professional improvement of journalists (with the diffusion of the concepts and techniques of

investigative reporting); ii) defense of the right of access to public information; and iii) the defense of freedom of expression.

- **Câmara Municipal de São Paulo.** Portuguese for *São Paulo City Council*. It is the legislative body of the municipality of São Paulo, in Brazil. It is composed of 55 councilmen, the maximum number established by the 1988 Constitution. Considered the largest municipal legislative house in Brazil, it was created in 1560 and is also one of the oldest.

- **CGM-SP** - acronym for the *Municipal Office of the Comptroller General of São Paulo City Hall*. CGM-SP “is responsible for ensuring the defense of public assets, promoting transparency, and preventing corruption in municipal management”¹⁷.

- **Conselho da Transparência do Estado de São Paulo** - *The Transparency Council of São Paulo state*, a body that “has an advisory nature and is responsible for proposing guidelines, methodologies, mechanisms, and procedures aimed at increasing institutional transparency in the state public administration, with a view to preventing the misappropriation of public resources, efficient management, and ensuring administrative morality”¹⁸.

- **ENCCLA** – Acronym for *National Strategy to Combat Corruption and Money Laundering*. ENCCLA is “the main Brazilian institutional articulation network for the arrangement, discussion, formulation and implementation of public policies and solutions to fight corruption and money laundering. The Strategy was created in 2003 and currently includes approximately 90 public institutions - belonging to the three Branches of Power (Executive, Legislative and Judiciary) and the Public Prosecutor's Office, also covering the federal, state and, in some cases, even municipal spheres - and entities”¹⁹.

- **Fiquem Sabendo (FS).** A journalism agency specialized in the Brazilian access to information (ATI) law. Their website reads: “Fiquem Sabendo is an independent data agency specialized in the Access to Information Law (LAI). Our mission is to battle to reveal data and documents hidden from society, while training

¹⁷My direct translation from Portuguese from the CGM official page:

https://www.prefeitura.sp.gov.br/cidade/secretarias/controladoria_geral/acesso_a_informacao/index.php?p=151865

¹⁸My direct translation from Portuguese from the Transparency Council official website. See <http://www.transparencia.sp.gov.br/Home/Conselho>

¹⁹My direct translation from Portuguese from the ENCCLA official website. See <http://enccla.camara.leg.br/quem-somos>

citizens capable of exercising control of public resources and services alongside our team.”²⁰ FS team collects government data via access to information requests and “curate” the documents and datasets they receive. They then send a newsletter twice a month to diffuse the datasets and files. They claim to have more than 5,000 readers of their newsletter and estimate that the vast majority are journalists.

- **FoccoSP** – Acronym for *The São Paulo Forum for the Fight against Corruption and Money Laundering*. Composed of 33 institutions from the three spheres of power [Executive, Legislative and Judiciary branches] that meet periodically to establish priority actions and develop them through working groups. It was created in 2013 “to foster dialogue and implement actions to strengthen the state in the fight against corruption and money laundering”²¹.

- **Fórum de Direito de Acesso a Informações Públicas**. Portuguese for *The Right to Access Public Information Forum*. It is “a coalition of civil society entities, media organizations, and researchers dedicated to making social control of the implementation of the Access to Information Law (Law 12.527/2011)”.

After identifying these communities and organizations, I selected accountability professionals who simultaneously fell into at least two of three categories below:

- work/have recently worked (less than 10 years) for accountability institutions based in São Paulo City (preferably but not necessarily in the four organizations mentioned above);
- have held leadership positions in their institutions; and
- have represented their respective institutions in at least one of the three “communities of practices” mentioned above.

The 20 accountability professionals interviewed are described in terms of gender and type of accountability activities as follows:

- Eight female, and twelve male professionals.
- Nine interviewees work or have worked for social accountability, non-for-profit institutions whose mission is transparency and anticorruption (“diagonal accountability”),

²⁰Translated from the official website: <https://fiquemsabendo.com.br/quem-somos-contato/>

²¹My direct translation from Portuguese from the FoccoSP official website. See <https://www.foccosp.org/apresenta%C3%A7%C3%A3o>

while eleven work for public sector accountability institutions (Court of Auditors, Prosecutor's Offices, Office of the Comptroller General, City Council).

Eighteen of the 20 informants have posed no opposition to disclosing their names; however, for safety and security purposes, the option was not disclosing names to preserve the two professionals who have asked for anonymity. The list below brings the background of the 20 professionals consulted. In average they hold 20,4 years of experience²². Each professional is identified by a code. The forum/organization column informs the communities of practices and institutions they are involved in.

Table 12 – List of accountability actors interviewed (encoded)

Code	Forum/Organization	Experience (# years)
G01	ENCCLA, CGM-SP (former CGM-SP Deputy Head, and former Head of the Information and Intelligence Production Team at CGM-SP)	21
G02	ENCCLA, FoccoSP, CGM-SP (former Head), CGE-MG (former Head), CGU (former national Secretary of Transparency)	23
G03	ENCCLA, FoccoSP, CGM-SP (former Head), MPSP, TJSP	33
G04	FoccoSP (former Executive Secretary), MPSP	20
G05	ENCCLA, MPF (former Head of Car Wash task force São Paulo's unit)	30
G06	FoccoSP, MPSP (Head of the Execution unit — multidisciplinary team)	20
G07	ENCCLA, FoccoSP, CGM (former Head), CGE-SP (former Head)	28
G08	ENCCLA, FoccoSP, CGM (former Head)	17
G09	Municipal Legislative House (former São Paulo councillor, former Speaker of São Paulo City Council)	30
G10	FoccoSP, CGM (former Head), CGU	21
G11	Former Municipal Legislative House (former Head of the Research Unit)	21
NG1	Investigative journalist, Fiquem Sabendo (co-founder)	18
NG2	Forum, Abraji (former President) Data-driven investigative journalist	17
NG3	Forum, CGM-SP (formerly), open data CSO (Executive Director)	14
NG4	Forum, data-driven investigative journalist, Abraji former President	16
NG5	Forum, Abraji (former President), data-driven investigative journalist	28
NG6	Forum, (former Executive Secretary), anticorruption CSO (program manager)	14
NG7	Data-driven investigative journalist, Fiquem Sabendo (co-founder), Abraji (Director)	9
NG8	Forum, Abraji (former President) open data CSO (former Executive Director)	13
NG9	ENCCLA, Forum, anticorruption CSO (former Executive Director)	15

²²As interviews occurred in 2022, the number of years of experience informed refers to 2022 minus the year they started working with the topics of this thesis. The interviewees informed the year of start during the interviews.

All informant but four fell into the three categories simultaneously. Two exceptions are the ones who have held leadership positions at the City Council, the Legislative branch of São Paulo. And the only category they fall out is the one related to communities of practice. There is an easy explanation for that: the City Council is not a member of the anticorruption communities of practice mapped, which is a symptom of how the local Legislative power have relinquished their oversight attribution. The two FS members selected (NG1 and NG7), investigative journalists who cofounded the institution, have not been part of communities of practice in their personal capacity, but one of their organization peers have been active in the Forum.

I conducted the interviews with the 20 accountability professionals selected between August and September 2022. I prepared two interview guidelines for state accountability actors (Annex 1) and diagonal accountability actors (Annex 2). The interviews occurred remotely via videocall software (Zoom). All the interviews have been recorded in video (camera open) and audio. In total, the recordings sum almost 16 hours (959 minutes). On average, each interview lasted 48 minutes. I transcribed all audio files into text.

After the interviews, during the last quarter of 2022, I consulted two additional accountability actors – one from a diagonal accountability institution and one from a horizontal accountability institution – who had been quoted by informants. These two other accountability actors have clarified specific news stories about using the IPTU dataset for anticorruption purposes.

Documents

Prior to the interview data analysis, I had analyzed documents from three sources: 1) municipal decrees, ordinances and administrative processes at the São Paulo City Hall as well as press releases; 2) administrative and judicial processes in local accountability institutions (Audit Courts and Public Prosecutors' offices), and City Council; 3) Civil society organizations and media report.

Two sets of documents from the São Paulo City Hall have been particularly relevant: i) infra-legal norms, that is, municipal decrees and ordinances related to the disclosure of the IPTU dataset, and ii) press releases issued by CGM-SP.

From the infra-legal norms, it is reasonable to infer the political motivations and to identify the legal reasons behind the disclosure of the IPTU dataset, particularly between 2014 and 16; later, from 2017, these norms present the arguments for removing the names of the

owners from the dataset, while bringing arguments for keeping the rest of the data publicly available. Finally, from the press releases issued between 2014 and 2016, it is possible to specify the anticorruption body (CGM-SP) point of view during Fernando Haddad's term.

TCM-SP (Municipal Court of Auditors) as well as MPF and MPE (Federal and State Prosecutors' Offices, respectively) website have been visited so as to search for administrative and judicial processes related to the IPTU dataset. Likewise, the City Council's official website has been explored to search for publicly available information about the accountability function of the Legislative related to anticorruption activities and possible uses of the IPTU dataset.

I also collected news stories and searched for civil society organizations reports using the IPTU dataset. I searched news pieces in Folha de São Paulo printed version using simultaneously the words "IPTU" and "Sao Paulo" and published Between March 2016 and March 2022. Folha de São Paulo is the widest distributed newspaper in São Paulo. I also looked for reports mentioning "IPTU" in the two most relevant anticorruption CSOs based in São Paulo: Transparência Brasil and Transparency International - Brazil,

Additional sources

I triangulate the interviews and document analysis with handwritten notes and email exchanges that I kept during the period when I worked at the São Paulo City Hall (2013-2015) and the following period (2015-2019) in which I acted, as a consultant for the Transparency International organization, as an interlocutor for the São Paulo City Hall in discussions about data openness, including the IPTU database. These notes and emails have been used to select the informants, compose the interview guides and analyze the interview data.

I had the opportunity, for example, to attend a meeting with the then São Paulo Mayor, Fernando Haddad, and the Comptroller of the Municipality on the 2015's International Anticorruption Day (December 9th), exactly the date in which the mayor himself announced the anticorruption package with four initiatives, being one of them the disclosure of the IPTU dataset.

Another opportunity I had to observe the context was in 2019 during the Abraji's International Conference of Investigative Journalism. I was invited to deliver a presentation and my lecture was entitled "Data-driven journalism in major corruption investigations". One of the slides of my presentation listed the IPTU database as one relevant source of information for countering corruption.

I have also attended plenary meetings at ENCCLA events, as observer, while working for Transparency International. ENCCLA plenary meetings are held once a year, in different locations across Brazil. I was present in four occasions.²³

3.4 Procedures for Data Analysis

Grounded theory (Strauss & Corbin, 1998) brings an empirical approach to studying social life through qualitative research and analysis. It is grounded in social construction in the sense that the only realities possible are those constructed through a shared language. Research using such approaches is usually committed to representing those studied on their terms and multiple realities.

This thesis uses a mixed-methods approach based on existing theory and data emerging from the field, the context analyzed, including document and interview data analyses.

The first step during data analysis was to search documents – news stories, NGO reports, audit reports, judicial and sanctioning administrative processes – published by the social and institutional accountability institutions.

Confirming what I had observed as someone familiar with this accountability ecosystem, very few documents explicitly mentioned the IPTU database as a relevant source of data for investigative news stories or administrative and judicial processes on corruption or money laundering.

The second step comprised the interviews and their transcriptions. For transcribing, I used the Google tool Pinpoint²⁴, that provides an automatic raw written transcription of the audio file. While correcting errors of the automatic transcription, I had my first contact with the interview data.

In order to guide my interview data analysis, I followed the coding process inspired on grounded theory (Strauss & Corbin, 1998). I conducted data analysis using the abductive approach, which is a mixed approach, more appropriate when the theory is incomplete or outdated (Small, 2011; Timmermans & Tavory, 2012). As previously stated, the scientific production on open data and corruption control using historical institutionalism lens is still scarce.

²³Representing Transparency International, I attended ENCCLA's plenary meetings in 2015 (in Fortaleza, a North-eastern coastal city in Brazil), 2016 (Natal, also a North-eastern coastal city), 2017 (another North-eastern city), and Foz do Iguaçu (a Southern city, bordering Argentina and Paraguay)

²⁴See <https://journaliststudio.google.com/pinpoint>

“The interrelationships between society and the individual, history, and biography”

In a classic 1959 text, Wright Mills advocates the “sociological imagination”: a quality of the mind necessary to understand society and social and personal phenomena.

For Mills, the sociological imagination is a prerequisite for those who seek to make sense of life and the world, for it provides the ability to understand the interrelationships between society and the individual, history, and biography (1959, p. 13)

In that sense, I opted not to use specific qualitative software tools, as I wanted to “dive” into the data freely.

After the full transcription, I applied two cycles of coding. In the first cycle it is advisable to sustain an “open mind”, let the data “speak”, paying attention to what emerges from the data, not what the researcher is looking for. In other words, let creativity and serendipity enter the process and adapt the additional sampling according to what emerges.

I then created a list of definitive and sensitizing concepts. According to Bowen (2006), definitive concepts are “deductive” concepts that refer more precisely to what is common to a class of objects, with the aid of a clear definition in terms of attributes. On the other hand, a sensitizing concept lacks such specification of attributes or benchmarks and, consequently, does not enable the researcher to move directly to the instance and its relevant content. Instead, it gives a general sense of reference and guidance in approaching empirical instances. For Bowen, “whereas definitive concepts provide prescriptions of what to see, sensitizing concepts merely suggest directions along which to look (2006, p. 7).

The deductive codes, taken from the analytical lenses adopted for the thesis, were the following:

- access to corruption-related data/information
- ability to interpret data/information
- capacity to challenge corruption practices – technical capacity
- capacity to challenge corruption practices – resources
- capacity to challenge corruption practices – independency and autonomy

The sensitive codes, on the other hand, were these:

- Legal gaps
- Institutional leadership
- Lack of incentives
- Perception of risks
- Perception of high costs
- Interested networks

For the second cycle, I made an effort to observe the patterns emerging: similarities, differences, frequencies, correspondences. While coming across different codes that described similar data, I chose the one that best defined the “essence”, but I avoided missing variability. During this second cycle, deductive and sensitive codes were adjusted or replaced.

I coded the interviews from the most experienced informant to the least. After the analysis of the 18th interview, no new code had been created, reaching data saturation.

Data saturation entails bringing new participants continually into the study until the data set is complete, as indicated by data replication or redundancy. In other words, saturation is reached when the researcher gathers data to the point of diminishing returns, when nothing new is being added. (Bowen, 2008, p. 140)

To organize and codify the interview data, I used “data tables”, where the quote and the codes have been registered. According to Cloutier and Ravasi (2021), tables “are practical for organizing and keeping track of qualitative data; efficient for sorting data in a way that facilitates comparisons and the noticing of patterns; and effective at presenting findings in a clear and convincing way” (2021, p. 127).

3.5 Validation Criteria

I sought to meet the criteria proposed by Lincoln and Guba (1985): confirmability, credibility, reliability, and transferability. As suggested by the authors, I sought, in addition to triangulation, in-depth involvement in data collection and analysis, checking each piece of information and returning to the field whenever necessary. In addition, I used interview protocols to increase the reliability and credibility of the research. To ensure confirmability, I sought to provide evidence to corroborate the findings, connecting each analyzed excerpt with the meanings of the interpretation.

4 PRESENTATION OF THE RESULTS

This section presents the thesis results from the triangulation document analysis, interview data analysis, and direct experience.

From the literature review that pointed to the centrality of communities of practices and my direct experience and engagement with the thesis topics in São Paulo, I selected “fora of actors” practices and pertinent accountability institutions and knowledgeable informants.

Following the research question regarding data use by accountability actors in São Paulo, the section starts with the presentation of relevant communities of practices. It is essential to describe these “fora of actors” so as to identify accountability institutions and accountability actors, the professionals working for these organizations, and involved in these “communities of practices.”

After presenting the relevant communities of practice, I also list relevant accountability organizations, both from public and non-profit sectors (horizontal and diagonal), based in São Paulo. The knowledgeable informants will mainly come from the communities of practice and the organizations selected. After presenting the communities and the institutions, I proceed to an analysis of the documents found in these institutions’ official websites. The search covered from March 2016 to March 2022, the period in which the IPTU database was fully open.

Next, I present the findings from the interview data analysis. I start by focusing on diagonal accountability actors, notably journalists and civil society organizations, offering their views and perceptions about the challenges of accessing and using data. I then go to institutional accountability actors and their perceptions around challenges regarding data use and the outcomes of such uses on corruption control. The section closes with a description of the “anticorruption data dynamics” in terms of data quality, the uses and barriers for reaching their expected and desired outcomes, including the failures in accountability, which in turn contributes to corruption persistence.

4.1 The Transparency and Anticorruption “Communities of Practice” in São Paulo

Given the centrality of “forum of actors” for this thesis topic, the data collection's starting point was to list the communities of practices of accountability in São Paulo as well as the most relevant institutions around them. From my direct experience, handwritten notes and email exchanges, I initially identified four communities of accountability practices and validated them with the interviewees during contacts previous to the recorded conversation.

However, according to multiple accountability actors, one community previously mapped — Sao Paulo's Transparency Council — should not be considered. Firstly, it does not cover São Paulo City (only state level), and secondly, it was almost inactive during the period of interest for this thesis (2016-2022).

In the specific case of the Transparency Council of... here in the state of São Paulo, I joined in 2016, and it... it is not a council that has binding power, you know? So, it makes recommendations, debates, things like that. So, this already empties its power a little bit. And it had parity [note from the thesis' author: parity between civil society and government in the distribution of seats for the Transparency Council], and then in the wake of those scandals, the governor, who was Alckmin [note from the thesis' author: "Alckmin" is Geraldo Alckmin, who was Sao Paulo state governor for 12 years in two periods: 2001-2006 and 2011-2018 and became vice-president of Brazil in January 2023], removed the parity and reduced it, he put more members from the government and fewer from civil society, so it became unequal; this already gave a loss of relevance to the council and, as time went by, the subsequent governments that came in gave it less and less importance. And then you see this in the people who are appointed to chair the council, for example, you know? Who are the people from the governmental portfolios that are appointed to participate, that don't have an active participation, they are not people concerned with LAI, with transparency, they are people that... so... People who are from the area, you know? So, you put a person there who has never worked with the issue of transparency, what is she doing presiding over transparency, if she has never worked with it? So, in the case of the Transparency Council, it has been losing its importance over time, you know? I think there were some positive moments when we had some debates that even generated policy recommendations that were partially implemented, for example, in the area of public security, in the area of transparency so that we could, I don't know, not have so many setbacks... it was a time when criminal occurrence reports were made available, then they made some changes in the active transparency website, the availability of requests for access to information... these were all agendas that... that arose in the Transparency council, they were all agendas that... that arose in the active transparency website, that came up in the Transparency council, but in the last two years it was really bad, and we couldn't do anything. To the point of, like, having a meeting and saying: "we have no agenda to discuss today, so I'm proposing to cancel the meeting". How can we cancel

the meeting? There is nothing to discuss regarding transparency?!? The State Government of São Paulo is not suggesting any agenda? It is the civil society that has to suggest an agenda, because for the government there is nothing to discuss about transparency! So, this was unimaginable before. The presidency of the Council is always relatively: "Ah, let's invite someone to make a presentation, so we can discuss such and such a point, such and such a policy". So, you had initiatives from people who were trying to do something. But then it started emptying and then the Transparency Council of São Paulo got really, really bad in the last two years, a year and a half. (Informant NG9)

In that sense, I disregarded Conselho de Transparência, and considered 1) ENCCLA, 2) FoccoSP, and 3) Fórum de Acesso a Informação Pública as relevant “communities of practice.”

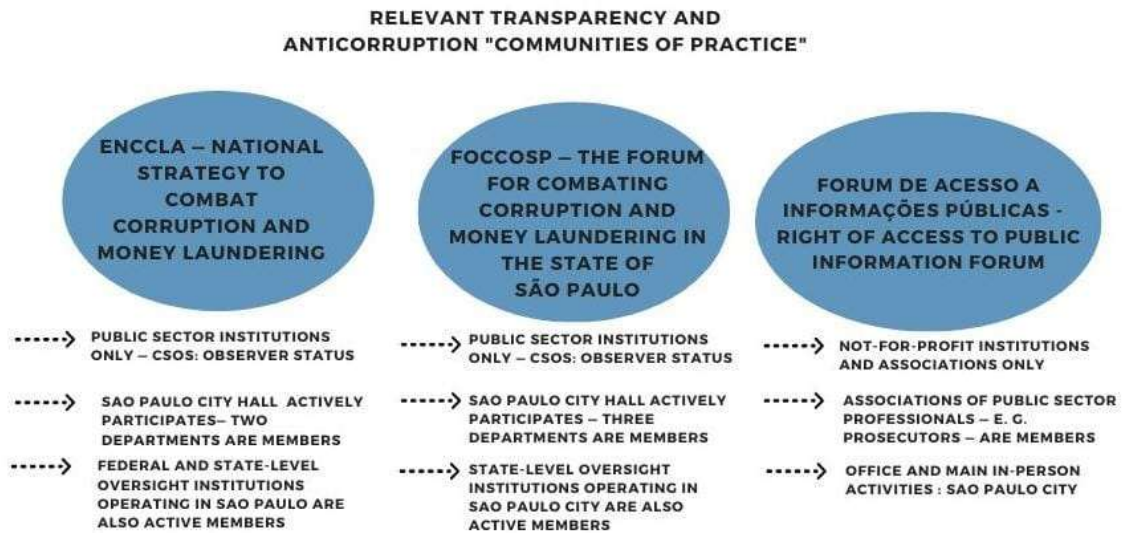
Adding to the communities of practice, I also identified four organizations in São Paulo’s anticorruption and open data ecosystem that are relevant to these discussions. Three are active participants of the “communities of practice” mapped. The only exception is the City Council. This is because the municipal Legislative branch has, in practice, outsourced its oversight role, as noted from document analysis and the interviews.

The four institutions are both from public and not-for-profit sectors and are based in São Paulo City. On the one hand, two crucial city-level public sector organizations that have accountability as one of their key attributions, namely the City Council and the Office of the Comptroller General (CGM-SP). On the other hand, two journalists’ associations, Abraji and FS — both entities constantly advocate for transparency and open government data and promote data use among journalists. As mentioned above, the City Council is not involved in communities of practices as an institution. Conversely, CGM-SP participates in both public-sector fora. Regarding the Civil Society Organizations (CSOs), Abraji coordinates the activities of the Forum, and FS is a member of the Forum.

Communities of practice

The three communities mapped include organizations that are based in São Paulo City but are not exclusively focused on the local level. One operates at the São Paulo state and city levels – FoccoSP – and the others (ENCCLA and Forum) are national fora with federal, state, and municipal participants.

Figure 4 – National and state level anticorruption “communities of practice” relevant to the context of São Paulo City



Source: Elaborated by the author of the thesis

Notably, there is no city-level community of practice on transparency and anticorruption. Although São Paulo City is a member of the Open Government Partnership initiative, no informant interviewed sees OGP São Paulo's activities as relevant to the anticorruption agenda. From my direct experience as a former civil servant at the City Hall, from document analysis, and the testimony of an informant more knowledgeable about the OGP São Paulo local program, anticorruption and data use only marginally were part of OGP São Paulo's goals.

The first two OGP action plans covered four years of the period of interest for this thesis, from 2016 to 2020. The third plan started in 2021 and is due to end in 2024.

The first OGP action plan (2016-2017) was more structural-oriented. Its five commitments, which aimed at establishing the grounds for civic participation, training (on topics related to open government), communication, institutionalization, and innovation, did not mention the word “corruption”.

The second plan (2018-2020) also brought five commitments, being of them “fight against corruption”. This commitment is read as follows: “Guarantee the accessibility of public data in an open format, through integration and betterment of the information made available on contracts, bids and budgetary/financial execution of the city of São Paulo, improving existing mechanisms, enabling better conditions for citizen monitoring of bidding processes”.

Finally, the third plan presents four commitments, one of which is related to data on bids and contracts.

As noted, there is no mention of the IPTU dataset.

Besides, the only organization dedicated to anticorruption decided to withdraw from the São Paulo OGP steering group. Its representative gave their reasons below.

We participated in the first... the first and the second Open Government action plans here at the São Paulo City Hall. [note from the thesis' author: the first action plan was created during Haddad's term and implemented during Doria's term, while the second plan was both elaborated and implemented during Covas's term, after Doria's resignation in April 2018]

I think that the City of São Paulo has always stood out for encouraging a lot of civic participation; this is even recognized in the OGP evaluations, but the difficulties we had were that, on the one hand, the first plan was made at the end of Haddad's administration, and it was Doria who was going to have to implement it.

Although there were commitments that were published in the Diário Oficial [official gazette], so it was an official commitment from the City Hall, and there were... there were ordinances, etc. decrees assuming the commitments, there was no political will to execute them, you know? So, things that were seen as "this one is very petista" ["petista" is the Portuguese term for PT — the Workers Party — supporters or affiliated; Fernando Haddad is from PT], they didn't.... they didn't make much effort to execute. "Ah, let's modify the commitment because if we can't do it, we won't have money, I don't know what, and so on"; so, we had many battles in this sense, to try to make the original commitments be fulfilled. Because in some cases there wasn't much interest, like a participative budget thing. It wasn't exactly that, but something along those lines, so they didn't want to do it: "Oh no, this is too petista" [Portuguese term for PT — the Workers Party — supporters or affiliated].

So, there were these difficulties. And then, when the Doria administration came in, it was already the second plan, the Municipal Office of the Comptroller General was downgraded, you know, which took care of this part a lot, to be part of SMIT, the Department of Innovation, so it was a sub-area there. And this was also reflected in less support to make the commitments, to make things happen on the part of the City Hall. (...). So, even for that reason, given the difficulties we were seeing, we thought that the effort was not worth it, and we, Transparência Brasil, did not continue with the

subsequent plans. “We are going to leave the municipal level because it is giving too much work for little result in terms of progress.”, you know, and then we did not follow up on what happened in the following years (Informant NG9).

As noted, the disclosure of the IPTU data has not been mentioned in the action plans, and anticorruption has not been a central part of the São Paulo Open Government program. Likewise, anticorruption actors, be it in civil society or formal oversight institutions, have not been involved in OGP in São Paulo City²⁵. Therefore, actors engaged in OGP are not knowledgeable on the topic of this thesis; hence they have not been consulted.

On the other hand, the three communities of practice taken as the first step for locating the accountability actors have transparency and anticorruption as top priorities. ENCCLA and Focco-SP bring the word “corruption” into their description. The Forum has the topic of access to information as a critical priority.

ENCCLA is regarded as a relevant community not only because its primary goal is the fight against corruption and money laundering but also because the federal government has led it, and it has been active since 2013. ENCCLA plenary meets once a year, at the end of each year, and sets annual goals. Many of these goals are related to transparency and data use. Hence, one can say that this is a forum for promoting data use and sharing among public sector institutions — but in general, exclusively among government agencies. That is so because only government agencies are part of ENCCLA. As we will discuss later, horizontal accountability institutions are more comfortable accessing data directly from other government agencies (via data-sharing systems and formal agreements) rather than collecting data from open sources — even when these sources are official websites.

Two São Paulo City Hall organizations are ENCCLA members: CGM-SP and PGM-SP (Procuradoria Geral do Município - Municipality's Attorney General's Office).

FoccoSP, like ENCCLA, is a 100% public sector coalition. Similar to ENCCLA, FoccoSP promotes data sharing among state accountability institutions. CGM and PGM are members, as well as the Department of Finance.

²⁵According to the 2022 “Open São Paulo” annual report, the following CSOs are members of the steering committee or “shared management forum” of the OGP São Paulo: Colaboratório de Desenvolvimento e Participação - COLAB-USP; Instituto Prospectiva – Inspro; Instituto Cidades Sustentáveis; Observatório Social do Brasil; Rede pela Transparência e Participação Social; Coletivo Ocupa Mãe; Delibera Brasil; Fundação Álvares Penteado. **None of these institutions participate in the three relevant transparency and anticorruption “communities of practices” identified.** See https://www.prefeitura.sp.gov.br/cidade/secretarias/casa_civil/relacoes_institucionais/coordenadoria_de_governo_aberto/index.php/

Forum is a coalition comprising 30 civil society organizations. Abraji coordinates the Forum, and Fiquem Sabendo is a recent associate. It is regarded as a relevant community because it was necessary for pushing the Brazilian access to information (ATI) law in the 2000s. Since 2011, when Brazil's ATI law was enacted, Forum has been the leading actor advocating for adequate implementation of this norm.

After selecting these three national and state levels “fora of actors” and the four accountability institutions based in the city of São Paulo, I looked for accountability actors, both from public sector institutions and from civil society.

I then identified and interviewed 20 professionals, being nine from non-state institutions – those performing “diagonal accountability” – and eleven from public sector institutions; hence performing “horizontal accountability”.

But before presenting the results from the interviews, the document analysis follows. As previously noted, CGM-SP, the internal oversight body of the São Paulo City Hall, has had a central role in opening data. The agency, created in 2013, was particularly active during the first years of operation and before being downgraded.

The City Council, on the other hand, seems to have passed entirely on the oversight attribution to the Municipal Court of Auditors (TCM-SP), a government agency responsible for auditing the municipal budget, and that is a technical arm of the local Legislative House.

Federal and state prosecutors' offices – MPF and MPSP, respectively – could also have used the IPTU dataset. For example, money laundering is a federal offense; MPF could use the available data to investigate this crime. In principle, MPSP can also use the urban land and property data to explore, for example, illicit enrichment of state-level civil servants.

A presentation of relevant documents related to the topic of this thesis appears below.

4.2 Horizontal Accountability Institutions: What They Proactively Inform About Data Use

I analyzed documents from three sets of institutions. Firstly, I looked at ordinances and administrative processes at the São Paulo City Hall as well as press releases. Second: administrative and judicial processes in the following horizontal accountability institutions: the Municipal Court of Auditors (TCM-SP), the City Council (Legislative branch) and the Public Prosecutors' offices, both federal and state levels (MPF/SP and MPSP; 3) Civil society organizations and media report.

TCM-SP as well as MPF and MPSP (Federal and State Prosecutors' Offices, respectively) website have been visited so as to search for administrative and judicial processes

related to the IPTU dataset. Six keywords²⁶ have been used to search for documents: “IPTU”, “tax”, “laundering”, “corruption”, “enrichment” and “improbability”.

Likewise, the City Council's official website has been explored to search for publicly available information about the accountability function of the Legislative related to anticorruption activities and possible uses of the IPTU dataset. In that sense, besides the six keywords, I also looked for parliamentary inquiries.

I also collected news stories and searched for civil society organizations' reports using the IPTU dataset.

Office of the Comptroller General (CGM-SP)

CGM-SP was created in 2013, the first year of Haddad's term as mayor. In this first year, because of data analysis and other activities, CGM found inconsistency in assets owned by civil servants and broke the “Audit Mafia scandal”.

CGM-SP was also crucial for pushing the city's transparency and open government agenda. For example, in 2013, CGM-SP led the redaction of the municipal decree aimed at operationalizing, at the local government, the national access to information law.

When the 2013 massive street protests erupted in Brazil, following my suggestions to the then Comptroller General, CGM-SP and other departments started meetings and consultations inside and outside local government to engage in the international open government agenda more formally. As a result, a decree²⁷ was signed by mayor Fernando Haddad in January 2014, establishing the program “São Paulo Aberta” (“Open São Paulo”). Less than two years later, another decree²⁸, on 9 December 2015, authorized the public disclosure of the IPTU dataset.

The CGM-SP press release²⁹ on that date reads as follows:

City Hall announces transparency of the real estate registry in the city

The City's Open Data Portal, a Municipal Policy for the Protection and Defense of Users, and an Integrity Program for companies were also launched by the Mayor;

²⁶In Portuguese: “IPTU”, “imposto”, “lavagem”, “corrupção”, “enriquecimento”, “improbidade”

²⁷Municipal Decree N° 54.794 (28 January 2014)

²⁸Municipal Decree N° 56.701 (9 December 2015)

²⁹See https://www.prefeitura.sp.gov.br/cidade/secretarias/controladoria_geral/noticias/?p=208146

measures are part of the celebrations for the International Day Against Corruption (12/9)

In an unprecedented initiative in Brazil, the São Paulo City Hall announced on 12/9 the disclosure of data on the ownership and address of properties in the city. (...)

Figure 5 – Reproduction of the official CGM-SP webpage announcing the disclosure of the IPTU dataset as an anticorruption initiative

The screenshot shows the official website of the Controladoria Geral do Município (CGM) of São Paulo. At the top left is the logo of the City of São Paulo and the CGM. A navigation menu includes 'Serviços', 'Mapa de Serviços', 'Acessibilidade', and 'Legislação'. Below the menu is a search bar with the placeholder 'Palavra-chave' and a red 'Pesquisar' button. The main content area features a news item titled 'Prefeitura anuncia transparência do registro de imóveis na cidade' dated 17:07 09/12/2015. A banner for dengue prevention is also present.

Later, in May 2016, marking CGM-SP third anniversary, the institution published a document³⁰ entitled “Office of the Comptroller General in Cases - innovative experiences in fighting corruption and promoting integrity in the city of São Paulo” with 17 activities. One of which was the disclosure of the real estate registry.

The initiative is described as follows:

“(…)

Despite being information considered of a public nature, the ownership of real estate was, until now difficult to access. To obtain the records, it was necessary to request a certificate from a notary, against payment. With the opening of the cadastral base, the information is now accessible online and free of charge.

The initiative is part of the context of policies and actions to combat corruption implemented since the creation of the Office of the Comptroller General of the Municipality in 2013. With the investigations related to the corruption case involving Tax on Services (ISS) inspectors, evidence of the existence of a fraud scheme involving the Property and Urban

³⁰See pages 10 and 11 of the document, that can be found at http://www.prefeitura.sp.gov.br/cidade/secretarias/upload/controladoria_geral/arquivos/CC_Final2.pdf

Territory Tax (IPTU) was discovered. The investigations revealed that owners of large enterprises, such as shopping centers and universities, paid bribes to inspectors to reduce the areas declared for the calculation of the IPTU.

After the discovery of the corruption scheme involving the property tax, the municipal administration took the initiative to increase transparency took the initiative to increase transparency and social and social control of the information related to this t, also aiming to improve the integrity of the data and integrity of the data and its management. (...)"

The document also states that other departments of the City Hall were “fundamental” for the disclosure of the dataset and mentions three actors: Municipal Department of Urban Development, Municipal Department of Legal Affairs, and the Municipal Department for Finance.

In addition, the “Cases” brings a box that reads the following:

Methodology

To provide legal security to the initiative of data publication the Executive Decree 56,701/2015, a regulatory framework that defines the parameters for the availability of information of the real estate registry, was elaborated. Legal opinions of the Attorney General of the Municipality from 2003 and 2015 also supported the decision to open the dataset because they brought the understanding that the real estate registration information does not have tax secrecy.

2016 was the last year of Fernando Haddad’s term. As mentioned above, he ran for reelection but was defeated by João Doria, who won in the first round. Doria obtained 53% of the votes, whereas Fernando Haddad got 16,7% in the elections held in October 2016.

João Doria took office as São Paulo mayor in January 2017. Elected with promises about austerity and reducing public expenditures, the new mayor allocated CGM-SP as a sector inside the Legal Affairs Department, so the Comptroller would no longer report to the mayor. In 2018 Doria resign his mayoral position to run for governor of São Paulo state. He won the election. Bruno Covas elected vice-mayor in the 2016 local elections became the mayor of São Paulo.

On the legal framework front, in the same year of 2018, Brazil enacted its privacy law (LGPD – Lei Geral de Proteção de Dados) primarily based on the European Union's GDPR (General Data Protection Regulation).

In September 2020, in order to comply with the national legislation, the São Paulo City Hall edited a decree³¹, regulating the application of said law within the scope of the direct and indirect Municipal Administration. As stipulated in the Decree, the CGM-SP oversees Privacy Data Protection.

One month later, in October 2020, PGM-SP published a legal opinion³² recommending the removal of the name of real estate and urban land owners from the public version of the dataset.

Some excerpts of the legal opinion clarify the views of the new political leadership of São Paulo City Hall in terms of the usability of the IPTU dataset (also known as CIF – Cadastro Imobiliário Fiscal, the Portuguese for *Fiscal Real Estate Registry*).

“The information (...) are subsidies for the technical work of the Municipality (...). It is also used for research in the area of urban planning by universities and educational institutions and also by the private sector, which uses this content for its professional activities (architects, engineers, real estate market, civil construction, civil law).

It is by accessing the fiscal cadaster, via GeoSampa, that citizens can consult, formally and free of charge, information about a certain lot associated with other non-tax data such as zoning, public areas, equipment location, road system, among others, since the system allows the integration of inter-departmental information. It is worth mentioning that municipal public services such as feasibility studies for business start-ups (Empreenda Fácil), electronic licensing and zoning studies already include in their user instructions the search of registration data in GeoSampa”.

The excerpt above seems to demonstrate the intended uses of the dataset: policy analysis (“by universities and educational institutes”) and business opportunities (“by the private sector”).

Time to recall the 2021 World Bank publication “Data for Better Lives”, mentioned above. The WB study mentioned three intended outcomes from the use of data: “greater accountability”; “better policy making and service delivery”; and “increased business

³¹Municipal Decree n. 59.767 (16 September 2020)

³²Legal Opinion PGM-SP n. 12.195 (5 October 2020)

opportunities. The PGM-SP legal opinion, as the excerpt above illustrate, listed the last two positive outcomes, but avoided the first.

In conclusion, the legal opinion saw no “public interest” in disclosing the names of owners, “whether natural or legal person” – thus extending the privacy protection to legal persons:

It is true that the LGPD was explicit in allowing the processing of personal data by the Administration to the extent necessary to achieve the public interest; in this case, however, there is no discernible public interest in the display by GeoSampa of personal data contained in the CIF.

Thus, (...) one must conclude for the possibility of disclosure through GeoSampa of data contained in the Fiscal Real Estate Registry (CIF), except for those that allow the identification of the owner of the property or taxpayer of the respective IPTU, whether natural or legal person."

In 2021, the City Hall decided to withdraw property owners' names. As the executive secretariat of the interdepartmental committee of access to information and simultaneously the agency in charge of privacy data protection, CGMSP agrees to, without mentioning the words “accountability” or “corruption”.

In January 2022, the IT department of the City Hall informed that the dataset, without the columns “name” and “id tax number”, is ready to be uploaded to GeoSampa, as shown below.

Figure 6 – Reproduction of an official document from the São Paulo City Hall’s IT Department (translation below)

São Paulo, 13 de janeiro de 2022.

SF/COTEC

Sr. Coordenador,

Informamos que a demanda D16035 foi executada com sucesso pela Prodam, conforme 057404901. Essa demanda corresponde à geração de novos arquivos extraídos do sistema Novo IPTU, sem os dados de sigilo, para atualização do GeoSampa. Os arquivos foram devidamente validados e disponibilizados por SF/SUREM /DECAD, conforme 057239970.

Desta maneira, resta pendente a atualização do sistema GeoSampa, de responsabilidade de SMUL.

Atenciosamente,

The image above reads, in English, as follows:

Mr. Coordinator,

We inform you that demand D16035 was successfully executed by Prodam [note: Prodam is the IT department of the local government], according to 057404901. This demand corresponds to the generation of new files extracted from the Novo IPTU system, without the secrecy data, to update the GeoSampa. The files were duly validated and made available by SF/SUREM/DECAD, according to 057239970.

Thus, what is still pending is solely the update of the GeoSampa system, which is SMUL's responsibility [“SMUL” is the Municipal Department of Urban Development]

I could not locate the exact data in which the digital file without the names of property owners was eventually published at the City Hall’s GeoSampa portal.

In a consultation via email³³, a top civil servant from the Municipal Department of Urban Development responded, “first quarter of 2022”.

Thus, at a certain point between mid-January and late March 2022, the IPTU dataset was no longer fully available. The dataset was kept online, but the names of legal and natural persons were no longer publicly available.

CGM-SP – Auditing

The São Paulo Office of the Comptroller General (CGM-SP) is divided into four subareas, and one of them is the auditing sector (AUDI). On their section³⁴ of the CGM-SP official webpage, it can be read:

The General Audit Coordination (AUDI) represents the macro-function responsible for verifying the legality and legitimacy of administrative acts and facts and evaluating the results achieved, regarding the aspects of efficiency, effectiveness and economy of the budgetary, financial, patrimonial, operational, accounting and finalist management of the direct and indirect administration units and of private law entities that receive resources from the Municipality of São Paulo (Third Sector).

This verification is done through programmed and special audits, based on the issuance of Service Orders (OSs), which are generated from the premises defined in the Annual Internal Audit Activities Plan

³³Email interview in June 2022 with the Technical Division Director of the Information Production and Analysis Coordination (GEOINFO)

³⁴See https://www.prefeitura.sp.gov.br/cidade/secretarias/controladoria_geral/a_cgm/index.php?p=153810

At this section, there is a dedicated webpage for “Audit Reports, Technical Notes and Monitoring Notes”. These documents range from 2013 to 2022. As seen on the webpage³⁵, there is no mention to IPTU.

On 22 December 2022, there were 123 audit reports, 62 Technical Notes and 48 monitoring notes listed on the webpage, along with a summary. None of these 233 documents mentioned the IPTU dataset.

São Paulo City Council and TCM-SP

A search using the initials “IPTU” at the São Paulo City’s Council official website³⁶ brings 80 press notes published between 1 March 2016 and 31 March 2022.

- 41 refer to IPTU taxation exemptions for specific groups of citizens
- 17 refer to the relief or cancellation of IPTU debt
- 14 refer to the increase of IPTU
- 4 refer to IPTU in the perspective of tax justice and transparency

However, the legislative debate on tax transparency is unrelated to the disclosure of the IPTU database. According to the press office, the topic under discussion (on 7 April 2017) was a bill turning mandatory for the municipality to make clear for taxpayers all the component elements of the formation of the final value of the IPTU.

One of the oversight mechanisms for the Legislative branches is a parliamentary inquiry committee. From March 2016 to March 2022, the City Council established 14 committees of this type, but none related to urban land corruption³⁷.

Moreover, unlike the São Paulo state Legislative House, the São Paulo City Council does not host a standing “inspection and control” committee.

At the local Legislative in Sao Paulo, there are seven standing committees³⁸, none exclusively dedicated to oversight. none exclusively dedicated to oversight. On the other hand, the state-level Parliament hosts 19 standing committees, one of which is the Inspection and Control Committee. This committee is "responsible for inspecting the acts of the direct or

³⁵See https://www.prefeitura.sp.gov.br/cidade/secretarias/controladoria_geral/relatorios/index.php?p=172575 (last visited on 18 January 2023)

³⁶<https://www.saopaulo.sp.leg.br>

³⁷The committees created in this period covered the following topics: delivery apps; falsification of goods; health plan (related to Covid-19), violence against LGPTQA+, violence against women, antennae, illegal car parking/valets, migration, the Municipal Theatre, environmental compensations, violence of soccer fans, major text debtors and tax evasion.

³⁸See <https://www.saopaulo.sp.leg.br/atividade-legislativa/comissoes/>

indirect administration of the State and of the concessionary public service companies, under the terms of the pertinent legislation, especially to verify the regularity, efficiency and efficacy of its organs in the fulfillment of institutional objectives, as well as to give an opinion on proposals regarding the Governor's accounts."³⁹

In conclusion, no official information about IPTU dataset being used for corruption control has been found in the City Council website.

The Municipal Court of Auditors (TCM-SP), according to what is read on their website⁴⁰, is responsible for the exercise of “external control, especially preventively and concomitantly, inspecting, judging, and guiding the management of public resources of the Municipality of São Paulo in order to ensure that they are collected and applied in accordance with the principles of legality, legitimacy, and economy, aiming at the improvement of municipal services for the benefit of society”.

The TCM-SP describes itself as “a body of popular representation - as it assists the City Council, elected by direct vote of the citizens. (...). The Court of Auditors of the City of São Paulo is part of the municipal public sector, but is autonomous and has its own budget, which guarantees it freedom to act within the constitutional principles of legality, impartiality and economy.

TCM-SP publishes annual activity reports. However, by December 2022, only one report⁴¹, from the previous year, was publicly available. A summary of the activities of the first quarter of 2022 was also available.

So as to try to locate uses of the IPTU dataset, I opened both reports and used the following keywords as search terms: “IPTU”, “tax”, “laundering”, “corruption”, “enrichment” and “improbability”.

The 2021 annual report, with 78 pages, returned the following results:

- IPTU: zero
- Tax: one mention, about tax collection
- Corruption: five mentions, all citing ENCCLA and FoccoSP meetings that TCM-SP representatives attended
- Laundering: three mentions, all citing ENCCLA and FoccoSP)
- Enrichment: zero

³⁹My translation of the committee webpage: <https://al.sp.gov.br/comissao/?idComissao=8509>

⁴⁰See “missão” (“mission”) area of the official website: <https://portal.tcm.sp.gov.br/Pagina/91> (last visited on 19 December 2022)

⁴¹See “transparency” area of the official website: <https://portal.tcm.sp.gov.br/Pagina/102> (last visited on 19 December 2022)

- Improbability: one mention, citing an event where a debate around the administrative improbity law occurred

At the summary of activities related to the first quarter of 2022 none of the six keywords returned results.

MPSP and MPF

Ministério Público, which is the Brazilian Public Prosecutor's Office, introduces themselves as follows, in their official website⁴²: “an essential function of Justice, for the defense of social and individual rights; the defense of the legal order, and the defense of the democratic system.”

The official website adds “The Brazilian Public Prosecution Service is composed of the Public Prosecution Services in the states (acting before the state courts), and the Public Prosecution Service of the Union (MPU), which, in turn, has four branches: the Federal Public Prosecutors' Office (MPF), the Labor Public Prosecutors' Office (MPT), the Military Public Prosecutors' Office (MPM), and the Public Prosecutors' Office of the Federal District and Territories (MPDFT).

MPSP is the one responsible for acting before the state court of São Paulo. They are responsible for São Paulo state and São Paulo municipalities.

MPF, the federal service, has units in each of the 27 Brazilian federal states. Hence, MPF/SP “is the Federal Public Prosecutor's Office (MPF) unit that handles cases within the jurisdiction of the federal courts in the state of São Paulo. The public prosecutors carry out their functions in civil and criminal lawsuits in federal courts”⁴³.

According to Brazil’s Constitution, the Ministério Público has budgetary and administrative autonomy. The head of the Prosecutors Office is nominated by the Executive branch of Government and confirmed by the Legislative Branch. Therefore, in São Paulo state, the Prosecutor General is nominated by the Governor. The Federal Prosecutor General, on the other hand, is nominated by the president of Brazil and confirmed by the Senate.

As these both Ministérios Públicos, state and federal level, are based in São Paulo City and are members of the communities of practices above-mentioned, searches on their official websites have been done to look for information regarding possible uses of the IPTU dataset.

⁴² See <https://www.mpf.mp.br/o-mpf/sobre-o-mpf/sobre-o-mpf-1>

⁴³ See <https://www.mpf.mp.br/sp/institucional>

Although this is a city-level taxation database, the name of property owners could help state level and federal level prosecutors in cases, for example, regarding state-level civil servants or regarding money laundering, which is a federal crime.

On the MPF website, no entries related to the topics of the thesis have been found using the same keywords mentioned above. On the other hand, MPSP have published documents relevant to the objective of this thesis.

On the MPSP website⁴⁴, 35 entries have been found: 15 press notes and 20 administrative documents unrelated to investigations.

Of the 15 press notes, six are related to the so-called “Audit Tax Mafia”, a corruption scandal broken in 2013 involving São Paulo City Hall civil servants. They have been published in 2017 (one), 2018 (two), 2019 (one), 2020 (one) and 2021 (one). The titles of the six press notes are below:

1. Gedec [*acronym in Portuguese for “Special Group on Economic Crimes”*] denounces nine for involvement in the IPTU mafia / Directors of a shopping mall gave bribes to avoid paying tax (4 May 2017)
2. MPSP denounces representatives of Shopping JK Iguatemi [*a luxury shopping mall*] for payment of bribery / Payments of R\$ 3 million to stay tax free (19 April 2018)
3. "King of tax inspectors" managed assets blocked, according to prosecutors / man arrested continued managing assets that were already blocked by Justice (21 June 2018)
4. Denounced by Gedec, former São Paulo tax inspector is convicted of money laundering / Judiciary imposed 13-year prison sentence (12 December 2019)
5. Gedec denounces seven people for fraud scheme involving supermarket chain / Civil servants received kickbacks to favor company (10 March 2020)
6. Gedec makes agreements not to prosecute and reverses almost R\$ 3.3 million for health / Agreements with companies occurred in the scope of investigations of the Audit Tax Mafia (12 April 2021)

One of the six notes (the one published in 2020 about the supermarket chain) mentions a partnership between MPSP and CGM-SP, the local Office of the Comptroller General:

⁴⁴Searches have been done in December 2022 at the “content” area of the website, filtering from 2016 to 2022: <https://www.mpsp.mp.br/web/guest/busca>

“The scheme was unveiled after an investigation started in 2013 by Gedec in partnership with the Comptroller General of the Municipality of São Paulo, which led to the discovery of a sophisticated criminal organization composed of municipal tax auditors in the Finance Department”⁴⁵

Another note, the one published in 2019 about the former tax inspector convicted, mentions money laundering via real estate properties:

“In relation to real estate acquired (..), Mendroni [the MPSP prosecutor] pointed out that there is no justified origin for the payments of most of the amounts used for the settlement of the units. Furthermore, there were clear divergences in the declarations of values of acquisitions/sales, demonstrating maneuvers of concealment of values for money laundering through these assets”⁴⁶.

Although there is no mention of the IPTU dataset, it is reasonable to hypothesize that either MPSP or CGM-SP used the urban land and real estate property data to investigate the properties purchased by the former tax inspector.

As noted, MPSP collected information from the Audit Tax Mafia scandal and used them on courts. It is then possible to affirm that there have been accountability procedures, including sanctions.

However, by analyzing the documents alone, it remains unclear if the prosecutors have gathered information on real estate property from the publicly available IPTU dataset or by other means. Interview data analysis added to the partnership between MPSP and CGM-SP mentioned above will reinforce the hypothesis that although prosecutors used data on property ownership, these pieces of information were collected from CGM-SP internally, without the necessity of visiting the open source dataset publicly available.

In conclusion, the local Legislative branch, the São Paulo City Council, and its technical arm, TCM-SP, seem to have not investigated real estate and urban land ownership during the period that the IPTU dataset was fully available. On the other hand, MPSP has apparently built cases upon the work developed by CGM-SP.

⁴⁵See <https://www.mpsp.mp.br/w/gedec-denuncia-sete-pessoas-por-esquema-de-fraudes-envolvendo-rede-de-supermercados>

⁴⁶See <https://www.mpsp.mp.br/w/denunciada-pelo-gedec-ex-fiscal-de-s%C3%A3o-paulo-%C3%A9-condenada-por-lavagem-de-dinheiro>

4.3 Civil Society and Media: Greater Accountability, Better Policy Analysis or Increased Business Opportunities?

Relevant media outlets based in São Paulo and at least one Civil Society Organization have used the IPTU dataset as a basis for technical studies and news reports about corruption and money laundering risks. Additionally, the availability of data was important for policy analysis on land uses and for startup enterprises.

Between March 2016 and March 2022, *Folha de São Paulo*, the widest distributed newspaper in São Paulo, wrote 266 news pieces in its printed version using simultaneously the words “IPTU” and “Sao Paulo”. The printed version of the newspaper was analyzed because the most exclusive, robust and investigative news stories are published in paper.

After reading all headlines, it was possible to affirm that the vast majority of the news stories are unrelated to transparency and discusses issues on taxation and the city. Only six news pieces explicitly mention that the reporters used the publicly available IPTU dataset.

Two news stories published in 2016 focus on inequalities. The first⁴⁷ one states that 0,1% of real estate properties in São Paulo, precisely 1.840 units, occupy urban land equivalent to the area of two Ibirapuera parks. Ibirapuera is the largest urban park in São Paulo and the most visited park in South America. The news report adds that residential buildings favoring 107,000 low-income families could be built in such a large urban area. The reporters write that they made the story “from 3,3 million properties of IPTU registry opened by Fernando Haddad’s administration”. The second, a continuation of the first, listed luxury mansions owned by celebrities and politicians.

Another story⁴⁸, published in 2018, focuses on city licenses regarding fire prevention procedures in buildings. A fourth news story⁴⁹, from 2019, informs about the expansion of evangelical churches.

The other two stories, published in 2021, relate more closely to this thesis's topics. One refers to the continuation of a scandal broken two years before, known as the “kindergarten mafia”. According to police investigations and news reports, charities in charge of kindergarten

⁴⁷See “São Paulo mansions occupy area of two Ibirapuera parks” (<https://www1.folha.uol.com.br/cotidiano/2016/07/1794827-mansoes-de-sao-paulo-ocupam-area-de-dois-parques-ibirapuera.shtml>) 24 July 2016.

⁴⁸See “Almost half of the buildings in SP predate tough fire regulations” (<https://www1.folha.uol.com.br/cotidiano/2018/05/quase-metade-dos-predios-de-sp-sao-anteriores-as-regras-duras-anti-incendio.shtml>) 11 May 2018.

⁴⁹See “São Paulo gains 2,433 new churches in 25 years with evangelical expansion” (<https://www1.folha.uol.com.br/cotidiano/2019/09/sao-paulo-ganha-2433-novas-igrejas-em-25-anos-com-expansao-evangelica.shtml>) 5 September 2019

schools were deviating public money. These schools had formal partnerships with the City Hall, thus receiving public funds. According to the news stories, one company suspected of issuing false invoices to the kindergarten schools transferred money to the then São Paulo mayor, Ricard Nunes⁵⁰. The reporters write that they found the connection between the suspect company and the mayor thanks to the IPTU registry.

The other 2021 news story⁵¹ is related to a health operator investigated during the Covid pandemic. The reporter investigated the company and the buildings they own in Sao Paulo and found that the information they provided to the municipality when requesting fire prevention licenses differed from the data of the IPTU registry.

As noted, journalists of the largest media outlet based in Sao Paulo have occasionally used the IPTU dataset. Four of six news articles, though, referred to analysis unrelated to ethical instances, fraud, corruption, or money laundering. These four news pieces are more closely related to policy analysis rather than greater accountability.

An example of the intense use of the IPTU registry for anticorruption and anti-money laundering purposes was a study I led while working for Transparency International Brazil. Published in 2017, the report, entitled “Does corruption live next door?”⁵², demonstrated that more than 3,400 real estate properties in the most valuable districts of São Paulo were owned by legal persons whose ultimate controllers were other legal persons – these latter registered in tax havens; that is, non-transparent jurisdictions. This is exactly the same pattern found in the first conviction related to the anticorruption investigation “Car Wash”: layers of non-transparent legal persons hiding the natural person who *de facto* owns a valuable asset. The report concludes that although it is not a crime *per se*, multiple layers of legal structures appearing as owners of real estate properties mean a high risk of money laundering and other illicit acts such as tax evasion.

⁵⁰See “Company investigated in the kindergarten mafia passed on money to the mayor of São Paulo, investigation says” (<https://www1.folha.uol.com.br/cotidiano/2021/10/empresa-investigada-na-mafia-das-creches-repassou-dinheiro-ao-prefeito-de-sao-paulo-aponta-investigacao.shtml>). Ricardo Nunes was mayor of São Paulo since May 2021. The 2017-2020 municipal administration, as previously explained, was led by João Doria and Bruno Covas. Elected vice-mayor, Covas became São Paulo mayor when Doria resigned in 2018. Re-elected in the 2020 municipal elections, Covas passed away in May 2021. Ricardo Nunes, vice-mayor elected, became then the city mayor.

⁵¹See “Parliamentary Inquiry Committee in the City Council will investigate possible ethical violations and lack of permits” (<https://www1.folha.uol.com.br/equilibrioesaude/2021/10/cpi-da-prevent-na-camara-de-sp-vai-apurar-de-eventuais-violacoes-eticas-a-falta-de-alvara.shtml>) 7 October 2021

⁵²See “Does Corruption live next door?” (<https://www.transparency.org/en/publications/sao-paulo-does-corruption-live-next-door>) 10 April 2017

While the watchdog uses of the IPTU dataset seem rare and occasional, the uses for policy analysis are more common. Besides news reports on urban land uses, at least one research center has used the IPTU dataset extensively.

A research center based in São Paulo, Centro de Estudos da Metrópole (CEM), has published a series of technical notes⁵³ analyzing trends related to urban land uses based on the IPTU registry. For example, technical note 14, published in April 2021, entitled “Spatial patterns of formal residential growth: municipality of São Paulo 2000-2020”. In the introduction, the authors write that they collected data “from the fiscal data of the registry of properties in the city of São Paulo produced by the Municipal Finance Department of the São Paulo City Hall Paulo (PMSP) for the purpose of assessing the Urban Land and Property Tax”.

As noted, the IPTU dataset allowed for analysis ranging from flaws in fire prevention and rise of evangelical churches to patterns of residencia growth.

Likewise, the IPTU registry seems to have been helpful for business purposes. It is reasonable to build the hypothesis that most private-sector uses remain unseen by public opinion. However, at least one real estate start-up firm has published about using the dataset. Loft has been writing analysis⁵⁴ on the real estate market in São Paulo based on the IPTU dataset.

After presenting the results from the document analysis, I proceed to the presentation of the interview data. Twenty accountability actors with solid expertise and experience have been interviewed — eleven professionals in horizontal accountability institutions and nine diagonal accountability actors.

4.4 The Perceptions of Diagonal Accountability Actors

This subsection explores the perceptions of diagonal accountability actors – investigative journalists and leaders of anticorruption civil society organizations – on the benefits and challenges of using open data to combat corruption at the local level. Before presenting their views on the role of open data to counter corruption in São Paulo, exploring

⁵³See

(https://centrodametropole.fflch.usp.br/sites/centrodametropole.fflch.usp.br/files/cem_na_midia_anexos/14-nota_tecnica_crescimento_residencial.pdf) It is noteworthy no stress that main author of this note is a former CGM-SP civil servant, what suggests that familiarity with the IPTU dataset is a driver for using it. For other technical notes by CEM <https://centrodametropole.fflch.usp.br/pt-br/publicacoes/notas-tecnicas>.

⁵⁴See for example “São Paulo built 1.2 million apartments in six decades”. 6.Jun.2022. In the news report, it is read that the data are from a survey “based on information from the IPTU”. (<https://www1.folha.uol.com.br/cotidiano/2022/06/sao-paulo-construiu-12-milhao-de-apartamentos-em-seis-decadas.shtml>)

the case of the IPTU dataset, I will recall the two institutions whose leaders have been interviewed and the community of practice relevant to diagonal accountability.

Abraji is an organization founded in 2002 that has been crucial in promoting access to information law. Today the institution, which gathers hundreds of Brazilian investigative journalists and whose office is in São Paulo City, still promotes transparency, monitors transparency policies, and trains journalists on access to information and data use. Almost 100% of the Annual International Conference Investigative Journalism has occurred⁵⁵ in São Paulo.

Fiquem Sabendo (FS) is a digital journalism agency specialized in transparency policies. Their goal is to push for government data to be more open and curate disclosed data to promote data use among journalists.

Forum has been a vocal community advocating for the passage of access to information law in the 2000s and early 2010s. However, after the sanctioning of the law, in 2011 and the first years of implementation, the Forum was emptied. Recently, however, when transparency laws and policies were perceived as being under attack⁵⁶ in Brazil, during Jair Bolsonaro's term as president (2019-2022), Forum was reactivated and revived.

The nine interviewees who are diagonal accountability actors hold or have recently held leadership positions at Abraji, FS, or Forum. All the journalists that have been interviewed are either former Abraji presidents or co-founders of FS — all are experienced in investigative and data-driven journalism and have been involved in training sessions for journalists in terms of data use. Civil society representatives consulted are or have been executive directors of relevant open data or anticorruption CSOs participating at the Forum.

The informants all live or have lived or worked in São Paulo City for many years and have experience working with local issues. In addition, these diagonal accountability actors have worked in areas or projects in media outlets or Civil Society Organizations.

I will present these actors' perceptions through the analytical lenses used for this thesis, the “publicity and accountability condition” and “corruption persistence”. However, before doing so, two “partial” success stories on open data use for corruption control will be showcased.

“Partial” success stories

⁵⁵The 17th Conference, in 2022, took place in São Paulo, as well as 14 out of the 16 previous events.

⁵⁶See, for example, a press note issued by Forum on 8 November 2021: “Organizations and news media speak out against using LGPD to deny information”. LGPD is the Brazilian privacy law (<http://informacaopublica.org.br/?p=4273>)

When asked to share success stories about data use for anticorruption purposes, one investigative journalist, FS cofounder, mentioned the following:

There was a case... I'm wondering if it fits in corruption, I'll tell you about it, then we'll think if it makes sense or not. But we published in Fiquem Sabendo the data related to the pensioners, right? (...). Because basically we knew until 2021 the total amount that the federal government spent on pensioners, but we didn't know who this money was going to, how, for what reason the person was entitled, since when the person received it, and so on. So Bruno denounced it, and there was a period of more than a year for the government to structure the data and put it into a database and correct the problems, and then they released this data. And then the population itself, you know, actually not the population, the journalists, the organized civil society started to look at this data after it was opened and several reports with very strange things started to come out. (...) But until then I remember that I was a little frustrated, because I thought: "Well, it is a great case of data opening, but it has not led to anything so far". But I found out that in July the Office of the Comptroller General [CGU] released a report - we published it - showing that after the opening of this data, they opened several investigations and identified irregularities in these payments. I don't know if it can be called corruption... maybe it is a way... it is a Brazilian way, which in a certain way is... I don't know if it can be called corruption, technically, but people who should not be entitled to certain pensions and are receiving them; so in the report they looked only at cases, let's say, a sample, and they have already identified there a deviation of 5 million reais. That's because they took a sample. I don't remember the exact number, but it is a small number of pensioners. And then what they say in the report is "if we expand this methodology to everybody, we will potentially save a lot more, we will revoke several, several pensions Because people couldn't be receiving for three different people or, I don't know, the daughter... the unmarried daughter who had the pension for the unmarried daughter there, and she got married but is still receiving... and things, things like that. So this was a case that took a long time, but I can see a beginning, a middle and almost an end, right? The data was closed, then we fought for it to be opened, it was opened, the journalists were all over it, they produced several reports about it. The CGU realized the importance of opening an investigation, identified irregularities and is in the process of beginning to punish. So, there were all... all the stages, and we will see if in fact they are going to punish, it has not happened yet... the "season finale" has not happened yet, but we are almost there (Informant NG7)

The case above is not related to a local-level dataset. The information was disclosed by the federal government, and, as the informant mentions there has been “beginning”, “middle” and “almost an end”.

Figure 1 represents the “stages” mentioned by the interviewee: data was initially out of the public sight; then a dataset is disclosed; next, data is used by accountability actors; then a sanction mechanism is put in place. However, as above-mentioned, the case has not reached an end.

There is another case of “partial success” reported by a journalist, and this one is related to the IPTU dataset in Sao Paulo. The news story is about top civil servants from São Paulo public universities who were supposedly receiving per diem irregularly:

I did one [news piece using the IPTU dataset] (...) I remember a specific news report that investigated the per diems that the employees of the state universities in São Paulo received. I had received information from a source that basically said that the university employees had... had not received raises for years, while the upper echelons of the state universities, more specifically USP and Unesp (Unicamp did not appear in this story) [USP, UNESP and UNICAMP are three public universities funded by São Paulo state], they, because they were there at the top of the chain, were able to keep changing the benefits, doing things there and almost hidden without transparency on the subject. And then the story they told me is that, for example, pro-rectors, the dean himself, these people lived in São Paulo for a long time, but they kept getting per diems from the University, as if they were moving from the interior because... I don't know, the guy has his office in Ribeirão Preto [Ribeirão Preto is a city in the interior of São Paulo state] (...), but he is dean, he's been dean at USP for "a hundred" years already. The guy has a residence in São Paulo. The same thing with the rector of Unesp (...) And then I remember that I showed it in the report; basically, I did this cross-referencing, right, people who received this aid, and theoretically should be moving around, but who had a property in São Paulo. I saw the database: this guy owns this house. And then I used this data to compare and say "Look, you have real estate in São Paulo, you have lived in São Paulo for so many years. Why do you ask for transportation assistance...(...) Sorry, the name is "per diem". And then they said that... basically, the answer is: "the legislation allows me, so I'll do it. (...) only that, so, the legislation says that it is for those who do not live in the city, for those who need to move around. And then this is also ... generated a complaint was to the Court of Auditors of the State. (...)

The Court the Court of Auditors recommended that there was return of these per diems ... also do not know what unfolded later ... it took a long time, (...), but the last thing I remember this episode is that the Court recommended, it was not an obligation, but recommended that they stop paying these daily rates and even there was this return, because it made no sense: if you live in São Paulo, you cannot receive travel per diem. (Informant NG7)

On the development of this case, I consulted⁵⁷ one the prosecutors of the Sao Paulo state who was in charge of the investigations. There was one case for each of the universities. This prosecutor is one of the professionals I consulted during the last quarter of 2022, after the series of recorded interviews.

The cases have been put on hold. The Court has adopted the reckless attitude of putting some cases on hold for years on end... There is no procedural delimitation of time limits and resumption of proceedings. Technically, there is no maximum term for a stay of proceedings. This option to suspend sine die the processing of certain cases is a potentially harmful grey area for external control. There is certainly political pressure from the rectors. (Informant not identified — asked for complete anonymity)

Asked how this political pressure manifests itself in practice, the informant responded the following:

Meetings directly with the Councillors [the São Paulo state Court of Accounts are led by a group of seven councilors], I assume, given the past context. I cannot say, because I have not witnessed it, nor have I had information from a reliable source about it. (Informant not identified — asked for complete anonymity)

Both cases seem to follow a similar path: data made available is used by investigative journalists. After the news reports, horizontal accountability institutions start a formal procedure. The cases, however, do not come to an end.

After these overviews on two partial success cases from beginning to an “almost end”, the following subsections presents the perceptions of journalists and civil society

⁵⁷Information provided via instant app messenger. The said prosecutor, who asked for anonymity, also signed the Informed Consent.

representatives about each “stage” through the analytical lenses used for this thesis: the “publicity and accountability condition” and “corruption persistence.

Publicity and Accountability Condition

Publicity condition comprises access and interpretation of information. The accountability condition refers to the capacity of actors to take action with the information to hold the powerful into account.

Journalists and civil society organizations have a more prominent role in the publicity condition, not only pushing for more transparency but also writing analysis based on the data. However, they have little accountability capacity, as these are more related to law enforcement.

According to the informants heard, journalists and civil society organizations do not access and use relevant anticorruption data, in general, given a variety of reasons.

Firstly, and in line with the literature, anticorruption and open data CSO leaders mentioned that data for anticorruption purposes are harder to access than data for other purposes.

We wanted to know, during the pandemic, which schools had adequate school infrastructure, and then monitor this. So, we took the data from the school census to look at that. And then we realized that they were very badly filled out, the quality was very bad.

They were not useful, they were not useful for control, you can't do social accountability. So somehow people are doing research, they are thinking that they can use them, they can monitor them and do something, but they can't really control them. On the innovation side, I have less visibility because we had little contact, you know? But the impression I have is that many companies use open data from what we hear. So I am also assuming that they can find it minimally reasonable, although it must have problems for them as well. But every time we went to do social control, but all of them, it is not a hyperbole, it is not a way of speaking. All of them. We saw many problems in the data and it was very difficult for us to do social control with what was available. So, Brazil is very well evaluated internationally in terms of data transparency, but the reality is that in practice... I don't know how other countries are, but in practice the data that is there does not reflect the reality; so in part it is as if it wasn't there, you know? You have a... it's not, it's not real there. Because you don't really know what information

is there. Supposedly it is, but you go... and it is very basic things like. "Ah, supposedly we have all the financial execution data, like how it is, payment, liquidation, etc.". And then you find out that, per item or per product purchased, it doesn't exist.

So, supposedly it seems that there is in the Transparency Portal, the federal government has everything there. But in fact, it doesn't... you have no way of knowing how much was paid for each item, you have no way of knowing, there is no active transparency, so every time we go to do this, we come across these problems. (Informant NG9)

Besides data on school buildings infrastructures, quality data on procurement is also hard to access, particularly at the subnational level, according to this informant:

One that is very... that is very specific and that Transparencia Brasil has tried to work on frequently is public procurement.

So, there is... we managed, for example, we managed with a lot of suffering... the only state, for example, we wanted the subnational data; so, the only state that had this available for all the municipalities was Rio Grande do Sul. Because the State Audit Court has this initiative, it imposes the use of a specific tool that they developed for all the municipalities to insert their purchasing and bidding data there. So, it is able to gather - as they should, as all audit courts should be able to do - this data and actively provide it to society. And it is indeed open data, it is structured, it is readable... almost beautiful. It has one or another little problem, nothing is perfect...but that was it. We managed to get the State Audit Court of Pernambuco, but, well, they have this data, but they don't make it available, they themselves don't seem to use it, so they didn't know about some of the problems with that data; for example, it was not possible to cross the bidding spreadsheet with the commitment spreadsheet... we said: "ok, you have this data, but we can't put it together". "Oh, no? The government produces the open data, but as they don't... they don't use it themselves, they don't know if they are doing something that is really good for fighting corruption, you know? And it is a control agency... imagine if they had this attention, how much they themselves could benefit from this in the fight against corruption?

And the rest of the states, they don't have it. It would have to be from municipality to municipality and look: how many municipalities, nowadays, have the structure or

know what open data is? Almost none. So, this is something out of the ordinary.
(Informant NG6)

Another informant highlighted the new privacy legislation as a barrier for public disclosure of corruption-related data.

Generally, to investigate corruption, you are investigating people, or companies... and this almost inevitably leads to the point of personal information, confidential documents, police investigations, administrative processes. So, I think that the great difficulty that already existed, and today this is only becoming more accentuated, is having access to things that the government itself, sometimes for technical reasons, cannot disclose. So, what is the biggest problem that I see today, that the access to information law, sometimes I feel, is even going backwards... because you still depend on sources to get some things. You can't ask for access to the police investigation, this is not disclosed at all through the LAI [Brazil's Access to Information Law], neither through open data, nor in the transparency portal. It does not exist. On the contrary, everything that involves... not even the Federal Police salary you can get on the Transparency Portal site of the federal government, because they understand that it has to have a special secrecy, a specific secrecy, right? And that makes us have to abandon these transparency mechanisms and go back to the journalism of the 2000s, I don't know, which is to call, which is traditional, which is to call the police chief and tell him to say: "Look, can you give me this inquiry? Can you give me this information?" "And then the person will give it to you off-line, I... he can't give it to you, so I think that this is the great difficulty (...)

So, these data related to government sanctions, so, Ibama [The Brazilian Institute of Environment and Renewable Natural Resources, Ibama, is a federal agency under the Ministry of Environment] notifications, police investigations, people that have already been arrested, this kind of thing... The OCCRP [Organized Crime and Corruption Reporting Project, a global network of investigative journalists] has a system called Aleph, which is open to journalists, you have to register, but once you are registered and approved, you have access to the database, which is very interesting, because they compile a lot of things, both public databases and others that are not public, but because they are journalists, they had access, so, for example, the Pandora Papers, Panama Papers, the Swiss Leaks recently... Swiss Secrets, which are basically

leaks that show bank accounts or offshore accounts of big companies or people in other countries. And then it is basically a search tool, like an internal Google, where I search there the name "João da Silva", and it will tell me if João da Silva appears in any of these leaks; if João da Silva has been fined by Ibama, if João da Silva has a house, if João da Silva has real estate... Anyway, it will give me a lot... it will almost make a mini dossier on João da Silva that I am researching there, and this is very complicated, it is very restricted today because we are dealing with sensitive information. So, it is not a database that is open to the whole society, for example. It is only for people that are approved by the OCCRP. There are all sorts of criteria for you to have access to it; I see that the future of investigative journalism and of these investigations depends on the cross-checking of this data. It is no good having separate databases, they need to be interconnected, you need to know that this person is in this, in this, in that; that this person has already been charged for this and that... isolated databases are not that useful, that is why tools like this end up being useful, because if you don't have them, then you depend on having a programmer, a person with some technical knowledge, who can cross-reference two or more databases (Informant NG7).

There are also concerns about the timeliness of the datasets, affecting the quality of the data:

And also to keep them up to date. Sometimes you create a portal there, you create a repository in an open data portal that eventually exists and you say "that's it, I've done my role, that's it, I'll come back and update it in two years, when the next open data plan comes out" (...) It's not much use... and it stops being open because if it's not updated, right? Because this [timely data] is one of the legs of open data. And as there is not so much demand, not so much use of open data, this is not so noticeable, "Ah, I'm here barely fulfilling my... my duty to provide open data, but nobody is watching, right? So why should I bother to do the right thing? Anyway, there will be no public embarrassment". (Informant NG6)

Another topic related to access and supply is the absence of dissemination or facilitation procedures to induce data use. These informants referred to the IPTU dataset expressly.

The city hall itself could have done it [a tool for easier external access to the data] ... a more friendly interface there, where you can make some filters and download only what you are interested in, just a region, right? I think this is a problem. (Informant NG3)

Generally speaking, about “supply side” lack of efforts in promoting the portals:

Now the issue of the offer.... governments don't bother to give more visibility to this, or to inform more about this, and also doesn't bother to actually offer information, sometimes you see agencies that have nothing open at all (Informant NG6).

As for the “demand side”, according to informants, there have been improvements in raising media and civil society capacities, as more training courses are teaching how to use data.

However, there is a relevant barrier to more actual data use: the crises of the journalism business model. Two elements connected to the business model have been mentioned. Firstly, media outlets tend not to invest in investigative journalism, which is not profitable.

Let's consider that there is a media crisis, you know, the large vehicles that will migrate to digital, will have great difficulty in defining, discovering a viable business model some, on the other hand you have new journalistic initiatives that start from scratch 100% digital with great potential, but that does not always have there this asset that is the experience of being a vehicle of communication experience of professionals with more... So you have... you have a scenario where you have to create a critical journalism that is relevant and professional, let's say? So you have... you have a scenario that is in some way unfavorable, the macroeconomic scenario of the media as a whole, and of the country at last. Even today you see the more traditional vehicles that move around like elephants, trying to situate themselves in the new times, but also, at the same time, trying to at the same time are the vehicles that have... have more professional journalism, many times because they have more professionals, more impact, more credibility. And this is also an asset for them, right? Let's put it this way... that they try to bring to the digital world (Informant NG4).

Second, the lack of professionals resulting in less time of dedication to data analysis..

There are many things that could and should be the target of reports, and that are asking someone to look at and put together, that end up being... being hidden for lack of manpower. I think that today there is a... a crisis of dedication time, right?

I think that we have improved in terms of journalists' capacity, journalists are more aware of these databases, they know how to handle them better, but we have worsened in one aspect which is: today, they have less time to do investigations that do not have an obvious lead, do not have a... which is what... you go on an exploratory journey of data without knowing very well what you are going to find... it was already a luxury in our time, but at least we could do it, and today I find it very difficult for anyone to do it (Informant NG5).

The high cost of hiring data science and data analytics professionals make data-driven investigative journalism projects more costly:

Another obstacle that has to do with the crisis in newspapers is the difficulty in hiring professionals from other areas. For example, at Estadão Dados, we had a programmer working with us... today, we can't, no newspaper can... the IT sectors of newspapers are depleted because banks, the financial sector, and even foreign companies are competing with these people to bring... these people that know how to work with public numbers, know how to work with code. So, the newspaper doesn't have the capacity to pay a professional from these areas today. And this is also a problem for working with large databases (Informant NG5).

The lack of leadership and strategic vision have also been mentioned:

In order to do good watchdog journalism and work with corruption, we need to invest in people, in good investigative journalists. Investigative journalists, time, so that these journalists actually have time to work on a story, and also investment in technical knowledge. So, you have an interdisciplinary team in the newsroom, or you have journalists with a greater technical capacity, which generally are professionals that end up being more valued, being more expensive, it is more difficult to maintain as well. So, in the case of journalism this lack of support from the top leadership, I think it is seen less as a lack of interest in the theme and more as a structural issue of journalism today, which is lack of resources, lack of investment, right? And there is a lack of strategic

vision from the leadership of how valuable this type of journalism is, right? And so, it is valuable in the sense of social impact, because if you think about the logic of the product, of the audience, it is the kind of journalism that generates much less revenue, I would say, for the vehicle than entertainment, sports journalism, you know? That's it: investigative journalism serves to generate social impact, but it is a market failure. High cost and even relatively low audience interest, you know? It moves public opinion, it moves the relevant actors in the ecosystem, but it is difficult to pay for this kind of journalism (Informant NG8).

Another barrier for data use, related to the scarcity of professionals in media and civil society, is the need to prioritize. According to the informants, diagonal accountability actors tend to prioritize the national level, leaving subnational entities unattended.

I think that it... this database [the IPTU database] ... it could have been better used. I will speculate about the causes: I think it has to do with the... back then, right, with a time when journalists were not so prepared to work with data. Maybe it has... Maybe the very existence of the Lava Jato campaign [Car Wash anticorruption investigation] has affected this in some aspect because investigative journalists were all looking at something else, you know? They were focusing a lot on campaign finance issues and so on, which didn't.... Necessarily...maybe it had some connection but not in such an obvious way it hit that database (Informant NG5).

Not only investigative journalists tend to focus on national affairs. That is a likely scenario for CSOs as well, according to this informant:

It is difficult to embrace all the problems; then you have to make choices, and choices are kind of random like this: you could go one way or the other... ah, but, someone gives you funding, you get funding, I think that funding directs a lot of what civil society organizations are going to do; then, if the funders are not interested in this, for example, in the territorial issue and social control, you are not going to deal with this because there is no funding.

So, if the funders are not interested in this, for example, in the territorial issue and social control, you are not going to work with this because you don't have it financially. I think there is a bit of this, and I think there is also the fact that our ecosystem is stronger for national issues than for local issues... and territorial issues

are local by nature. There are some things that can be, of course, like demarcation of indigenous lands, for example, which is federal. But the database you are commenting on, for example, is typically municipal and can be in several municipalities, right? I think we would need to have many... more organizations with this expertise to work with big data problems at the municipal level so that you would have many uses and a lot of social control. (...) At the subnational level, those who work with local social control have less capacity to work with databases of this type (Informant NG9).

In sum, there is limited capacity to interpret data installed locally. And even when such skills exist, and there are experienced professionals interested in local-level contexts, diagonal accountability actors based in São Paulo tend to focus on the national level, leaving local-level data underused. In addition, the fact that between 2016 and 2018 — the first two years of the IPTU dataset disclosure — the “Car Wash” operation was high on the agenda diverted the media attention to the federal level (federal government and federal state-owned enterprises). Investigative journalists and anticorruption CSOs were then looking at the national level, including those based in São Paulo.

Another barrier to anticorruption data use is the low economic incentives: few media outlets prioritize investigative journalism because this kind of product is not profitable. So, although there are initiatives promoting data use and data-driven journalism – and many journalists are interested and building capacity — the number of news stories is low. Informants mentioned that, because of the crises of the business model of journalism, fewer professionals work for media outlets, so it is hard to spare time for a lengthy news story, and media owners do not prioritize these kinds of news reports.

This is the data table related to publicity condition, according to the perceptions of diagonal accountability actors:

Table 13 – The “publicity condition”, according to data-driven journalists and CSO representatives

Second order themes	First order concepts
Low quality data for anticorruption purposes	Non-availability of the history of each property Non-updated and incomplete information Restrictions for anticorruption data - privacy and tax secrecy laws
Little promotion of data use	Supply side: lack of promotion and dissemination Supply side: absence of filters and tools for better data
Little incentives for local-level data use	Demand-side: Crises of the journalism business models Demand-side: Investigative journalism products as market failure Demand-side: High cost for hiring data professionals Demand-side: Focus on the national level

Accountability condition

The concept of accountability condition used for this thesis refer to the capacity of accountability actors to effectively use the information available to challenge corrupt practices.

According to informants, one barrier for data-driven journalism is the deficit in knowledge about legal types of corruption or the functioning of public administration. That is, journalists can be skilled in using software to read and interpret data but not proficient in having insights or raising reasonable hypotheses about corruption by looking at a dataset. As a result, the actual uses against corruption tend to be limited.

Informants also mentioned that diagonal accountability actors in general have little knowledge on public management, making it harder to fulfill the accountability condition.

And there is also a deficit in the training for data interpretation, you know, the proper use of data, so... one knows how to use, in some cases... In some cases, we see that you know how to use the tools, right? So, you know how to use spreadsheets or even R or advanced statistics, but when it comes to interpreting it within the context of public administration, how often is the data dirty or bad or reliable or not? And this sometimes gets a little choked up. So, this connection between the structure of the administration, how the public administration works, is missing a little bit, because this influences the way you look at the data.

If you take the general scenario of professionals today, and here I am going to focus specifically on journalism professionals, who I end up having a greater knowledge of, there are few of them that have not only the technical knowledge, you know, let's say, of the technological tools to work with data, but also the legal knowledge that I think you have a legal framework that needs to be known even for you to be able to formulate the necessary hypotheses for a good data analysis when you are working with social control of corruption. And this I think is missing a lot in journalism professionals, so that they can act effectively in this social control. Right? And even so, I think that within the watchdog function of journalism to check corruption cases. So, I think that, except for very obvious things, we can see the question of "rachadinhas" [popular term given to "diverting the salary of an advisor". In practice, it is a transfer of part of the salary of the civil servant to the congressman or secretary or minister who hired him or her based on a previously established agreement] or illicit enrichment, right? Or without... without justification of public figures that are more prominent like this... you have cases that

come to light, but that are very evident and obvious, and I think there are several other aspects of corruption that... that are not covered by the media and mainly because I believe that there is this lack of knowledge, you know, of how to gather the necessary data for this coverage of how to formulate good hypotheses about it, right? Where... What are the official sources to check this information and even the very unavailability in many cases of data or the need for you to cross-reference many databases, making the job super-complex.

In sum, these are the themes emerging from interview data when it comes to the accountability condition among diagonal accountability actors

Table 14 – The “accountability condition”, according to data-driven journalists and CSO representatives

Second order themes	First order concepts
Deficit of knowledge on criminal law	Lack of legal understanding around corruption and money laundering as crimes
Deficit of knowledge on public administration	Little understanding of the dynamics and structures of public administration

Corruption persistence

In terms of the persistence of corruption, even in a scenario of more transparency, diagonal accountability actors understand that networked interests and low institutional capacities lead to impunity, which, in turn, helps explain the resilience of corrupt practices.

Regarding networked interests, the quotes below refer to groups with similar interests related to urban land and real estate properties.

This informant shares experience related to the “Panama Papers”, a set of leaked documents from a Panamanian offshore law firm and corporate service provider. These documents are the origins of a series of investigative journalism stories published from 2016 onwards by various media outlets from different countries.

"A data story that caused political problems within the newsroom was the Panama Papers. We participated in this global investigation, and we ended up catching, among the offshore company owners and so on, the biggest advertiser of the newspaper (...) there was one of the [surname of the family who owns the media outlet — the surname has been omitted so as to protect the informant] there too". The articles were

published... maybe they didn't give it the prominence it deserved, the international prominence. The prominence abroad, in other vehicles, it was bigger.... but the articles were published. And what happened is that when we were invited to be part of the Panama Papers 2, let's say, the subsequent investigations, more leaks and such.... the newspaper didn't want to be part of it, without giving any further explanations... so, it became clear to us that it had been traumatic, and that they preferred... in this case transparency would not be important or good, because it would expose the interests of people with ties to the company itself."

(...)

"I vaguely remember that it also caused a certain discomfort, because [surname of a well-known businessmen who own a chain of private schools and colleges – the surname was omitted to protect the informant] was... had relations with the newspaper, he was also an advertiser and so on, but she ended up leaving as well.

(...)

Database investigations focused on fighting corruption is... there is a risk that you end up hitting interests of the elite that dominate the Brazilian press. Yes, it is quite possible that this has happened in other newsrooms, it happened in [a well-known newspaper was mentioned, but the reference was omitted so as to protect the informant] in this example that I mentioned and... Yes, I think that this is an obstacle to be considered, that... that... the thirst for... the search for information, and the flag of transparency, raised by the newspapers, that it has a certain limit, right?" (Informant NG5)

This informant quoted below refers to another aspect of networked interest: class factor.

And when it comes to urban property, there is still a class factor, real estate, that clashes with the interests of an economic elite that owns a lot of real estate and is not sensitive to this kind of transparency, you know? And the class of people... I remember the first... one of the first reports that came out was about the judges who owned real estate, you know? So if we also look at the prosecutors... I think that they are actors that do not necessarily have an interest in this... or have a conflict of interest there, right? If there is no firm movement there to... to keep this on the air, they are not going to make much of an issue of it (NG3).

The interviewee quoted below highlights the conflict of interest:

Does anyone have any doubt that SECOVI [association of real estate companies] closely monitors what we are talking about? So, for example, Flávio Amary was recently Secretary of João Doria [Mayor of São Paulo between 2017 and 2018]! He leaves the presidency of SECOVI and will be Head of Department of Housing. The jump from SECOVI to Department of Housing, is direct: the guy didn't pass through any other place before, he left from here to here. Presidency of SECOVI, Housing Secretary. This movement is not subtle, and it is not without... without its symbolism; so, the guy that was lobbying is now the secretary. So if he comes across someone from the administration, below him, trying to make a social control against the interests of where he was and probably will come back as soon as he leaves office, he will not walk in the public administration. On top there will be his hits and his natural composition among these peers, among these actors, so that social control is stopped there. (Informant not identified — asked for complete anonymity)

The quotes below refer to barriers for accountability institutions to properly activate their sanction mechanisms. These barriers operate in various ways:

- The shaping of institutions, regarding, for example, the process of nominating high leadership for the local oversight body (TCM-SP)
- The lack of priority in terms of capacity building for effective data use for accountability — the leadership controls the strategic actions of the oversight body and may not want to prioritize proper use of data and evidence.
- The lack of capacity could be fabricated not only internally. External groups can influence in the functioning of an accountability institution, by nominating allies for strategic positions or by hiring experienced law firms to judicialize investigations and postpone or stop collect of evidence and decisions.
- These internal and external barriers may lead to a sense of impunity that motivates more illicit acts.

The informant quoted below stresses the process of nomination of the high leadership of the Municipal Court of Accounts (TCM-SP).

We have to understand, everything is politics, right? (...) I'll give you an example. TCM has its seven councilors, their political choice is what is written in the City's Organic Law. But, what is how is it formed? How is the A, B or C seat on the TCM occupied? The councilmen appoint them, preceded by agreements conducted by the executive, or conducted by the legislative, which is where things get even uglier. Basically, a TCM of seven councilors - I tell you from experience that I know some of them – to mess with this? There are only one or two [that would investigate]! Because they are also involved with this kind of thing. So, how is the TCM going to handle an investigation of this size, if it is going to touch a hornet's nest and it can even hurt back, hurt TCM's own guys?

So, imagine that the TCM has today almost all of these councilors were city representatives in the Legislative House, the ones who are there [at TCM] today, right? So, you're going to hit the scandal of.... scandal that will affect the urban property tax?!? It will involve big values! It will involve CNPJs [tax id for legal persons] with some properties that have been empty for years, with no fines levied, some unpaid values.... then you will go back all the steps, you will hit a confusing urban general layout, and some of those guys who should investigate voted for that law proposition [that made a “confusing” layout be legalized]. So, that is it... this kind of database [the IPTU database] – coming back and summarizing in one sentence: it is kind of made for nobody to mess with, because it involves something very serious in São Paulo, which is property. And it is the issue of real estate speculation. Which is the occupation of the most expensive square meter in Brazil, involving very big things. And something of this size, when you have no way to explain it, it won't be the TCM that will explain it, nor the Public Prosecutors' Office, much less, so... Let it go, do you understand? (Informant not identified – asked for complete anonymity)

This informant mentions the need for building specific and strong capacities for data analysis and the lack of political interest in doing so:

I think that this barrier that I talked about, the lack of capacity, is not only technical, but also investment to be able to do it. I need a project to do this investigation, right? I don't think it is something simple that a person will look at the base, identify... they will need to make multiple crossings like the one Transparency International did on... "Corruption Lives Next Door", right? They had to take several commercial boards

and make robots to scrape them, so... that base by itself, I think it is not enough to do a process like this. So, if by any chance someone identifies some irregularity there, she, right? I don't know, a person, a case of illicit enrichment of a public official, they would have to gather a lot of evidence to instruct a process so that it could have some progress. This would require an effort, an investment, a capacity that I find very difficult to exist on the part of society; and, on the part of the government, I see the institutions weakened to do this. The Municipal Comptroller's Office itself, which could be actively doing this actively, right? So, I don't know, but if I were to say, what is the main barrier, I think it is more this point of... the difficulty of accessing these institutional actors or these... these means of denunciation, with information already instructed to be accepted, to have admissibility, to have some follow-up... it is not something simple to characterize, right? It requires a great effort. Look... at the municipal level I think that the political choice seems to me to be for the minimum effort of: "I don't want to investigate". "I don't want to open this flank", right, and reproduce a model that already existed in the state, for example, in the state of São Paulo... does not have an internal control system as developed as it was proposed at City Hall and as it is in the federal government. And I think that with the change of government there was a logic of: "No, why all this?", you know? "We are going back to being a little more passive here", right? There was no news of any operation, nothing like the level of the "Audit Tax Mafia" in the previous administration, so I think it was an option for a model of internal control that was emptier, more... less independent than the one we had in the previous administration, in the previous management, right? More weakened and by... I think it was a political choice of not wanting to give an outlet to these issues, right? On the other hand, there wasn't a public opinion pushing for this kind of action, right? So, I think there was some occasional criticism, and things went on, becoming less and less under the spotlight and... and that's how it stayed (Informant NG3).

This journalist mentions barriers to effective sanctioning mechanisms based on previous experience in digging into documents and having close contact with law enforcement and oversight agencies:

So, for example, when I interviewed people, they said: "Look, this is the strategy of the companies: they put in a bunch of law firms, very expensive ones, the best lawyers in the area, because they know that if they keep basically fooling around, they know they

are wrong, they know they committed a crime. So you just have to keep stalling, keep holding back... because at some point you will get a weakness, I don't know, "ah, you should have done the metrics with this standard and not this one, so it's wrong. Your report is unfeasible; in other words, you can't fine me anymore". Or, you can keep on fooling around, fooling around, until Ibama simply doesn't have the people to inspect and then... or to take the process forward, then it expires for lack of action. These are the strategies, so, it involves being a powerful company, both in terms of influence in the government and in terms of money, to pay these lawyers and so on; it involves having this influence.

And the last element is the lack of strength. I think that it has a lot to do with it, because when you deal with a big company, it puts several obstacles in front of them, and they have the advantage that they can do whatever they want, while the public agency can only do what the law says, right? They can only do what is in the Internal Regulations. So, if the guy keeps fooling you and you, by law, are obliged to answer all the letters or by law you are obliged to notify the person at the right address, blah blah blah, the guys will catch you with this: "I can't find it, I can't even find it"? There are many cases that they can't even find the person to, to notify, and then the action is time-barred because they can't find it. And then you hit the lack of power of the agency, right, of "what else can I do? Okay, I didn't find it, so I'll publish it in the newspaper and if it doesn't appear in a long time it will be fined and that's it; I'll go after this person's assets, I'll find a way to punish him. Or, I don't know, "increase the statute of limitations, because we are not being able to do it, increase the number of servers"; I don't know, so it is a little bit of everything... it is the powerful, it is the network, and it is the lack of strength of the institutions (Informant NG7).

This experienced informant hypothesizes about the motivations of corrupt actors:

My hypothesis, if it can be a... something very synthetic, is impunity, right? It's the... the barbarities that are done, even when there is light on them, are based, in my interpretation, on the certainty of impunity... or the bet of impunity. What we are seeing now with the so-called secret budget, right? It is secret up to a certain point, but there are many connections that are being shown, that are being evidenced. It is evident that there are benefits to electoral bases, that there are contracts that are absolutely suspect, public money is destined to actors that are completely muddy, and even so the guys keep

doing it... I mean, that is where the certainty of impunity comes in. And that's where, in the current case, the control bodies are completely dismantled. The government promoted a dismantling in these areas, in all... in all the branches of the state, where there was some kind of action against crime, against illegality. Even the environmental agencies (Informant NG5).

Therefore, regarding corruption persistence, we have the data table below:

Table 15 – Persistence of corruption, according to the perceptions of diagonal accountability actors

Second order themes	First order concepts
Powerful networked interests	Greater capacity to align interests - political and economic elites
Weak law enforcement	Low sanctioning capacity
Gaps in administrative and legal frameworks	Capacity to postpone administrative and court decisions

4.5 The perceptions of horizontal accountability actors

This subsection explores the perceptions of professionals with large experience in horizontal accountability institutions. Three of these institutions are focused on the city level: Office of the Comptroller General (CGM-SP), which is the agency of the Executive branch in charge of anticorruption and transparency polices; the Legislative branch (São Paulo City Council), and the Municipal Court of Auditors (TCM-SP), which is a technical agency linked to the City Council.

Professionals from the federal and state prosecutors' office (MPF and MPSP) have also been consulted. All of them are based in São Paulo and their institutions cover São Paulo City.

The only relevant institution that do not participate in communities of practices is the City Council, the Legislative branch of the local government.

Legislative has three functions: oversight, representation, and lawmaking. Pessanha (2009) argues that oversight is an overlooked function in Brazil: document analysis and interview data analysis seem to confirm this proposition.

São Paulo City Council as an institution apparently has no relevant activities on anticorruption data use, neither using the data nor promoting the use. There is no register or official document indicating that councilors or technical personnel have used the IPTU dataset or have significantly used other anticorruption-related data for accountability in the period between 2016 and 2022. This lack of activity is corroborated by two informants who have had

relevant positions at the City Council: a former councilor, who for one term hold the highest political position – the Speaker of the Legislative House –, a former research director. They gave the following testimonies:

The oversight role, this fundamental role that society has given to the Parliament, the Parliament of São Paulo doesn't exercise it. Literally, it doesn't exercise it. Until today the inspection and control committee, which at my time was translated into a theme of our Municipal Constitution, you know, in our organic law, hasn't been constituted yet. We had advances, such as a goal plan for some degree of physical control of what the Executive performs, but the Parliament doesn't even do the fiscal accounting control, the one that is recurrent, in the activity of the Court of Auditors that controls contracts, purchases and so on, but it also doesn't control the organic development process of the city. To the extent that it has failed to establish any measurement environment of these processes, then it does not control investment of the public resource, it does not control the outcome of this investment nor the impact that these hypothetical investments produce in the city for you to correct your course. By not controlling any of them, it also doesn't control corruption, right? So, the task of controlling the Executive — and this control that many times can be subtle — doesn't happen; on the other hand, the strident inspection does. So, the CPIs [parliamentary inquiry committee] end up being an open point of control that makes a lot of noise but doesn't solve anything. So, you set up the CPI to produce more corruption (Informant G9)

We can think about the oversight role of the Legislative, I think, from two fronts: one that is... what has been outsourced to the Court of Auditors (...). So, I think that this role is exercised well, but not exactly by the councilmen, because they have their own body [Municipal Court of Auditors] that helps with this. And it makes sense in a city the size of São Paulo that this is not played... that the State Court of Auditors that already evaluates the accounts of the State Government and 644 other municipalities. I think that the size of the capital city, the size of the state of São Paulo, somehow justifies the existence of this institution [the Municipal Court of Auditors].

So, from the perspective of public accounts, the councilors are not so dedicated to this, but they have their own body. Then there is the other role of oversight, which is to call the Secretaries [Heads of Department] to provide clarification, follow up on bills, eventually, if necessary, install some CPI and so on... and this is much more linked to

the mayor's ability to build his coalitions and have a more pacified or less pacified relationship (...)

So, this role of overseeing, I think it has a little bit to do with our political culture and a little to do with the formation of coalitions. The political culture, I mean, thinking about the English Parliament, for example, because the English Parliament is the complete opposite: not all parliamentary countries are like this, but in the English Parliament the presence of ministers there in the Legislative Plenary, discussing bills and being held accountable, is a daily occurrence, and here [in Brazil] inviting a minister, a municipal secretary, to come to the Legislative, to debate a public policy, always sounds like "we are sending a message". Even if it is an invitation, you know, it sounds almost like a threat... a thing of our culture. This could be more peaceful, considering the structure of our coalition governments at all levels... and that it is quite common for municipal secretaries, state secretaries and ministers of state to come from the Parliament. But culturally... we did not create this culture. So, certainly this side of inspection is the weakest arm of our Legislative (Informant G11).

The quotes above are additional pieces of evidence about weak checks and balances, particularly at the local level.

The other nine informants from the horizontal accountability cluster are active and experienced professionals based in São Paulo. Five of them have been Comptrollers at São Paulo City Hall and are familiar with the transparency policies of the municipality in general, and the IPTU disclosure in particular. Another one has been Deputy Comptroller at São Paulo City Hall while leading the “Information and Intelligence Production Advisory Team” – and is specialized in asset investigation. The other three informants are leaders of the MPSP, MPF/SP and FOCCOSP. The informant from MPF/SP has been a leader of the Car Wash Task Force in São Paulo – “Car Wash” was a huge anticorruption investigation in Brazil. The interviewee from MPSP leads a team of professionals who support prosecutors working in the defense of public budget and public assets. The other informant is secretary general of FOCCOSP.

In terms of the publicity condition, the institutional accountability professionals all agree that data is a crucial aspect of their daily work and many of them agree that data publicly available is interesting for potentially including society in accountability procedures. However, in terms of access, they argue that, as public-sector accountability professionals, they have other means for accessing relevant information – in general, they do not use open-source data, like governmental open data portals. This is because they can access systems and datasets directly

from the governmental agencies who prepare the data. There are challenges, though, particularly in terms of tax secrecy, privacy restrictions, and other challenges for data sharing.

Regarding use, these professionals argue that their colleagues and leadership have a law mindset and are not trained in technology and there is shortage of personnel, hence analyzing data is more challenging than accessing data.

On data quality and precision:

But, when we talk specifically about asset investigation, it is important that we can know the value of that asset and when the person acquired it. This is a key point, you know? Knowing that a person has a property that today is worth, I don't know, let's say, a million [1 million Brazilian reais], a property that is worth a million, right? So, it is like... the accuracy of the information would be very interesting if there was a history in which this was made available, but then I don't know in terms of amount of information, how it would be... the volume of information to be made available, but just having it is already nice, right? But for us, if we had this calibration, then it would be better.

I don't remember searching for information in this specific case [the IPTU dataset], in these investigations and information from the City Hall, because we use a lot of information from the registry's office itself, and there it is possible to identify the origin, the step by step, the history of the property, and it seems that the information from the City Hall about the IPTU is very much a photograph of that moment, right? about the person who is occupying there, but I don't have a history, I don't have a negotiation value, I don't have ... precisely focused on the issue of money laundering, right? (Informant G1)

On structured data:

Finding structured data... I always joke, I will give an example here, working with electoral data: the TSE [Tribunal Superior Eleitoral, the Superior Electoral Court], it organizes the election, it registers the election, it is not concerned with who will later take these election records and work with them. As much as it organizes this in comma separated format, it has a dictionary of variables, it has a series of things that for us sound like inconsistencies, but that for electoral justice make sense, right?

So, these are the two challenges: transforming information records into workable databases and knowing how to ask questions to the data. I think these are the challenges for those who are in any position, whether in the analysis of public policy or in the formulation... in the formulation of public policy, in the analysis of results and performance of public policy, from an academic point of view or from the point of view of civil society control (Informant G1).

On the non-prioritization of open-source data:

If this information is open or if my agency has this information, obviously I will first try to look for it within my agency. Because for me, maybe the reliability, especially the updating of the database, is greater than what is open, right? (...)

I don't know what level of knowledge the citizen would have in order to be able to... but I think that there are people who obviously, you know, even if they are from society, not linked to public agencies, who can do this and know, you know, give a... give a destination to this information. But my impression is that the agencies prefer what comes directly from the agency, and not what is in the... what is open, you know? (Informant G1).

On restrictions for data sharing between public institutions:

(...) this is the main obstacle to this kind of work. Because data is power, it means power, and the agencies don't want to give up their power, when it comes to requests made by an agency that has some power with this type, such as a control agency. So, I remember the sacrifice it took to access, for example, the Denatran [national agency that controls, supervises and inspects all traffic policies in Brazil] database. It was a fundamental database, which we needed to know the vehicles of those being investigated for corruption. To know if the guy had a Ferrari.... it was a sacrifice, we had to involve even the minister to get access to this base. A base in the public service that is right next to us and that we could not get access to... it was extremely difficult, this is just an example, right? Other bases that we later had access to: the Navy base, to have access to the Navy service... the boat service... I don't know exactly how it was. But it was very difficult to have access to this kind of thing. Inside the government itself. Imagine when you go to another sphere of government. For example, when you are in

the state or in the municipality and you want to have access to this kind of thing? It is very complicated. For example: the Receita Federal [Brazil's Revenue Service] is an agency that is extremely closed to grant access to any kind of data it has. Any kind of data. (Informant not identified — asked for complete anonymity)

Still on the challenges for data sharing among government agencies:

There is a lack of a better, let's not say partnership, but a bigger sharing between the control agencies... This results from a series of factors, whether they are of an investigative nature that says "ah, if the information leaks, suddenly the investigation won't be productive, right? The evidence to be collected will not be produced. Another question, at this point, is about vanity itself, right? "Oh no, it's my investigation, I'm not going to share it."

So, there is this question of sharing information about these two... two problems that I see and that exist, they are real, you see, Fabiano? they are real. "Ah, who is going to coordinate the investigation?" "Ah, no, it's my investigation, I am going to appear because I did a good job". "Ah, no, if I leak this information...". And it is really worrying. If I share it, maybe my investigation won't have the effect... right? (Informant G4).

On governments failures in promoting the use of data

(...) maybe the Office of the Comptroller General [CGM-SP]... far be it from me to criticize those who succeeded me, but... they don't do dissemination anymore, right? "Look, it is available!". What the databases are, you know? We planned that famous database catalog that we thought of back then, but what is the dissemination that you do? Nobody knows... when I say this in the lectures I give etc., nobody knows. Not even the political agents know. Not even the political agents know! (Informant not identified — asked for complete anonymity)

Still on the lack of promoting for data reuse:

(...) it is fundamental that the data producer, especially when we talk about the public sector, has a project to disseminate and disseminate widely. So much so because

you... generally those who search for data have a very specific profile, we are talking about academic researchers, and even in the public service it ends up being those who have some bias or a background or a more academic bias, or the press. And obviously we are talking about a press that has evolved a lot in the last years and... In terms of its capacity to search and work with data in large quantities and not only in anecdotal facts or specific cases, right? I'm saying that it is not only anecdotal because I'm thinking now for example of the UOL article that searches the president's real estate records, there are few cases, but it is not anecdotal, but obviously this is.... this is different from having a greater demand... and the training that people need to have to be able to work with a large amount of data. So... if I just have my data available there in an API for you to download and I just leave that API and don't tell anybody, it is a very restricted and very specific audience that is going to use it (Informant G11).

Regarding data use, which is the second element of the “publicity condition”, informants mentioned as a barrier the low proficiency of institutional accountability professionals in technology and data. Most of these professionals have an academic background in Law and have not been trained in technology or IT systems.

The agencies, the public institutions, in the Public Prosecutor's Office, the Public Defender's Office itself, I doubt that they have made good use of it [the IPTU dataset]. And this is not for lack of will. It is for lack of preparation. It is for lack of specific training in this sense. If you take the schools of the Public Prosecutor's Office, of the Judiciary, of the Public Defender's Office, which train not only the prosecutors and the defenders, but also the technical staff that goes there to do the research to take the people that are working there. They don't have specific training for this, they don't know this... at most, they know how to consult the COAF [Brazil's Financial Intelligence Unit]. The COAF is the most consulted agency for institutions that have passwords and that have... that have the cooperation term that can use that database. Only the COAF. The others, I doubt (Informant G3).

On data reuse and lack of knowledge on technology and data

This is the crux of the matter: not many people understand the importance of having a database... because you see... people: "Oh, I wanted to know if that person....

- is very specific, right? - if that person has a property. Right? And so, he thinks like this... for him, it wouldn't matter if it was a base or an access to a specific system, right? But those who are already thinking about doing this with a large volume of data to be able to check if many public agents have... they are few, but they understand the importance.

Because investigation is when you are in the case and you are diving, right? But this idea: "Oh, this is necessary to do, a permanent and broad follow-up for all public agents, for example, from your agency", right? For example: "I am doing this for all officers or for the Public Prosecutor's Office, I am doing this for the whole Public Prosecutor's Office. I don't hear this speech, you know? It's more like: "What should we do in the cases that you are willing to investigate? How to go deeper, right? But this constant monitoring with some databases, right? which I call bases that are not protected by tax secrecy, right? you get this large mass of data and cross-reference with data of the civil servants in your agency, right, in your institution and do this follow-up? I don't see it. (Informant G1)

Institutional accountability actors also mentioned the shortage of personnel and the high cost for hiring data professionals as barriers to data use within oversight agencies.

(...) I think that we have access to the data. What is our difficulty, and it has been getting worse over the last three years, especially, we don't have people to work with. So, this team that you referred to, "ah, you have an investigation team"... it is a team of, I don't know, between 5 and 8 people to serve 1,200 Prosecutors. This part of the expertise and analysis we end up in a line that often takes a long time because we don't have the staff to do this work; so there is no point in me, as a prosecutor, trying to do it because I don't have the necessary expertise (Informant G5).

Still on challenges around data interpretation and lack of human resources.

I have been in the Public Prosecutors' Office for 20 years and I can tell you: before, we didn't have data, it was difficult, we didn't even know if... there was no transparency like there is today, access to this information... today we have a lot, a lot, of data. The great difficulty is to relate this data, to sift through it and see what interests you; so it is putting the data together so that they give you information. So today we

swim in data. But you need to have a vision. Then you need to have the practice to understand what this data wants to tell you. Crossing them doesn't give you much information, so we really use it, it is very important. We are trying to transform it so that instead of training all the initial or intermediate prosecutors, you know? We are trying to create a nucleus of technicians that can do this initial part (...)

There are people who find it easy to deal with; there are people who are good at investigation and there are people who find it easy to understand the internet, to search for data, right? And then when you get people who are very much from the Law, it is more difficult for you... to train for an infinite number of things; so, you see, I have 20 years [of career], and I still think that I am the person who is interested in a lot of subjects, so... I am there in many, many areas involving the MP in digital investigation and so on, and even so I feel difficulty (Informant G6).

Shortage of personnel is also connected to the difficulties for hiring professionals with such skills.

There is a lack of manpower to read the information, to extract... because, for example, I have a Law degree, my background is in Law.

I need to extract an accounting data, many times. So I... I can also have a daily experience, a reading, an identification, but I can't use it officially... I need an answer from someone who is technically qualified. And the person who has the technical qualification, many times sees, extracts and passes me the data so that I can transform that into law, into a process and use it. So, many times this expertise of the person who is extracting the data is missing, which today is the data analyst. You see, private companies have the same difficulty with this workforce today, it's... because it's... the guy who analyzes the data, he extracts the necessary information and gives you the elements to read and think what you, in your expertise, in your function, in your area need (Informant G4).

In sum, we have the data table below, considering the perceptions of horizontal accountability professionals based in Sao Paulo for the “publicity condition”.

Table 16 – The “publicity condition”, according to the perceptions of horizontal accountability actors

Second order themes	First order concepts
Problems of data quality for anticorruption purposes	Non-availability of the history of each property Non-updated and incomplete information Lower confidence on open data portals
Little promotion of data use	Government fails in disseminating open data policies Restriction for data sharing: tax secrecy Restriction for data sharing: privacy legislation Need to build confidence among government agencies
Constrained institutional capacities on technology	Law professionals untrained in technology and data Shortage of personnel High cost for hiring data professionals

Regarding the “accountability condition”, informants mentioned a lack of institutional leadership for countering corruption using new technologies and data. They also mention lack of institutional capacities, including specialization in anticorruption, to counter specialized law firms.

I think that those who are still in leadership, and here I will include myself in this aspect. I don't participate in any leadership, but I am in the age group of those who are in leadership. And this and this generation was not a generation I am putting here of people who are in their fifties and upwards who have lived with this reality. That they learned the importance of these databases, of making them public, of working on top of it even for the institution itself, to be able to handle it more efficiently. So I think that it is a cultural issue... it is not in the culture of my generation and of the generations that preceded me from now on, this culture is growing, I think. It is a subject that has been in the headlines in the country, and that has been investigated... the need-to-know data, to make the data publicly available, but my generation did not participate in the formation of this subject, I never heard about this in the Law school that I attended. This was not even talked about, it didn't exist (Informant G3).

Still on institutional capacity

For me, one barrier is that we still have a work structure from the last century. You have a natural prosecutor that takes care of everything, and you are going to face lawyers that are extremely specialized in this issue of corruption. First, you already have a person that is... that is already hiding all this. He is already doing this, many

times, with all the initial information; so, in order to fight the wrong, always doing that very Cartesian thing, you don't get there. So, you need a little more audacity to do it, and you need specialization. So, I guess that's how it is, my opinion, the structure that today we have the Public Prosecutors' Office and even the Judiciary, we have to be specialized in everything, we talk about everything and then, the most complex thing we have today is a corruption investigation; and if you have, I don't know, all the crime, all of it... you don't have time to specialize in this, so you do what you can, on the surface, and leave it there. And if you don't go deeper, you don't find it. So I think that we will have to change... and it is very complicated, the public structure is very complicated to change... that we would need to have regional public prosecutors for people who are highly specialized to work on a certain thing. For me, today it is more important to have a regional public prosecutor's office for the public patrimony, for the environment with people that only do this all day long, employees that do this, and so on, than to have a prosecutor in the place of damage, in a certain city (Informant G5).

Table 17 – The “accountability condition”, according to the perceptions of horizontal accountability actors

Second order themes	First order concepts
Insufficient anticorruption institutional expertise	Lack of specialization in anticorruption
Lack of data-minded leadership	Absence of technology and data mindset at management level

Persistence of corruption

Assuming that power imbalance and path dependence shape institutions and society across history, I assume in this thesis that accountability institutions may not be properly equipped to hold corrupt actors to account. Personal and professional connections (networks) built over the years and organizational mechanisms constantly adapted find loopholes that allow corrupt practices, while weak or controllable checks and balances ensure impunity.

According to institutional accountability actors based in São Paulo City some these assumptions are valid, particularly the limited capacity for accountability institutions to effectively sanction corrupt practices even when data and evidence abound.

Beyond this perception of inefficacy, informants mention cultural and political challenges. The perception of widespread corrupt practices and tolerance toward illicit and unjust acts are seen as disincentives for political leaders to build effective anticorruption

systems. In addition, these accountability actors argue that impunity leads to demotivation and the search for “comfort zones” for their peers.

The perception that fraud and corruption is not only tolerated but put in practice, particularly at the local level government, is seen in the testimonies below:

There is a degree of tolerance in this curse of corruption; so, the feeling I have is that part of society goes in there and says: "no, everybody steals a little bit, let's leave it and so on and so forth. So, this... is something that is impossible to make acceptable by a control agency. How can a control agency say "no, it's just a little corruption", right? Because... So, this is the first problem we have, which is a cultural problem; then, society accepts it and has been accepting it for a long time - Brazilian society, you know? - and has been accepting this; so, this is a first clipping. Without interpreting it, it is difficult to say the following: "but is there a fight against corruption? Yes, there is. But we have a society that is tolerant to corruption, dammit! They talk, talk, talk, but they practice it, right? They shout, they go to the streets, there are a lot of very strong movements against corruption, when you go to see the guy's life... the guy is pickpocket... so it is difficult, difficult (Informant G5).

On the perception of widespread corruption and the lack of political incentives to effectively counter corruption

I think that we need to evolve as a society as a whole. (...) ... when the discussion gets to this point, especially informal ones, in bars, all around, and people know what I work with and start criticizing, everything and then I ask them: "when you went to do a construction project in your house or a friend of yours or a relative, what is the standard procedure? Do you go there and ask for the correct instructions at the City Hall? [Ask] someone who had a company that knows how to do it? Then they are allowed only do one floor, and the person does two floors, right? Isn't that the way our society works? Our society, the informal part of it, is very developed, isn't it? I tell them... this is our [accountability professionals] day-to-day life, you know? We start talking to the people who are on the other side. This is the day to day... I remember... I think that trying to talk about something that does not go too far, that has already been reported, about big jobs that we did in some secretaries in the city government about integrity that generated... in the end what does it generate? They change the

secretary, because things stop working. You institute the controls that were supposed to be instituted, improve the procedures and everything... but then it all stops working. We had cases like this, right?

The secretary said "let's make an integrity plan, improve our governance... let's institutionalize controls, let's improve procedures", and there was a general outcry from society, which was the user of the service. So, it stopped working! And then what does the mayor do? And what does the mayor do now? Sometimes it seems that we... the impression is that we need to evolve to make things work in a better way and... sometimes the shock doesn't work, the shock... I don't know if we have to go in parts... my experiences with shocks were not good, you know? On both sides. Because we know that there is corruption on both sides. No civil servant is there demanding money from the honest ones, right? On the other side they are also trying to do things faster, wanting to do what they can't, to take advantage and... unfortunately, many cases that I saw is that everyone thinks it's good, many people think it's good this way, from both sides... "this is better, this way, we can solve it". But it is usually unfair to those who are outside this circle, isn't it? Because those who take advantage of this are the ones who are already there, who already have more financial power, more contracts... they make things move. And then everything moves (Informant G10).

Perception of widespread illegal practices and impunity, even with more transparency, lead horizontal anticorruption actors to demotivation. There has also been mentions of searching for “comfort zones”, the quest for more accommodated ways of facing their tasks.

(...) knowledge has not changed people's conduct... the misconduct is not something seriously reprehensible, something that people... I think it's very serious to choose to work with... I'm already getting discouraged about fighting corruption, because I think it's much more serious if you, you know, someone, a group that embezzles 500 million [Brazilian reais] from a... from a poor city, from health care, than, in fact, someone who steals with a gun, but this guy goes to jail... who is stealing [with a gun] there..., the sensation that Brazilians have of this reprehensible conduct, it's smaller in the fight against corruption, you know?

I don't know, now, this is a very personal opinion, which is not very technical. I don't think the judiciary really cares about making... changing... changing our culture. I think that's so ingrained in our culture of being corrupt that I have some difficulties to

understand; so I'll give you an example of a judgment that I saw last week in the STJ [National Superior Court, last court of appeal], which was about a "ghost employee", saying that it is not embezzlement because the person is hired, he never works, he only receives a salary. It is a culture that we have... everyone knows, because in a small town everyone knows.

(...) So it is not something that is hidden, it is something that is even publicized. So in this case, this employee specifically, she received almost 500... 500 thousand [reais] was... I think it is like this, from the time that we were able to demonstrate, half a million [reais] and then the unanimous decision said: "this is not a crime because the money is hers; she is not embezzling, it is hers, because it is hers, she has the right to it". Nobody, nobody has the right because you only have the right to receive money if you work! So much so that we don't receive it before! It's not like condo fee that you pay before and per month, right? You only receive it afterwards, if you, if you really worked! (Informant G6).

On the high costs for countering corruption

A certain comfort zone... Yeah... a certain comfort zone. You face this kind of issue and work harder, looking at this, it not rarely generates conflict. At the point of the São Paulo City Hall, you have also witnessed this, I am asked why I was breaking the secrecy of the municipal civil servants by accessing the IPTU database. That is it. So maybe it would be much more comfortable to keep quiet, right? Doing that traditional audit and not touching this hornet's nest (Informant not identified — asked for complete anonymity).

On impunity:

No, there was not [no sanction for any illicit act]. Because as the task forces [Car Wash task forces] were exterminated, you know?, I stopped working with these cases, the person who was in charge of these cases just finished the investigations; There are occasional cases in which even with this consistent wealth of data in their hands, you choose not to investigate, for legal or technical reasons, not always supported, you know? (...) For other reasons, for example: you have such a large database and when you have these elements in your hands, when you go to organize this data for the

accusation, you realize, for example, that the crime is time-barred; then you don't You don't start the criminal action or you don't start the criminal action for a reason, a legal understanding, that the member of the Public Prosecutor's Office sustains, saying, no, I don't know, I think that this is not a crime. Like, "I don't want to stay with this investigation, I need to find an excuse". But these are very specific cases (Informant G5).

Political leaders may consider that it is not interesting for their careers to lead the implementation of transparency and anticorruption practices, given the perception of tolerance toward corruption and the challenges and risks of establishing integrity practices. At this point, it is essential to remind that Fernando Haddad, the politician who created the CGM-SP and gave the green light to open data policies, including the disclosure of the IPTU dataset, has lost re-election for Sao Paulo mayor at the end of 2016. Two years later, he ran for President and lost again. Finally, in 2022 he was a candidate for governor of Sao Paulo state and again was not elected by popular vote. This anecdotal evidence suggests that being an anticorruption champion does not necessarily mean political and electoral gains, at least at the local level.

One informant mentioned these disincentives:

One thing that I think has become kind of evident in these years, I think it is one of the obstacles, I think it is the main obstacle: I think that, as a society, we have a lot of difficulty, I see, to understand these transparency movements, so, what do I see from the manager? "Ah, I am going to open the data, then they will... then suddenly they do a job and discover that the IPTU is outdated, there is an error in the collection". He receives much more criticism for the error exposed than for opening up to try to correct it, right? I don't know, in my mind it was always very difficult to understand this... and in fact this is the reaction that I always see, you know? Usually the manager that exposes himself and creates an Office of the Comptroller General, that opens data and... he is not valued for this... as much as he is punished for this act... for what this act exposes, you know? Unfortunately, we are still very slow in public management, aren't we? There is a lot to improve: there is corruption, there are mistakes, there are deficiencies... and it seems, to society, I see that the manager who hides is more... he ends up more valued. "Man, this guy doesn't...." These things: "this one doesn't steal, this one doesn't have problems, this one we don't know...". Hey? I don't know how to deal with this, but it is something that is clear to me, you know? It is very difficult today... not only in the City

Hall, but also in the Court, in a private company... when you discuss any kind of opening of data, it is... a kind of action that is sometimes a little more innovative, this transparency thing from There is always a back foot, and you end up understanding the manager's side. This is much more criticized for what they will potentially discover, even if they fix it, even if they have measures to fix it, than what they will than what he will get in terms of political return... because there is no way, there is a political level that decides this, right? About having had this initiative. This is more or less the general summary. I think that bad faith, really... corruption and bad faith, is the minority of the problem in this... I don't think that every manager is there in cahoots, is getting paid.... but the political risk of doing this is a big obstacle for us to improve in this area (Informant G10).

This lack of incentives for political leaders to lead effective anticorruption institutions and transparency policies seem to lead to underfunded and underequipped accountability institutions, especially at the local level.

The technical issue is a challenge for us, the budget issue is a challenge, the lack of planning compromises too much because we, who are civil servants, stay, but every cycle of change of government, we see a lot of things being dismantled and then when it is finishing, finishing to start structuring, then it all starts from scratch again. So, this brings an institutional commitment that I think is very serious, you know? At the federal level I feel that there may be more continuity, they have a tradition, public careers may have more strength in this sense, you know?

At city and state level, I don't feel this institutional issue as much, and as I said, it ends up compromising; so, we don't have much strength yet, in the sense of creating systems that are more agile, to assimilate this importance of making data available, of working with data, of taking ownership of this data, you know? This ends up being a very individual thing, for each professional (Informant G8).

Lack of institutional capacity.

Just to give you an idea — and this is already a result of my academic research — I applied a questionnaire to... there were 28 courts of auditors in Brazil. And only three carry out asset investigation. That's it... my questionnaire was applied in 2018,

right? The law... the law that provides for the possibility of the Courts of Accounts to investigate the assets of high authorities is law 8730 of 1993. And in 2018, it was 25 years old. Twenty-five years of law... and what did I imagine? Well, all the Audit Courts do this kind of investigation work because it is in the law... (Informant G1).

On expertise:

Conducting an investigation has a process of preparation, of investigative intelligence... investigating is not for everyone. It is not. You don't become an investigator overnight. It takes technique, science and I think that our accountability institutions were not prepared for this.

So, when I say: it has to be exemplary, it is because you have to be prepared for this. The administration in São Paulo, the Comptroller General in São Paulo, today is not prepared to carry out the investigations that it was 10 years ago! It has been losing its investigative capacity, so it's the other way around. Very competent accountability institutions, such as the State Public Prosecutor's Office, have not deepened their investigative capacity. So, Eliot Ness used to say: "Follow the money! Follow the money!". Man, is it so hard to follow the money? (Informant G9).

These perceptions of widespread corruption and social tolerance to corruption translated into low political incentives for adequate institutionalization of accountability agencies – particularly at the local level – seem to result in the malfunction of these institutions, with loopholes and contradictory decisions, both in the administrative and judicial contexts.

The testimonies below report some of these loopholes and flaws.

The open data policy, it is a necessary policy, but it is not enough, right?, to guarantee everything that is desirable in this area. You mentioned some situations that will then extrapolate the very scope of the control of the administration, they will slide into judicial situations, privileged jurisdiction [in Brazil, the holders of certain political leadership positions are subject to investigation, prosecution and trial by a previously appointed judicial body, which is not the same for people in general], situations in which there is also a whole procedural complexity to prevent the conclusion of these procedures, right? To make more room for the prescriptions to take place. So there is a gear historically used in favor of impunity. This mechanism has not been dismantled.

In the administrative sphere too, you know? What happened at the CGU, from the point of view of disciplinary processing, did not happen in a similar way in the states and municipalities. So, the CGU can be more effective in holding federal civil servants accountable, but mainly in situations of, let's say, ordinary acts of corruption. And in the state of municipalities, this is more difficult because in the federal system the CGU investigates and then holds them accountable and applies sanctions, right? So the Federal Inspector General's Office has the competence to investigate, to hold accountable, and then to apply a sanction. In states and municipalities, no: the comptroller's offices only do a preliminary investigation; then the accountability stage goes to the public attorney's office, to the public prosecutors; and then the sanction is applied by the secretary of each area. So, this creates back and forth processes, you know? So, in the City Hall of São Paulo, I saw this happen several times, you know? in São Paulo there is this... this deficiency too, you know? I tried to correct it, but I faced a lot of resistance. (...) what I saw happening in the city of São Paulo was that many times the Municipal Inspector General's Office would conclude a preliminary investigation; it would forward the case so that the public servant would be held accountable; and when it went to the competent Attorney's Office, which is called Procedural, in the city, the Attorney's Office for disciplinary proceedings is many times. The Attorney who was there didn't understand a lot of times that there was still a lack of instruction and that it was necessary to better characterize this or that circumstance of the investigation. Many times, he would return the case for further investigation, so a lot of time was lost in these comings and goings, and also this distancing of the one who is going to hold the attorney accountable from what was done during the investigation ended up creating an opportunity for the sanction not to be applied. And, even so, in the end, the one who applied the sanction was not even the one who had done the investigation nor the one who had done the accountability process. So, this is what still happens in the states and municipalities: an excessive segmentation of the disciplinary process, making it take several years; allowing room for the statute of limitations and for corporativism, you know, for that spirit of teamwork of those who don't want to take on any colleague. "Ah, it's already there in the environment you live in", you know? (Informant G7)

On loopholes in the legal and institutional frameworks

And then I ask you to listen very carefully and understand why in all these hypotheses, ok? All of them are very well grounded legally. And we know that there are legal options all the time. Especially when we work with a... with a "constitutionalization" of the law. And there is a tendency to judge based on principles. When... you know very well: when we work... when we judge based on principles, rejecting laws that have a constitutional presumption, because they are not considered adequate due to such and such principle, this opens a range for any type of decision based on the Constitution. And when we have a broad constitution, like ours, it is easy to justify any issue... so if you ask me.: "look, have you seen this specifically?" For that, no. I've never noticed it. But if you ask me "Oh, but could it have been that?" It could. But it's my legal option... it's hard to substantiate, to speculate "oh, there's a powerful one. Oh, why not, because nothing would happen here". I don't know, because it was a legal option of the judge. My function is not to judge. I cannot criticize the one who judged, who has the power to judge, a valid legal option that is not teratological, right? So it is very difficult, so I think that the problem would be more in the legal framework to face these issues than from a personal point of view. "Look, nothing here can't be investigated because..." Really, I don't think this is the focus. I think that our legal framework allows us to eventually understand situations like this. But, on paper, it is properly justified based on our legislation, nothing teratological, they are legal options. And legal options based on constitutional principles, which are very broad, right? So, I think that this is a flaw today... our general problem (Informant G4).

I will say: the case, in the concrete case, because it is public information, you know, of a process in the area that I work... many times you, the Court of Auditors, you know, it... in relation to the accounts of the Head of the Executive: governor and mayors. The court does not judge, it issues a brief opinion, which is sent to the Legislative Assembly or the City Council, because this is the political system. Like it or not, that's it.

Many times, a fact, for example, which is the same, for a mayor, let's say, from a neighboring municipality, is one outcome in the City Council, for another it is another outcome, which is a local political decision, but the fact is the same. Ah, that mayor is more powerful [than the one from the neighboring municipality]? Does he have more influence? I don't know, these are facts, I can't get into this kind of opinion, you know? when I speak technically, you know? You have to have proof, you have to have it, right?

So, you see the same fact, the same situation with different aspects... internally in the Court of Auditors there is a jurisprudence.

(...). Technically, you have to try to understand the distortions in our legal system rather than specific facts... specific facts, "look at this guy, he wasn't punished, a lawsuit wasn't filed against him, because he is so-and-so", this I've never seen, never. "Ah, it was instituted, but he didn't have... no punishment because it's him, nothing was going to happen". I've never seen it, ok? Never seen. What you see is the use of the system by means of appeal, by means of one system, the other it goes from... maybe the biggest problem is the system, it's not an enforcer, ok? Obviously, it is not a general rule, we have problems eventually identified, right? What is the problem? I think that the framework is set up in a way that does not allow you to assign the proper responsibility to those power managers with more responsibility. I think this is a big legal difficulty (Informant G8).

As noted from these testimonies, horizontal accountability institutions have the capacity to access data, including some anticorruption data. These professionals, however, prefer accessing data directly from other government agencies, rather than open government portals. Although they list some problems in data quality, informants mention that interpretation and analysis of data is more challenging than accessing data, given shortage of personnel and insufficient capacity to work with technology and data.

And in the occasions when these barriers are overcome and data is analyzed and produce evidence, the investigations do not necessarily result into sanctions, according to the interviewees. They list institutional flaws and loopholes, both in administrative and judicial instances that are used by law firms and other actors to avoid the sanctioning mechanisms.

In sum, here is the data table regarding corruption persistence.

Table 18 – Persistence of corruption, according to the perceptions of horizontal accountability actors

Second order themes	First order concepts
Accountability professionals' demotivation	Impunity reducing accountability professionals motivation
Cultural norms	Perception of tolerance toward corruption
Gaps in administrative and legal frameworks	Postponement of administrative and court decisions Controversial and contradictory court decisions
Low institutional capacity in sanctioning	Low capacity or ability to properly sanction
Political disincentive	Perception of no gain for integrity champions

4.6 From “publicity and accountability conditions” to “techno-legal and socio-political conditions”

The “publicity and accountability conditions” proposed by Lindstedt and Naurin in 2010 initially framed the conditions needed for transparency to impact corruption control. In this thesis, based on more recent literature about data, technology, and corruption control, I advanced the concept to include the elements of “data” and “communities of practice.” The notion of a “more complex” accountability, discussed by McGee (2019), also brings more nuances to the level of efficacy needed to hold powerful actors into account, particularly at the local level. Finally, the elements of power imbalance and path dependence, derived from the historical institutionalism approach, allowed for more specific analytical views regarding the functioning of institutions.

This case study demonstrates that “publicity and accountability conditions” provides a rich framework for approaching the relationship between transparency and corruption control. However, we need more detailed elements to improve its explanatory power, especially in times of big data and particularly at the local level.

Document analysis and the interview data reveal that actors in charge of oversight, both from inside and outside the public sector, currently have access to various datasets. Nonetheless, there are limits when it comes to anticorruption-related data. Therefore, the first element of the “publicity condition” needs more nuance. It is not simply “access information”. Not all information helps push anticorruption investigations. Data needs to be complete and timely to serve as reliable evidence.

Regarding the second element, the capacity to interpret the information received: while technology allows for more information to be available, it brings the technical challenge, both in technological and legal terms, of adequately making sense of the data. That is, creating valid inferences with a high degree of accuracy to serve as solid evidence in administrative and judicial procedures.

When it comes to the “accountability condition”, the empirical data collected demonstrates that the mere existence of sanctioning mechanisms, and the possibility to activate them via elections or court, is insufficient to state that the condition of accountability is present. Political parties and messages may confuse voters, reducing their choices or misguiding their perceptions. Administrative and justice courts are in place, but powerful actors may have the chance to reduce the effectiveness of sanctioning mechanisms. Tolerance toward illicit practices and the perception of widespread corruption may inhibit engagement. In that sense,

“the accountability condition” element is a valid basis, but it is insufficient to adequately explain the failures in controlling corruption.

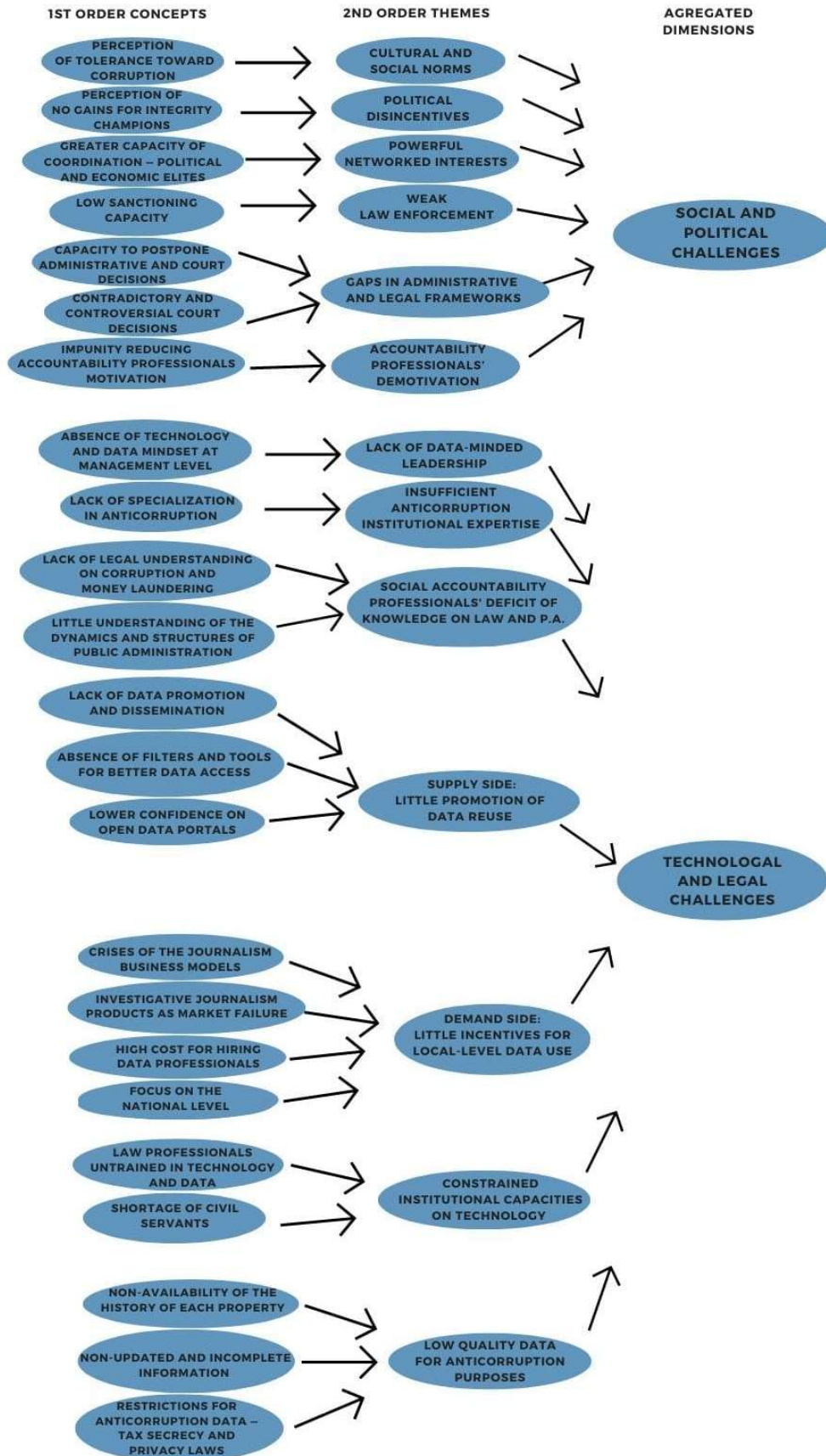
Therefore, I propose the framework of “techno-legal and “socio-political” conditions as a novel framework to address the role of open data to counter corruption at megacities of the Global South.

The “socio-political condition” refers to social consensus around integrity combined with political conditions for leaders to establish effective transparency policies and accountability institutions. A shared understanding of the relevance of fairness and impartiality for the proper functioning of life in society can incentivize political leaders to implement adequate transparency and integrity policies. Although these conditions seem unfeasible in more complex and diverse societies, it is reasonable to consider that those agreements could be reached locally. More effective accountability institutions, whose functioning is not affected by maneuvers, would be able to use the available data properly. Sanctioning mechanisms would use data and evidence to address malfeasance coherently. The “social-political condition” operates on a macro level.

At the organizational level (or meso-level), the “techno-legal condition” refers to the technological and legal instruments for producing, disseminating, and reusing data to address illicit acts. Both diagonal and horizontal accountability institutions must be able to prepare and reuse data to hold corrupt actors into account. Accountability institutions need to establish multidisciplinary teams with complementary skills, particularly in information technology, law, and public administration.

These two conditions emerge as the counterpart of the empirical evidence on challenges collected. The Figure 7 presents the challenges at the macro-level (the social and political challenges) and the meso-level (the technological and legal challenges).

Figure 7 – Challenges for open data to more likely produce expected outcomes on corruption control in megacities



The next chapter discusses the results.

5 DISCUSSIONS OF THE RESULTS

From the documentation, direct experience, and interviews, it is possible to examine the challenges for local-level accountability actors to reuse data made available via open data initiatives, build inferences and try to use these analyses to counter corruption.

The “partial” success stories presented illustrate how distant the relationship between open data and corruption control is at the city level. Even when diagonal accountability and institutional accountability actors effectively put forward a corruption case based on data, there is no guarantee that the case will be concluded soon. While it is unthinkable to consider that each accusation is correct and would necessarily result in sanctions, it is noteworthy that when it comes to corruption-related cases, there are various steps for data to bring about desired and expected outcomes regarding corruption control. The limited number of “partial” success cases at least bring examples of effective data use — notwithstanding the limited production of results.

On the other hand, the “unsuccessful” stories are numerous. They demonstrate that subnational level accountability actors have i) difficulties accessing good quality data, ii) struggle to publish news stories or reports or present cases before the Judiciary or administrative courts, and; iii) may find deterrence, suspensions, delays, and loopholes within sanctioning mechanisms.

Good quality anticorruption data is hard to access, particularly at the local level, confirming what was found in other studies (Laboutková, 2018; Žuffová, 2020). Government agencies publishing data online seldom promote data use actively and do not build tools for better data visualization or collection. In addition, datasets can be incomplete and outdated. Moreover, there are important restrictions related to tax secrecy and privacy legislations. Regarding institutional accountability actors, they prefer to access data from other government agencies via data-sharing agreements, but these deals may take time because it demands mutual confidence.

When data access obstacles are overcome, data analysis challenges appear in a subnational context. There are few accountability professionals for data interpretation. Although there has been more training for investigative journalists and non-profit professionals, there are few incentives for these actors to analyze local-level datasets: investigative journalism products are seen as a market failure, and CSOs may be diverted from local issues because of donors funding national-level topics. Institutional accountability organizations have insufficient data professionals, and they may lack specific expertise on anticorruption, as

oversight agencies have other topics to deal with. Despite these difficulties, there are occasions in which data can be accessed and used, confirming the optimistic findings of a few empirical studies found in the literature review (Lourenço, 2016; Mendoza & Ko, 2015; Royo-Montañés & Benítez-Gómez, 2019).

Nonetheless, whenever data analysis challenges are overcome, and an administrative or judicial case is put forward, the presence of data and evidence do not assure effective sanctioning, at least at the local level. Impunity is the observed outcome even in the presence of well-documented cases of wrongdoing and even when critical junctures occur, as political leadership change or anticorruption mechanisms are put in place – see, respectively, the case studies in Mexico (Aguirre-Ochoa & Gómez, 2021) and Italy (Di Mascio et al., 2020) mentioned in the literature review section. This resembles the power asymmetry and path dependence elements of historical institutionalism lenses. Formal procedures, rules, and institutions are shaped across time, in a chain of reactions and counter-reactions, with actors with political and economic power as well as social connections eventually prevailing. Specialized law firms opposing law enforcement agencies with insufficient capacities in the courts and political appointees in strategic positions within government agencies – these are some of the strategies to protect special interests. These tactics relate to the “stability and enforcement mechanisms that reduce uncertainty around controversial operations”, making them appear “less morally censurable”, as suggested by [Vannucci \(2015\)](#). These flaws and loopholes also connect to weakening of checks and balances at the local level, as proposed by Meza and Pérez-Chiqués (2021).

Strong institutions matter. A professionalized oversight institution within the Executive branch, with political support, can promote open data policies and effectively use and share relevant data, as CGM-SP demonstrates. The legislative branch at the municipal level, however, may not adequately exercise its oversight role. Outsourcing check and balance activities to a technical arm cannot be the solution, however, if the leadership of this specialized agency is closely connected to powerful interests. It is noticeable that the focus here is not a tiny municipality — this case unveils weak accountability institutions in a megacity.

The “publicity and accountability conditions” is an interesting concept to examine the relationship between open data and corruption control. However, it does not capture some of the nuances mentioned above, particularly in terms of the “accountability condition”. For Lindstedt and Naurin, “[t]he most important sanctioning mechanism for citizens in a political system (besides perhaps revolution) is the ability of people to choose their government in general elections. Political accountability through elections is also complemented by legal

accountability through the courts. Democracy and the rule of law are thus crucial for accountability” (2010, p. 305). Elections, though, can compensate corrupt leaders, as the evidence demonstrates worldwide. And the rule of law needs effective institutions to materialize; otherwise, it becomes a vague, senseless, and abstract idea.

In that sense, this thesis proposes the analytical model of “sociopolitical and technological conditions” to examine the relationship between open data and corruption, building upon Lindstedt and Naurin proposition. These authors’ contribution was precious, particularly considering the moment it was put forward: in 2010, a moment in history when there was great enthusiasm around transparency and open data. However, as Matheus and Janssen (2020) stress, it was only around 2014 that scholars started being more realistic about the limits of open data. The figure that closes the previous section summarizes the content of tables used for translating the concepts and themes that emerged from the empirical data. Empirical data presented the flaws and gaps in adequate transparency and accountability for the São Paulo case study. They also showed the drive for corruption persistence.

The themes presented are discussed below.

5.1 Social and Political Conditions

The social and political conditions for open data to bring about desired outcomes regarding corruption control in the context of a megacity comprise six themes:

- Cultural and social norms towards corruption acceptance
- Incentives and disincentives for political leaders
- The level of coordination of powerful interest groups
- The level of weakness or strength of law enforcement agencies
- The effectiveness or failures of administrative and legal frameworks
- The level of motivation of accountability professionals to carry out corruption cases

Corruption studies using the economic paradigm focus on institutions and structures of incentives (Chuah et al., 2020; Granickas, 2014; Klitgaard, 1991; Lambsdorff, 2003; Rose-Ackerman, 2013;). In contrast, literature on corruption from the cultural paradigm focuses on social and moral values (Barr & Serra, 2010; Filgueiras, 2009; Gonçalves & Andrade, 2019; Torsello & Venard, 2016).

According to interview data, society and institutions tolerate corruption, particularly at the local level. However, the fact that certain social practices are in place is not exclusively connected to moral values or preferences: other drivers may also be present. For example, the perception of unfair or unfeasible norms, especially in jurisdictions where society is disconnected from the political and decision-making processes (Torsello & Venard, 2016). As March and Olsen (1983) argue in their seminal article on new institutionalism: interests, powers, and the rules of the game are neither entirely exogenous nor endogenous to institutions. Integrity programs and procedures in the public sector may be perceived as unnecessary obstacles to economic development. As such, political leaders and public sector managers may be incentivized to loosen the norms and procedures. In that sense, the most plausible explanation for continuous social acceptance of corruption is related to the development of the institutions and social perceptions across time in a cycle of reactions and counter-reactions. Therefore, although the anticorruption rhetoric is a political weapon and is frequently used in the political discourse, the incentives for political leaders to establish transparency and integrity policies may be negligible. As a result, the perception of widespread corruption and the presumed tolerance for malfeasance is associated with high costs for implementing anticorruption policies.

Associate with these chains of perceptions and incentives, another theme comprising sociopolitical conditions for open data to produce anticorruption outcomes is networked interests. Professional associations, for example, may have robust government relations departments to defend their interests actively. Moreover, economic and political elites have interests in common related to the increase of their gains. Politicians need money for their political campaigns, for example, while businesspeople and rent seekers need to have the chance to shape regulations – or at the very least to minimize losses. As Ang (2020) proposes, “access money” are reward offered by business actors to politicians or public sector managers for exclusive access and valuable privileges. These networked interests seem to be even more present at the city level, where frequent personal and professional contacts allow for more opportunities for coordination and alignment. These powerful interests may be able to shape institutions, such as in the nomination process for crucial positions in accountability organizations.

Over the years, the influence of special interests may surgically shape organizations, including oversight institutions and law enforcement agencies. While opposing the creation of anticorruption agencies or the implementation of transparency and open data policies would probably bring reputational costs, powerful and resourceful actors learn how to navigate into

these institutions and norms and protect themselves, identifying loopholes. Even when some levels of transparency are in place, and there are oversight institutions, corrupt forces will find ways to avoid sanctions, given the unequal dispute between accountability institutions and powerful, unscrupulous actors. Court decisions may adopt unorthodox, albeit valid, approaches vaguely based on constitutional principles, resulting in impunity. Administrative disciplinary procedures also have loopholes and gaps, allowing for postponements and contradictory decisions.

All this leads to a perception of impunity which in turn leads to the demotivation of crucial actors: the “principled-principles”, that is, the accountability actors (Peiffer & Alvarez, 2016). In a context of impunity, these accountability professionals may feel it is more appropriate to search for “comfort zones” and decide it is not beneficial to engage in anticorruption. The findings of this thesis confirm collective action theorists’ arguments about the high costs of countering corruption (Kubbe, 2018; Mungiu-Pippidi, 2013; Persson et al., 2013), even when there are anticorruption institutions in place.

5.2 Technological and Legal Conditions

The technological and legal conditions for open data to produce the expected outcomes regarding corruption control in the context of a megacity comprise seven themes:

- Data-minded institutional leadership
- High-lever anticorruption institutional expertise at horizontal accountability institutions
- Adequate institutional capacity in terms of technology and data at horizontal accountability organizations
- Knowledge of criminal law and public administration at diagonal accountability institutions
- “Supply-side” (government agencies) active promotion of data reuse
- Anticorruption data quality
- Incentives for the “demand side” to reuse local-level data

In times of data abundance and legislation promoting transparency, horizontal accountability institutions should be data-driven: these agencies should easily collect and

effectively use evidence for preventing and sanctioning corruption. However, as examined in the previous subsection, there are few political incentives for establishing effective anticorruption institutions. Moreover, institutional accountability agencies are usually led by professionals trained in law, with little knowledge of the potential of data and technology, particularly at the subnational level. The lack of political incentives and the absence of data-minded leadership at the management level in public sector oversight organizations lead to poor open data policies, with governments implementing the minimum requirements. Data quality is poor, particularly at the local level. Given the lack of priority and the lousy quality, the government, and the suppliers of data, tend not to promote their open data portals and initiatives actively.

In contrast to what often occurs in government agencies, data-minded professionals tend to lead diagonal accountability institutions. However, investigative journalists and CSO professionals focus on national issues rather than local issues. Moreover, with limited resources, these groups need to channel their time and energy, which leaves local-level open data underused. Moreover, while investigative data-driven journalism and the not-for-profit sector are relevant for activating formal accountability mechanisms (Bernhard et al., 2017; Lührmann et al., 2020), they do not have the “teeth” to counter corruption.

Another theme is related to specific areas of expertise. As a result of the lack of political incentives for establishing effective sanctioning mechanisms, horizontal accountability institutions may not have the specifics for working with anticorruption as they deal with other topics (such as monitoring various public policies). On the opposite side, powerful corrupt actors have the resources to hire law firms with substantial and specific experience in anticorruption. Although social accountability organizations are sometimes relatively well-equipped in data analysis, they may not have the necessary expertise in criminal law and public administration. As a result, news stories and CSO reports may not bring robust evidence of crimes or illicit acts to the public domain, which in turn will not facilitate the work of law enforcement agents.

It is important to note that from a long-term perspective, the thirteen elements comprising the sociopolitical and the techno-legal conditions were probably better at the beginning of the 2020 decade than they were twenty, thirty, or forty years before. There must be advances and setbacks if one assesses each element separately. However, improvements in institutional capacities and better access to data have not been able to make open data policies effective interventions against corruption at the local level, particularly considering the sanctioning aspect.

5.3 Theoretical and Policy Implications

The role of open data for tackling corruption is an understudied topic. On a recent paper analyzing various technologies and their role in fighting corruption, Adam and Fazekas (2021) conclude that “rigorous evidence on the impact of big and open data on transparency portals is still relatively scarce”. The authors added that “we need a more nuanced and detailed understanding of how transparency portals providing open data can have an impact on corruption” (2021, p. 10).

This thesis intended to fill this knowledge gap, providing a more nuanced and detailed understanding of the relationship between open data and corruption control.

The basic element, from which the research question was elaborated, emerges from the literature review and current debates about the characterization of corruption. Is corruption a result of power asymmetry, as collective action theorists argue, or a result of information asymmetry, as institutionalists sustain?

Based on the principal-agent model, institutionalists argue that the citizens need to have more information about the people to whom we delegated representation power as well as more information about their actions conducting public institutions and public affairs. More data would empower citizens and increase the risks of political agents to engage in corrupt practices.

Collective action theorists criticize this approach, stating that the cost of challenging corruption is high, particularly in contexts where there is a perception and an expectation of corruption. In that sense, it is not adequate to assume that the mere availability of data will allow society to act against corruption. Because there will likely be no actors capable of confronting established networks and interests.

Given the centrality of data availability and accountability performance in this debate, as well as the perception that corruption at the subnational level is more entrenched, particularly in the Global South, the research question for this thesis is the following: *Have accountability actors been using open data to prevent and repress corruption and money laundering cases involving urban land and real estate in megacities of the Global South? If so, what is the role of data on corruption control?* This broad question is built upon the case of São Paulo’s database on real estate property, which was fully available in public domain for six years (2016-2022).

The central hypothesis is the following: due to the expectation of impunity, historically given, added to the lack of skills and resources to interpret data, accountability actors that could challenge local-level corrupt practices tend not to use the available information, even when a

key dataset, such as real estate ownership, is openly and freely available to be re-used. And this lack of action is present even in large global cities, not only in smaller municipalities.

A combination of two analytical lenses was used in the thesis. Firstly, “the publicity and accountability condition”, a concept put forward by on a paper published in 2010 to examine the relationship between transparency and corruption control. The authors argue that two conditions must be in place for information to have impact on corruption: publicity, defined as the capacity of actors in society to access and interpret government-held information, and accountability, defined as sanction mechanisms (elections or court) available to society. This concept is useful for the examination of i) data availability, ii) capacity to interpret data and iii) capacity to activate the sanctioning mechanisms, traditionally vote – society would punish corrupt politicians – and judicial conviction (courts would sue corrupt leaders). This lens relates to the information asymmetry view of corruption.

A second analytical lens derived from the historical institutionalism. Considering the power asymmetry approach to corruption, two elements from historical institutionalism – power imbalance and path dependence – were used. These elements are interesting because they can provide insights about barriers for accountability actors and institutions. Powerful actors in a corrupt context can shape the institutions so as to protect themselves from being held to account. In the long run, path dependence makes processes and norms crystallize, making changes and transformations virtually impossible, even when critical junctures appear.

Corruption is apparently more connected to power imbalance than to information asymmetry. After years of transparency legislation and open data portals, it is safe to hypothesize that the information asymmetry between citizens and political agents is smaller today. Even so, the perception of corruption is still high in many contexts, particularly at the local level in countries of the Global South.

This does not mean that information is irrelevant. Apparently, there have been attempts for reducing transparency, especially when it comes to corruption-related data. In addition, there are various constraints for effective use of data for anticorruption purposes. Social accountability actors may find it difficult to adequately engage with corruption-related data given their limited knowledge on criminal law and public administration. Moreover, investigative news pieces are not profitable, and the journalism business model is in crises; hence fewer media outlets would allow their professionals to engage for days into these types of data. On the other hand, institutional accountability actors may find it hard to engage with data given constraints in terms of personnel specialized in data analysis. The same dataset may be more often used for policy analysis and private-sector innovation.

Although information asymmetry still exists and seems to be relevant for explaining failures in countering corruption, the findings of this thesis is more suggestive about power imbalance. Apparently, powerful actors have the capacity to influence in the design of institutions and, even when these institutions try to act, powerful interests seem to be able to reduce the strength of the sanctioning mechanisms, finding procedures and gaps to delay or escape any form of accountability. As a result, corrupt actors are able to defend themselves, even when there is data and evidence of malpractices.

This thesis contributes to the literature in three ways:

- It contributes to the theory on open data for anticorruption purposes, and theoretical discussions on data for development.
- It adds to the academic debate on conceptualization and mis-conceptualization of corruption, particularly recent discussions on the centrality of information asymmetry or power asymmetry as a central element of the concept.
- This thesis also adds to the growing corpus of research on local level corruption.

Open data for anticorruption

The main focus of the thesis is to examine if open data impacts corruption control in a megacity. A literature review on emerging technologies and their impact on corruption control (Adam & Fazekas, 2021, p. 10) have found three main barriers: mismatch between supply-side and demand-side, few communities of users, and absence of resources and sanctions.

A literature review on open data and corruption conducted in this thesis found two approaches. On the one hand, scholars discuss the potential of open data to counter corruption but present no empirical evidence of the impact. On the other hand, more realistic authors argue that anticorruption-related data are seldom made publicly available. Additionally, in rare cases where there is some relevant data availability, open data portals will only impact corruption control if the “publicity and accountability conditions” are in place. However, none of the papers examined advanced a more detailed or sophisticated understanding of the elements that comprise the “accountability condition”.

This thesis fills this gap, as it brings a more nuanced explanation of the necessary conditions for open data to be impactful. In addition, while confirming Adam and Fazekas (2021) about the existence of the three barriers above-mentioned (mismatch between supply-

side and demand-side; few communities of users; and absence of resources and sanctions), this thesis introduces more details about these obstacles.

The transformation potential of data is not easily achieved. Firstly, for open data to bring about the expected outcomes, there must be intense and systematic use of the information contained in datasets. However, there are several barriers to data use:

- quality of data: timeliness, completeness
- lack of promotion from the "supply side"
- From the "demand side", that is, the potential users, the barriers are manifested in three ways:
 - Lack of capacity to interpret data: journalists, in particular, may fail when it comes to building valid insights and inferences related to crimes or administrative failures, given lack of knowledge about criminal Law or public administration.
 - Lack of opportunity to publish data-driven reports or news stories, given low financial resources and the need to focus on a few issues, generally national-level affairs.
 - Lack of capacity and resources from horizontal accountability actors – as their background is usually Law, they don't have the training for data analysis and may rely on a few data science professionals occasionally hired by the public sector.

Whenever accountability actors have the chance to overcome the above-mentioned barriers, the reports or judicial or administrative cases resulting from data analysis encounter inefficient accountability mechanisms. These obstacles, in turn, lead to disengagement and disinterest in further data use, given the expected limited impact, creating a vicious cycle. In such a context, many institutional accountability actors, whose role is to hold the powerful into account, tend to search for “comfort zones” instead of challenging the status quo. More transparency and data will not impact corruption control in such a context.

The barriers above-mentioned recall the “publicity and accountability conditions”; however, they bring more subtle elements.

However, there are other types of obstacles related to social and political contexts. The challenges mentioned above are encountered at the organizational, institutional level.

Nonetheless, as March and Olsen (1983) argue, interests, powers and rules are not exogenous to institutions. In that sense, the functioning of a public sector organization, including accountability agencies, is dependent on the social and political environment. Tolerance towards corrupt practices in society and lack of political incentives to build effective oversight institutions can help explain the failures in sanctioning mechanisms.

These barriers mentioned above for data use also adds to the discussion about data for development. According to the World Bank (2021) and enthusiasts of the open government agenda, data would benefit society in three ways: greater accountability, increased business opportunities and better policy analysis and public services delivery. Document analysis and direct experience, however, indicate that uses of data for greater accountability are less frequent than the other two. One plausible explanation is that policy analysis and business innovation do not face all the above-mentioned obstacles. While data quality and lack of capacity for data interpretation may also be relevant barriers to policy analysis or private sector innovation, these uses do not face failures in accountability mechanisms. In this vein, this thesis also adds to the academic and policy debate about data for development.

Theories on corruption

The economic approach, to which the principle-agent model is the central pillar, still informs most of the anticorruption initiatives, according to Villeneuve et al. (2020). Based on the assumption that corruption results from information asymmetry and principles (accountability actors) will benefit from more transparency, these initiatives are often information tools. However, there has been recent academic debate around this concept of corruption. Some authors (Heywood, 2017; Marquette & Peiffer, 2020; Pyman & Heywood, 2020; Wedel, 2012) argue that the mis-conceptualization of corruption leads to wrong assumptions and inadequate policy prescriptions. These scholars sustain that corruption is a collective action problem, and in contexts where corruption is expected, there are high costs for any actor to challenge corrupt practices. The findings of this thesis seem to confirm the feasibility of the former approach, as the disclosure of relevant anticorruption-data reduced information asymmetry but have not brought about expected outcomes on corruption control.

In that sense, the thesis also contributes to the current theoretical debate on corruption, adopting an underexplored lens to analyze corruption: historical institutionalism. Elements of power imbalance and path dependence seem to help examine corruption, particularly when looking at the capacity of institutions to counter corruption while confirming the high costs of

challenging corrupt practices. The argument that accountability is “more complex” than anticipated also seems valid. This thesis demonstrates that it is reasonable to hypothesize that entrenched interests are relevant to shaping accountability institutions' functioning.

Conceptualizing corruption as an information asymmetry problem does not adequately capture the phenomenon. Information is relevant, but power asymmetry seems even more critical to understanding corruption's persistence. In that sense, the notions of policy capture (Fazekas & Tóth, 2018; Kaufman & Vicente, 2011; Olver, 2021) and “power extraction and preservation” (Amundsen, 2019) may help explain corruption.

This thesis also contributes to the growing academic production on city-level corruption. The power imbalance is even more relevant at the local level than at the national level. Accountability actors are more prone to use data related to national issues; therefore, national or federal open data policies tend to be more efficient. In other words, as data use is more intense, corruption due to information asymmetry is more likely to be noticed at the national level rather than locally. Several law enforcement agents based in São Paulo used the expression “comfort zone” to describe a situation where accountability actors prefer not to take risks while doing their work. These descriptions highlight the relevance of the approach that criticizes the focus on information asymmetry.

The power imbalance is more present at the local level for two reasons. Firstly, given the proximity of powerful actors, the context allows them to coordinate their interests better. Second, accountability institutions tend to be less developed at the subnational level when compared to the federal or national level.

Policy implications

Lastly, the thesis will also have implications in policy discussions because it brings a more nuanced and sophisticated analysis of the potentials and limits of open data interventions in countering corruption, particularly at the local level. Disclosing data without anticipating challenges for data interpretation and use seems not to be an advisable policy. Gaps in accountability institutions and mechanisms need to be addressed with priority. Policy prescriptions focusing on open data as an anticorruption intervention must focus on the following:

- Data quality: timely and complete data is essential. In terms of anticorruption, completeness of data also refers to precise information about natural and legal persons. However, given privacy legislation and tax secrecy laws, datasets

sometimes omit the names. In that sense, policy prescriptions need to be calibrated to determine the relevance of personal data for the public interest. While public interest data does not refer to intimacy (sexual options or religious beliefs, for instance), names of lobbyists actively pursuing influence decision-making in governments should not be regarded as private data. Ownership of legal companies also seems to be public interest information, and the fact that names are in the public domain does not seem to affect intimacy. Regarding accountability institutions, they need to share relevant data among themselves responsibly.

- Data use: on the “data supply side”, governments need to promote data use actively by communicating the existence of relevant datasets, ideally “curating” the data, and presenting possibilities of impactful use. On the “demand side”, users must be trained not only on how to interpret data, technically and legally, but also need space and incentives to publish their data-driven analysis.

- Accountability mechanisms: transparency and data are essential; however, evidence-based anticorruption cases may remain unsolved or constantly delayed if accountability institutions do not map their inconsistencies, flaws, errors, and contradictions and solve them. Accountability institutions must efficiently conclude cases in a fair, professional, and timely manner, respecting the rule of law.

As Schedler (2008) states, accountability should be a “modest concept” because we have to assume, from the outset, “that politics is a human enterprise, which as such is characterized by elements of freedom and indeterminacy”. For the author, accountability must rely on carefully constructed rules, but it must not suffocate the exercise of power in a regulatory straitjacket. After all, dishonest actors could use one institution with sanctioning authority for unlawful purposes. On the other hand, the lack of sanctions, particularly in an era of greater transparency, leads to distrust in democratic institutions, opening the door for extremism.

Therefore, increasing transparency without a careful reality check regarding the sociopolitical and techno-legal conditions in place, particularly at the local level, will not affect corruption control.

6 CONCLUDING REMARKS

Governance and public administration studies have long dedicated energy to examining the relationship between transparency, accountability, and corruption. However, studies about open data, including the technological aspects, are not abundant, particularly at the subnational level. This thesis helps advance the understanding of the interconnection of open government data, accountability, and corruption control at the subnational level, focusing on a megacity: São Paulo, which is internationally recognized as a reference in open government.

São Paulo City Hall announced at an event marking the international anticorruption day in 2015 that one of their policy interventions to tackle local-level corruption was to fully disclose on the internet, in reusable and cost-free format, the urban land and real estate dataset. This is a crucial dataset for controlling corruption and money laundering in the real estate sector. The dataset was fully available, including the names of property owners, from early 2016 until early 2022, when the local government decided to withdraw the names of legal and natural persons from the publicly available dataset.

This thesis investigated whether accountability actors have used the information freely available for six years. Recent academic debate on corruption conceptualization demonstrates that some scholars understand corruption as a result of information asymmetry. In contrast, others criticize this approach and argue that corruption results from power imbalance and the lack of actors able to challenge corrupt practices. Therefore, accountability actors are a central element in the discussions around corruption.

Suppose the principle-agent approach is correct, arguing that the leading cause of corruption is information asymmetry. In that case, the radical transparency over properties in São Paulo City must have affected corruption control because accountability actors must have used the available data and must have actioned sanction mechanisms. On the other hand, if corruption is a collective action problem, then power imbalance may have prevented the accountability actors in the city from activating anticorruption mechanisms.

One limitation of this study is the possibility that some uses of the disclosed dataset may have been left unnoticed, given the lack of transparency and information about methodologies and data collection. However, although success stories may have been overlooked, this study demonstrates, from document analysis and direct experience of the ground, followed by confirmation with knowledgeable informants, that accountability actors have not consistently used urban property data while it was fully open. In addition, in-depth interviews with experienced accountability actors from the government and civil society based in São Paulo

City helped assemble detailed and nuanced elements on the barriers that prevented more intensive data use.

Open data is not enough to counter corruption at the local level, even in megacities like São Paulo. In these global cities, some actors in society are well-trained and could potentially impact the proper functioning of local accountability institutions. However, they may devote their time and energy to national affairs, leaving local issues under attention. Moreover, where tolerance to corruption is still a norm, political incentives for pushing for broad institutional reforms are scarce. Therefore, local interests feel free to shape local institutions — weak checks and balances assure impunity and leave room for more corrupt practices, creating a vicious cycle. These are findings related to global cities, not small or medium-sized municipalities. In such cities, data access and use barriers should be even higher.

In megacities of the Global South, the distance between the availability of data on the internet and the desired outcomes regarding corruption control is not negligible. For open data to produce effects on governance at the local level, essential changes must occur, both in society and in public sector institutions. There must be socio-political conditions and adequate institutional capacity for data to bring about the desired and expected outcomes on corruption control at the subnational level.

The fact that access to critical anticorruption data is still a challenge indicates that information remains relevant. However, this thesis demonstrates that accountability is also a significant challenge, particularly at the local level. And power can decisively interfere with the functioning of accountability mechanisms and institutions.

Future research can better explain the low capacity of local-level societies to sanction mismanagement and corruption adequately. Interesting avenues are discussions about political corruption associated with state and policy capture, power extraction, and preservation. In addition, researchers should analyse if and how political issues influence the shaping and functioning of local-level institutions more often. The impact of the reduction of the civic space and the crisis of investigative journalism are two other possibilities for future research on the relationship between open government data and anticorruption.

Open data policies with negligible effect on corruption control may even bring unintended consequences. As the citizens are better informed about mismanagement and corruption while seeing impunity and no transformations, confidence in democratic institutions may decline. Future studies can bring more subtle elements about the relationship between accountability failures and trust in democracy.

It is about time for the governance world to more realistically face not only the potentials but also the pitfalls of open data role in countering corruption, particularly at the subnational level.

REFERENCES

- Acosta, S. I., & Plata Pineda, O. J. (2016). Participación, colaboración, transparencia y datos abiertos. Tras la definición de gobierno abierto. *Teuken Bidikay*, 7(9), 209–222. <https://revistas.elpoli.edu.co/index.php/teu/article/view/1016>
- Acus, A. (2012). Kaip ištrūkti iš korupcijos aklavietės? Čilės sėkmės istorija ir Argentinos agonija. *Politologija*, 3(67), 143-200. <https://doi.org/10.15388/Polit.2012.3.914>
- Adam, I., & Fazekas, M. (2021). Are emerging technologies helping win the fight against corruption? A review of the state of evidence. *Information Economics and Policy*, 57, 100950. <https://doi.org/10.1016/j.infoecopol.2021.100950>.
- Adu, K. K. (2018). The paradox of the right to information law in Africa. *Government Information Quarterly*, 35(4), 669-674. <https://doi.org/10.1016/j.giq.2018.10.003>
- Aguirre-Ochoa, J., & Gómez, M. (2021). Institutional Weakness and Anti-Crime Experience in Mexico. *URVIO Revista Latinoamericana de Estudios de Seguridad*, (29), 45-57. <https://doi.org/10.17141/urvio.29.2021.4407>
- Albuquerque, J., Anderson, L., Calvillo, N., Coaffee, J., Cunha, M. A., Degrossi, L. C., Dolif, G., Horita, F., Klonner, C., Lima-Silva, F., Marchezini, V., Martins, M. H., Pajarito-Grajales, D., Pitidis, V., Rudorff, C., Tkacz, N., Traijber, R., & Zipf, A. (2021). The role of data in transformations to sustainability: A critical research agenda. *Current Opinion in Environmental Sustainability*, 49, 153–163. <https://doi.org/10.1016/j.cosust.2021.06.009>
- Altayar, M.S. (2018). Motivations for open data adoption: An institutional theory perspective. *Government Information Quarterly*, 35(4), 633-634. <https://doi.org/10.1016/j.giq.2018.09.006>
- Amundsen, I. (Ed.). (2019). *Political corruption in Africa: Extraction and power preservation*. Edward Elgar Publishing. <https://doi.org/10.1002/pa.2180>
- Ang, Y. Y. (2020). Unbundling corruption: Revisiting six questions on corruption. *Global Perspectives*, 1(1). <https://doi.org/10.1525/gp.2020.12036>

- Bachtiar, A., Suhardi, & Muhamad, W. (2020, October 19-23). *Literature review of open government data* [Conference paper]. 2020 International Conference on Information Technology Systems and Innovation (ICITSI), Bandung, Indonesia.
<http://dx.doi.org/10.1109/ICITSI50517.2020.9264960>
- Bader, M. (2012). Trends and Patterns in Electoral Malpractice in Post-Soviet Eurasia. *Journal of Eurasian Studies*, 3(1), 49–57. <https://doi.org/10.1016/j.euras.2011.10.006>
- Barr, A., & Serra, D. (2010). Corruption and culture: An experimental analysis. *Journal of Public Economics*, 94(11-12), 862-869. <https://doi.org/10.1016/j.jpubeco.2010.07.006>
- Bernhard, M., Jung, D-J., Tzelgov, E., Coppedge, M., & Lindberg, S. I. (2017). Making embedded knowledge transparent: How the V-Dem dataset opens new vistas in civil society research. *Perspectives on Politics*, 15(2), 342-360.
<https://doi.org/10.1017/S1537592717000056>
- Binder, A. (2019). All exclusive: The politics of offshore finance in Mexico. *Review of International Political Economy*, 26(2), 313-336.
<https://doi.org/10.1080/09692290.2019.1567571>
- Borges, L. B. O., & Silva, J. M. P. D. (2019). Corrupção urbanística em sistemas de planejamento e gestão urbanos: um estudo de caso. urbe. *Revista Brasileira de Gestão Urbana*, 11, e20180056. <https://doi.org/10.1590/2175-3369.011.002.AO06>
- Bovens, M., Goodin, R. E., & Schillemans, T. (Eds.). (2014). *The Oxford handbook public accountability*. Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199641253.001.0001>
- Bowen, G. A. (2006). Grounded theory and sensitizing concepts. *International Journal of Qualitative Methods*, 5(3), 12-23. <https://doi.org/10.1177/160940690600500304>
- Bowen, G. A. (2008). Naturalistic inquiry and the saturation concept: a research note. *Qualitative Research*, 8(1), 137-152. <https://doi.org/10.1177/1468794107085301>
- Brandt, R., Affonso, J., & Macedo, F. (2015, 26 Maio). Cerveró é condenado em 2ª sentença contra a antiga diretoria da Petrobrás. *O Estado de São Paulo*, 136(44416), A4.

- Braunschweig, K., Eberius, J., Thiele, M., & Lehner, W. (2012, April 16-20). *The State of Open Data - Limits of Current Open Data Platforms* [Short paper]. 21st World Wide Web Conference 2012, Web Science Track at WWW'12, Lyon, France.
https://wwwdb.inf.tu-dresden.de/opendatasurvey/www2012_short.pdf
- Bukhtiarova, A., Dukhno, Y., Kulish, G., Kurochkina, I., & Lypchanskyi, V. (2019). Ensuring transparency of key public finance authorities. *Investment Management and Financial Innovations*, 16(2), 128-139. [http://dx.doi.org/10.21511/imfi.16\(2\).2019.11](http://dx.doi.org/10.21511/imfi.16(2).2019.11)
- Bvuma, S., & Joseph, B.K. (2019). Empowering Communities and Improving Public Services Through Open Data: South African Local Government Perspective. In M. Rodríguez Bolívar, K. Bwalya, & C. Reddick, C. (Eds). *Governance Models for Creating Public Value in Open Data Initiatives* (pp. 141-160). Springer. https://doi.org/10.1007/978-3-030-14446-3_7
- Chayes, S. (2017). *When corruption is the operating system: the case of Honduras*. Carnegie Endowment for International Peace.
https://carnegieendowment.org/files/Chayes_Corruption_Final_updated.pdf
- Chen, P. L. (2019). Follow the Money: The Buck Stops Where?: A Historical Analysis of Transparency and Campaign Finance Law in Taiwan (1935-2004). *National Taiwan University Law Review*, 14(1), 101-156.
<http://dx.doi.org/10.3966/181263242019031401003>
- Chornous, G., & Iarmolenko, I. (2019, October 8-11). *The Influence of an Information Environment on Consumer Behavior in E-Commerce*. 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T) (pp. 359-363). Kyiv, Ukraine.
<https://doi.org/10.1109/PICST47496.2019.9061479>
- Chuah, L. L., Loayza, N., & Myers, C. B. (2020). The fight against corruption: taming tigers and swatting flies. *Research and Policy Briefs*, (27), 1-4.
<http://hdl.handle.net/10986/33171>
- Cloutier, C., & Ravasi, D. (2021). Using tables to enhance trustworthiness in qualitative research. *Strategic Organization*, 19(1), 113-133.
<https://doi.org/10.1177/1476127020979329>

- Coates, K. (2019). Earthquake Losses in Schools: a review of events from 1988 to 2018. *Geophysical Research Abstracts*, 21, EGU2019-14444.
<https://meetingorganizer.copernicus.org/EGU2019/EGU2019-14444.pdf>
- Coelho, T. R., Silva, T. A. B., Cunha, M. A., & Teixeira, M. A. C. (2018). Transparência governamental nos estados e grandes municípios brasileiros: uma “dança dos sete véus” incompleta?. *Cadernos Gestão Pública e Cidadania*, 23(75), 235-260.
<http://dx.doi.org/10.12660/cgpc.v23n75.73447>
- Collier, D., & Mahon, J. E. (1993). Conceptual “stretching” revisited: Adapting categories in comparative analysis. *American Political Science Review*, 87(4), 845-855.
<https://doi.org/10.2307/2938818>
- Cook, K. S., & Emerson, R. M. (1978). Power, equity and commitment in exchange networks. *American sociological review*, 43(5), 721-739. <https://doi.org/10.2307/2094546>
- Cunliffe, A. L. (2011). Crafting qualitative research: Morgan and Smircich 30 years on. *Organizational Research Methods*, 14(4), 647-673.
<https://doi.org/10.1177/1094428110373658>
- Cunha, M. A. V. C. da, Rosina, M. S. G., Teixeira, M. A. C., Silva, A. P. da, Lazzari, E. A., Silva, M. C. F. da, Karolczak, R. M., Cantoni, S.L., Coelho, T. R., Silva, T. A. B. da, Spinola, L., Marinho, L., & Scheliga, N. R. (2015). *Projeto Democracia Digital. Volume 3: Dados Abertos nos municípios, estados e governo federal brasileiros*. CEAPG-Centro de Estudos em Administração Pública e Governo (EAESP) e GEPI-Grupo de Ensino e Pesquisa em Inovação (Direito SP).
<https://bibliotecadigital.fgv.br/dspace/handle/10438/16373>
- De Graaf, G. (2007). Causes of corruption: Towards a contextual theory of corruption. *Public Administration Quarterly*, 31(1/2), 39-86.
<https://research.vu.nl/ws/portalfiles/portal/2286235/Causes+of+Corruption+++Toward+s+a+Contextual+Theory+of+Corruption.pdf>
- De Willebois, E. V. D. D., Halter, E. M., Harrison, R. A., Park, J. W., & Sharman, J. C. (2011). *The Puppet Masters: How the Corrupt Use Legal Structures to Hide Stolen Assets and What to Do About It*. World Bank Group. <https://doi.org/10.1596/978-0-8213-8894-5>

- Di Mascio, F., Maggetti, M., & Natalini, A. (2020). Exploring the dynamics of delegation over time: Insights from Italian anti-corruption agencies (2003–2016). *Policy Studies Journal*, 48(2), 367-400. <https://doi.org/10.1111/psj.12253>
- Díez Garrido, M., Campos Domínguez, E., & Calvo, D. (2021). La transparencia de los partidos políticos como estrategia electoral. Una evaluación de sus promesas y sus páginas web. *Tripodos*, (44), 83-104. <https://doi.org/10.51698/tripodos.2019.44p83-104>
- Dincer, O. C., & Johnston, M. (2020, November 6). Pas de Deux of Illegal and Legal Corruption in America. *Institute for Corruption Studies Working Papers*, (2), 1-36. Available at SSRN 3726358. <https://dx.doi.org/10.2139/ssrn.3726358>
- Doshi, S., & Ranganathan, M. (2019). Towards a critical geography of corruption and power in late capitalism. *Progress in Human Geography*, 43(3), 436-457. <https://doi.org/10.1177/0309132517753070>
- Eco, U. (2011). *Apocalípticos e Integrados*. Debolsillo.
- Erlingsson, G.Ó., Bergh, A., & Sjölin, M. (2008). Public Corruption in Swedish Municipalities: Trouble Looming on the Horizon? *Local Government Studies*, 34(5), 595-608. <https://doi.org/10.1080/03003930802413780>
- Farha, L., & UN Human Right Council (2017). *Report of the Special Rapporteur on Adequate Housing as a Component of the Right to an Adequate Standard of Living, and on the Right to Non-discrimination in This Context*. United Nation, Human Right Council. <https://digitallibrary.un.org/record/861179>
- Fazekas, M., & Tóth, I. J. (2018). From Corruption to State Capture: A New Analytical Framework with Empirical Applications from Hungary. *Political Research Quarterly*, 69(2), 320-334. <https://doi.org/10.1177/1065912916639137>
- Filgueiras, F. (2009). A tolerância à corrupção no Brasil: uma antinomia entre normas morais e prática social. *Opinião Pública*, 15(2), 386-421. <https://doi.org/10.1590/S0104-62762009000200005>

- Fox, J. (2007). The uncertain relationship between transparency and accountability. *Development in practice*, 17(4-5), 663-671.
<https://doi.org/10.1080/09614520701469955>
- Gonçalves, V. B., & Andrade, D. M. (2019). A corrupção na perspectiva durkheimiana: um estudo de caso da Operação Lava Jato. *Revista de Administração Pública*, 53(2), 271-290. <https://doi.org/10.1590/0034-761220180192>
- Granickas, K. (2014). Open data as a tool to fight corruption. *European Public Sector Information Platform Topic Report No. 2014/4*.
https://data.europa.eu/sites/default/files/report/2014_open_data_as_a_tool_to_fight_corruption.pdf
- Gritzalis, A., Tsohou, A., & Lambrinoudakis, C. (2017). Transparency-Enabling Systems for Open Governance: Their Impact on Citizens' Trust and the Role of Information Privacy. In S. Katsikas & V. Zorkadis, (Eds.). *E-Democracy – Privacy-Preserving, Secure, Intelligent E-Government Services. e-Democracy 2017* (pp. 47-63). Springer.
https://doi.org/10.1007/978-3-319-71117-1_4
- Gross, J., Leib, M., Offerman, T., & Shalvi, S. (2018). Ethical Free Riding: When Honest People Find Dishonest Partners. *Psychological Science*, 29(12), 1956–1968.
<https://doi.org/10.1177/0956797618796480>
- Grossi, G., & Pianezzi, D. (2018). The new public corruption: Old questions for new challenges. *Accounting Forum*, 42(1), 86-101.
<https://doi.org/10.1016/j.accfor.2016.05.002>
- Hall, P. A., & Taylor, R. C. (1996). Political science and the three new institutionalisms. *Political Studies*, 44(5), 936-957. <https://doi.org/10.1111/j.1467-9248.1996.tb00343.x>
- Hellman, J. S., Jones, G., & Kaufmann, D. (2003). Seize the state, seize the day: state capture and influence in transition economies. *Journal of comparative economics*, 31(4), 751-773. <https://doi.org/10.1016/j.jce.2003.09.006>
- Herala, A., Vanhala, E., Porras, J., & Kärri, T. (2016, July 13-15). *Experiences about opening data in private sector: A systematic literature review* [Conference paper]. 2016 SAI computing conference (SAI), London, UK. <https://doi.org/10.1109/SAI.2016.7556060>

- Heywood, P. M. (2017). Rethinking corruption: Hocus-pocus, locus and focus. *The Slavonic and East European Review*, 95(1), 21-48.
<https://doi.org/10.5699/slaveasteurorev2.95.1.0021>
- Hood, C., & Heald, D. (2006). *Transparency: The key to better governance?* (Vol. 135). Oxford University Press.
- Hulstijn, J., Darusalam, D., & Janssen, M. (2017). Open Data for Accountability in the Fight Against Corruption. In M. Baldoni, C. Baroglio, & R. Micalizio (Eds.), *Computational Accountability and Responsibility in Multiagent Systems (CARE-MAS 2017)* (Vol. 2051, pp. 52-66). CEUR Workshop Proceedings. <http://ceur-ws.org/Vol-2051/paper5.pdf>
- Jakóbič, A. (2020). Stackelberg game modeling of cloud security defending strategy in the case of information leaks and corruption. *Simulation Modelling Practice and Theory*, 103, 102071. <https://doi.org/10.1016/j.simpat.2020.102071>
- Jancsics, D. (2019). Corruption as Resource Transfer: An Interdisciplinary Synthesis. *Public Administration Review*, 79(4), 523-537. <https://doi.org/10.1111/puar.13024>
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information systems management*, 29(4), 258-268. <https://doi.org/10.1080/10580530.2012.716740>
- Johnsøn, J. (2012). *Theories of change in anticorruption work: A tool for programme design and evaluation*. U4 Anti-Corruption Resource Centre; Chr. Michelsen Institute. U4 Issue, 2012, (6). <https://www.u4.no/publications/theories-of-change-in-anti-corruption-work-a-tool-for-programme-design-and-evaluation>
- Kakileti, S. T., Manjunath, G., Dekker, A., & Wee, L. (2020). Robust Estimation of Breast Cancer Incidence Risk in Presence of Incomplete or Inaccurate Information. *Asian Pacific journal of cancer prevention: APJCP*, 21(8), 2307–2313.
<https://doi.org/10.31557/APJCP.2020.21.8.2307>
- Karklins, R. (2005). *The system made me do it: corruption in post-communist societies*. Routledge.

- Kaufmann, D., & Vicente, P. C. (2011). Legal corruption. *Economics & Politics*, 23(2), 195-219. <https://doi.org/10.1111/j.1468-0343.2010.00377.x>
- Klitgaard, R. (1991). Political Corruption: Strategies for Reform. *Journal of Democracy*, 2(4), 86-100. <https://doi.org/10.1353/jod.1991.0058>
- Kitchin, R. (2021). *The Data Revolution: A critical analysis of big data, open data and data infrastructures* (2nd ed.). Sage.
- Köbis, N. C., van Prooijen, J.-W., Righetti, F., & Van Lange, P. A. M. (2017). The Road to Bribery and Corruption: Slippery Slope or Steep Cliff? *Psychological Science*, 28(3), 297–306. <https://doi.org/10.1177/0956797616682026>
- Kubbe, I. (2018). Europe’s “democratic culture” in the fight against corruption. *Crime, Law and Social Change*, 70, 217-240. <https://doi.org/10.1007/s10611-017-9728-9>
- Laboutková, Š. (2018). Open government partnership: unutilized potential in post-communist EU members?(Case of the Czech Republic). *Innovation: The European Journal of Social Science Research*, 31(3), 350-376. <https://doi.org/10.1080/13511610.2017.1415803>
- Lambsdorff, J. G. (2003). How corruption affects productivity. *Kyklos*, 56(4), 457-474. <https://doi.org/10.1046/j.0023-5962.2003.00233.x>
- Langley, A., & Abdallah, C. (2015). Templates and turns in qualitative studies of strategy and management. In G. B. Dagnino & M.C. Cinici (Eds.), *Research methods for strategic management* (1st ed., pp. 155-184). Routledge. <https://doi.org/10.4324/9780203676615>
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511815355>
- Leontieva, L.S., Khalilova, T.V., Gaynullina, L.F, & Khalilov, A. I. (2015). Social-Communicative Innovations in Anti-Corruption Activities (Regional Aspect). *Asian Social Science*, 11(7), 387-393. <https://doi.org/10.5539/ass.v11n7p387>

- Lin, C., Adolphs, R., & Alvarez, R. M. (2018). Inferring Whether Officials Are Corruptible From Looking at Their Faces. *Psychological Science*, 29(11), 1807–1823.
<https://doi.org/10.1177/0956797618788882>
- Lincoln, Y.S., & Guba, E.G. (1985) *Naturalistic inquiry*. Sage.
- Lindstedt, C., & Naurin, D. (2010). Transparency is not Enough: Making Transparency Effective is Reducing Corruption. *International Political Science Review*, 31(3), 301-322. <https://doi.org/10.1177/0192512110377602>
- Lopez, S. (2017). Is Transparency an Effective Way of Reducing Corruption? Evidence Around the World. *International Journal of Economic Perspectives*, 11(2), 186-191.
- Lourenço, R. P. (2016). Evidence of an Open Government Data Portal Impact on the Public Sphere. *International Journal of Electronic Government Research (IJEGR)*, 12(3), 21-36. <http://doi.org/10.4018/IJEGR.2016070102>
- Lührmann, A., Marquardt, K. L., & Mechkova, V. (2020). Constraining governments: New indices of vertical, horizontal, and diagonal accountability. *American Political Science Review*, 114(3), 811-820. <https://doi.org/10.1017/S0003055420000222>
- Madison, J. (1961). *The Federalist Papers*: No. 51. New American Library.
- Mahoney, J. (2000). Path dependence in historical sociology. *Theory and society*, 29(4), 507-548. <https://doi.org/10.1023/A:1007113830879>
- March, J. G., & Olsen, J. P. (1983). The New Institutionalism: Organizational Factors in Political Life. *American Political Science Review*, 78(3), 734-749.
<https://doi.org/10.2307/1961840>
- Marquette, H., & Peiffer, C. (2015). *Corruption and Collective Action* (Research Paper No. 32). U4 Anti-Corruption Resource Centre; Developmental Leadership Program, University of Birmingham. <https://www.u4.no/publications/corruption-and-collective-action.pdf>
- Marquette, H., & Pfeiffer, C. (2020). *Corruption Functionality Framework*. Global Integrity. <https://ace.globalintegrity.org/wp-content/uploads/2021/01/GI-ACE-Research-Paper-Corruption-Framework.pdf>

- Martini, M. (2017). *Doors wide open: Corruption and real estate in four key markets*. Transparency International.
https://images.transparencycdn.org/images/2017_DoorsWideOpen_EN.pdf
- Matheus, R., & Janssen, M. (2020). A systematic literature study to unravel transparency enabled by open government data: The window theory. *Public Performance & Management Review*, 43(3), 503-534. <https://doi.org/10.1080/15309576.2019.1691025>
- McGee, R. (2019). Rethinking accountability: a power perspective. In R. McGee & J. Pettit (Eds.), *Power, empowerment and social change* (1st ed., pp. 50-67). Routledge.
<https://doi.org/10.4324/9781351272322>
- Mechkova, V., Bernhard, M., & Lührmann, A. (2019). *Diagonal Accountability and Development Outcomes*. V-Dem Institute; Open Government Partnership.
<https://www.opengovpartnership.org/wp-content/uploads/2019/09/Diagonal-Accountability-Development-Outcomes20190909.pdf>
- Mendoza, R. U., & Ko, A. R. A. (2015). Preliminary insights from the Philippine Bureau of Customs imports database. *World Customs Journal*, 9(1), 83-90.
[https://worldcustomsjournal.org/Archives/Volume%209%2C%20Number%201%20\(Mar-Apr%202015\)/WCJ_V9N1%20Mendoza%20and%20Ko.pdf](https://worldcustomsjournal.org/Archives/Volume%209%2C%20Number%201%20(Mar-Apr%202015)/WCJ_V9N1%20Mendoza%20and%20Ko.pdf)
- Meijer, A. (2013). Understanding the complex dynamics of transparency. *Public administration review*, 73(3), 429-439. <https://doi.org/10.1111/puar.12032>
- Meza, O., & Pérez-Chiqués, E. (2021). Corruption consolidation in local governments: A grounded analytical framework. *Public Administration*, 99(3), 530-546.
<https://doi.org/10.1111/padm.12698>
- Michener, G., & Bersch, K. (2013). Identifying transparency. *Information Polity*, 18(3), 233-242. <https://doi.org/10.3233/IP-130299>
- Mills, C. W. (1959). The promise. In W.C. Mills, *The Sociological Imagination*. Oxford University Press. <https://sites.middlebury.edu/utopias/files/2013/02/The-Promise.pdf>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group (2009). Reprint--preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Physical therapy*, 89(9), 873–880. <https://doi.org/10.1093/ptj/89.9.873>

- Moloney, K., & Chu, H. Y. (2016). Linking Jamaica's public service motivations and ethical climate. *The American Review of Public Administration*, 46(4), 436-458.
<https://doi.org/10.1177/0275074014557022>
- Mungiu-Pippidi, A. (2013). Controlling corruption through collective action. *Journal of Democracy*, 24(1), 101-115. <https://doi.org/10.1353/jod.2013.0020>
- Murillo, M. J. (2015). Evaluating the role of online data availability: The case of economic and institutional transparency in sixteen Latin American nations. *International Political Science Review*, 36(1), 42-59. <https://doi.org/10.1177/0192512114541163>
- Nikiforova, A., & McBride, K. (2021). Open government data portal usability: A user-centred usability analysis of 41 open government data portals. *Telematics and Informatics*, 58, 101539. <https://doi.org/10.1016/j.tele.2020.101539>
- Nowak, W. (2017, September 15-16). *Economic Development in Asian Least Developed Countries* [Proceedings]. In M. Cingula, M. Przygoda, & K. Detelj (Eds.). Economic and Social Development 23rd International Scientific Conference on Economic and Social Development (pp. 128-137). Madrid, Spain. https://www.esd-conference.com/upload/book_of_proceedings/Book_of_Proceedings_esdMadrid_2017_Online.pdf
- Nowak, W. (2018, May 14-16). *Development Trends in Least Developed Countries* [Proceedings]. In N. Grünwald & M. Zakrzewska (Eds.). Proceedings of the 5th International Scientific Conference on Modern Economics (pp. 41-48). University of Vigo, Spain. http://www.wneiz.pl/katedry/makro/proceedings/Vigo_Proceedings_2.pdf
- Obama, B. (2009, January 26). Presidential Documents. Freedom of Information Act. Memorandum for the Heads of Executive Departments and Agencies. *Federal Register*, 74(15), 4683-4684.
<https://www.ims.gov/sites/default/files/presidentmemorandum620.pdf>
- O'Donnell, G. A. (1998). Horizontal accountability in new democracies. *Journal of democracy*, 9(3), 112-126. <https://doi.org/10.1353/jod.1998.0051>
- Olver, C. (2021). State Capture at a Local Level. In N. Callaghan, R. Foley & M. Swilling (Eds.), *Anatomy of State Capture* (pp. 173-187). African Sun Media.

- Park, C. H., & Kim, K. (2020). E-government as an anti-corruption tool: Panel data analysis across countries. *International Review of Administrative Sciences*, 86(4), 691-707.
<https://doi.org/10.1177/0020852318822055>
- Peiffer, C., & Alvarez, L. (2016). Who will be the “principled-principals”? Perceptions of corruption and willingness to engage in anticorruption activism. *Governance*, 29(3), 351-369. <https://doi.org/10.1111/gove.12172>
- Persson, A., Rothstein, B., & Teorell, J. (2013). Why Anticorruption Reform Fails: Systemic Corruption as a Collective Action Problem. *Governance*, 26(3), 449-471.
<https://doi.org/10.1111/j.1468-0491.2012.01604.x>
- Persson, A., Rothstein, B., & Teorell, J. (2018). Getting the basic nature of systemic corruption right: a reply to Marquette and Peiffer. *Governance*, 32(4), 799-810.
<https://doi.org/10.1111/gove.12403>
- Pessanha, C. (2009). Controle externo: a função esquecida do Legislativo no Brasil. In L. F. Schwartzman, I. F. Schwartzman, F. F. Schwartzman, & M. L. Schwartzman (Eds.), *O sociólogo e as políticas públicas: ensaios em homenagem a Simon Schwartzman* (pp. 243-258). FGV. http://www.schwartzman.org.br/simon/fest12_charles.pdf
- Pierson, P. (1997). *Path dependence, increasing returns, and the study of politics*. Minda de Gunzburg Center for European Studies, Harvard University.
https://ces.fas.harvard.edu/uploads/files/Working-Papers-Archives/PSGE_WP7_9.pdf
- Pinho, J. A. G. D., & Sacramento, A. R. S. (2009). Accountability: já podemos traduzi-la para o português?. *Revista de Administração Pública*, 43(6), 1343-1368.
<https://bibliotecadigital.fgv.br/ojs/index.php/rap/article/view/6898>
- Porumbescu, G., Meijer, A., & Grimmelikhuijsen, S. (2022). *Government Transparency: State of the Art and New Perspectives*. Cambridge University Press.
<https://doi.org/10.1017/9781108678568>
- Pozsgai-Alvarez, J., 2020. The abuse of entrusted power for private gain: meaning, nature and theoretical evolution. *Crime, Law and Social Change*, 74, 433-455.
<https://link.springer.com/article/10.1007/s10611-020-09903-4>

- Pyman, M., & Heywood, P. M. (2020). *The Sector Focus & Reformulation Approach* (SFRA). Global Integrity. <https://curbingcorruption.com/wp-content/uploads/2020/12/210111-Pyman-and-Heywood-2020-Sector-Focus-and-Reformulation-Approach.pdf>
- Relly, J. E., Rabbi, M. F., Sabharwal, M., Pakanati, R., & Schwalbe, E. H. (2020). More than a decade in the making: A study of the implementation of India's Right to Information Act. *World Development*, 136, 105088. <https://doi.org/10.1016/j.worlddev.2020.105088>
- Remeur, C. (2019). *Understanding money laundering through real estate transactions* [Briefing]. European Parliamentary Research Service (EPRS). https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI%282019%29633154
- Ridgway, J. (2016). Implications of the Data Revolution for Statistics Education. *International Statistical Review*, 84(3), 528-549. <https://doi.org/10.1111/insr.12110>
- Rose-Ackerman, S. (2013). *Corruption: A study in political economy*. Academic press.
- Royo-Montañés, S., & Benítez-Gómez, A. (2019). Portales de datos abiertos. Metodología de análisis y aplicación a municipios españoles. *Profesional de la Información*, 28(6), 1-13. <https://doi.org/10.3145/epi.2019.nov.09>
- Schedler, A. (2008). *¿Qué es la rendición de cuentas?* Instituto Federal de Acceso a la Información Pública (IFAI). <https://infocdmx.org.mx/capacitacion/documentos/JURIDICO08/LECTURAS/MODULO%202/RENDICIONDECUENTAS.pdf>
- Schnell, S. (2018). From information to predictability: transparency on the path to democratic governance. The case of Romania. *International Review of Administrative Sciences*, 84(4), 692-710. <https://doi.org/10.1177/0020852316648756>
- Schuster, C., Meyer-Sahling, J. H., & Mikkelsen, K. S. (2020). (Un)principled principals, (un)principled agents: The differential effects of managerial civil service reforms on corruption in developing and OECD countries. *Governance: An international journal of Policy, Administration, and Institutions*, 33(4), 829-848. <https://doi.org/10.1111/gove.12461>

- Skocpol, T., Evans, P., & Rueschemeyer, D. (1985). *Bringing the state back in*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511628283>
- Small, M. L. (2011). How to conduct a mixed methods study: Recent trends in a rapidly growing literature. *Annual Review of Sociology*, 37, 57-86.
<https://doi.org/10.1146/annurev.soc.012809.102657>
- Sriyakul, T., & Jermstittiparsert, K. (2020). Role of E-Government and open Data to Enhance Transparency. *Journal of Security and Sustainability Issues*, 10(Oct), 337-347.
[http://doi.org/10.9770/jssi.2020.10.Oct\(26\)](http://doi.org/10.9770/jssi.2020.10.Oct(26))
- Stake, R.E. (1998). Case Studies. In N. K. Denzin & Y.S. Lincoln (Eds.). *Strategies of Qualitative Inquiry* (pp. 455-454). Sage Publications.
- Stevenson, A. & Forsythe, M. (2020, August 12) *Luxury Homes Tie Chinese Communist Elite to Hong Kong's Fate*. The New York Times.
<https://www.nytimes.com/2020/08/12/business/china-hong-kong-elite.html>
- Strach, P., Sullivan, K., & Pérez-Chiqués, E. (2019). The garbage problem: Corruption, innovation, and capacity in four American cities, 1890–1940. *Studies in American Political Development*, 33(2), 209-233. <https://doi.org/10.1017/S0898588X19000087>
- Strauss, A. L., & Corbin, J. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Sage Publications.
- Strauss, A.L., & Corbin, J.M. (2018). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Sage Publications
- Taylor, L., & Schroeder, R. (2015). Is bigger better? The emergence of big data as a tool for international development policy. *GeoJournal*, 80, 503-518.
<https://doi.org/10.1007/s10708-014-9603-5>
- Timmermans, S., & Tavory, I. (2012). Theory Construction in Qualitative Research: From Grounded Theory to Abductive Analysis. *Sociological Theory*, 30(3), 167-186.
<https://doi.org/10.1177/0735275112457914>
- Toro-García, A. F., Gutiérrez-Vargas, C. C., & Correa-Ortiz, L. C. (2020). Estrategia de gobierno digital para la construcción de Estados más transparentes y proactivos.

- Trilogía Ciencia Tecnología Sociedad*, 12(22), 71-102.
<https://doi.org/10.22430/21457778.1235>
- Torsello, D., & Venard, B. (2016). The Anthropology of Corruption. *Journal of Management Inquiry*, 25(1), 34-54. <https://doi.org/10.1177/1056492615579081>
- Trantidis, A., & Tsagkroni, V. (2017). Clientelism and corruption: Institutional adaptation of state capture strategies in view of resource scarcity in Greece. *The British Journal of Politics and International Relations*, 19(2), 263-281.
<https://doi.org/10.1177/1369148117700658>
- Trautvetter, C. (2020, May 11). *Why many Berlin real state owners remain secret despite new transparency laws*. Tax Justice Network. <https://taxjustice.net/2020/05/11/fatf-anteportas-why-many-berlin-real-estate-owners-remain-anonymous-despite-new-transparency-laws/>
- UNESCO; Mendel, T.; & Notess, L. (2020). *The Right Information in Times of Crisis: Access to Information – Saving lives, Building Trust, Bringing Hope!* UNESCO.
<https://unesdoc.unesco.org/ark:/48223/pf0000374369>
- Valle-Cruz, D. (2019). Public value of e-government services through emerging technologies. *International Journal of Public Sector Management*, 32(5), 530-545.
<https://doi.org/10.1108/IJPSM-03-2018-0072>
- Van der Merwe, T., & Transparency International (2020). *Beneficial ownership registers: Progress to date*. Transparency International the global coalition against corruption, 4U Anti-Corruption Resource Centre.
<https://knowledgehub.transparency.org/helpdesk/beneficial-ownership-registers-progress-to-date>
- Vannucci, A. (2015). Three paradigms for the analysis of corruption. *Labour & Law Issues*, 1(2), 1-31. <https://doi.org/10.6092/issn.2421-2695/5468>
- Vijauakumar, G., & Bharathi, R.K. (2021). Reducing Road Accidents in India by Predicting Vehicle Defects and Black Spots. In V. L. N. Komanapalli, N. Sivakumaran & S. Hampannavar (Eds.) *Advances in Automation, Signal Processing, Instrumentation*,

and Control (pp. 1651-1660). Springer Nature Singapore.
https://doi.org/10.1007/978-981-15-8221-9_154

- Villeneuve, J-P., Mugellini, G. & Heide, M. (2020) International Anti-Corruption Initiatives: a Classification of Policy Interventions. *European Journal on Criminal Policy and Research*, 26(4), 431–455. <https://doi.org/10.1007/s10610-019-09410-w>
- Vitvitskyi, S.S., Andrieieva, O. M., Stychynska, A. B., Karashchuk, M. G., & Polovko, O.O. (2020). Social communication of public authorities in the system of potential manifestations of corruption. *Revista San Gregorio*, 1(42). 61-69.
<http://dx.doi.org/10.36097/rsan.v1i42.1547>
- Watson, T. J. (1995). Rhetoric, Discourse and Argument in Organizational Sense Making: A Reflexive Tale. *Organization Studies*, 16(5), 805-821.
<https://doi.org/10.1177/017084069501600503>
- Wedel, J. R. (2012). Rethinking corruption in an age of ambiguity. *Annual Review of Law and Social Science*, 8, 453-498. <https://doi.org/10.1146/annurev.lawsocsci.093008.131558>
- World Bank (2017). *World Development Report 2017: Governance and the Law*. Washington D.C.: World Bank Group. <https://www.worldbank.org/en/publication/wdr2017>
- World Bank (2021). *World development report 2021: Data for better lives*. Washington D.C.: World Bank Group.
<https://openknowledge.worldbank.org/bitstream/handle/10986/35218/211600ov.pdf>
- Žuffová, M. (2020). Do FOI laws and open government data deliver as anti-corruption policies? Evidence from a cross-country study. *Government information quarterly*, 37(3), 101480. <https://doi.org/10.1016/j.giq.2020.101480>

ANNEXES

Annex 1 – Questionnaire to Government Officials

A - Question on attributions and ability.

1. What position do you hold in the government agency? Since when?
2. Do you or your team have the attribution to initiate investigations about corruption or money laundering? Do you need any special authorization from the superior hierarchy?
3. Do you or your team use government data and information to inform your investigations? If so, how often? If not, why not?
4. Do you or your team usually request information to the government as one of the initial steps in an investigation? Or do you look for open-source data before any information request?
5. Do you sense that you and your team are adequately trained to read official data, including statistics and quantitative data?

B - Question on the knowledge about the dataset

6. Are you and your team aware of critical datasets that could be used to inform and enrich investigations and possible future sanctions?
7. Are you and your team aware of the dataset on real estate and urban land taxation published by the City Government of Sao Paulo in 2016? If so, when did this come to your attention and knowledge? If not, why do you think this information has not reached you or your team?

C - Question on data interpretation

8. Have you been using open-source data for a long time? If so, how long? If not, why not?
9. How would you describe your and your team's ability to interpret data and datasets? Regular? Good? Excellent?
10. Have you or your team ever worked on interpretation and analysis of the dataset on real estate and urban land taxation published by the City Government of Sao Paulo?
 - a. If so, can you describe this or these specific effort(s)? When and under what circumstances and context did this/these effort(s) occurred?
 - b. If not, why not?

D - Questions on data use (or barriers to usage) for accountability purposes

11. Have any of your or your team's investigations used the dataset on real estate and urban land taxation in practice?
 - a. If so, can you describe this or these specific investigation(s)? When and under what circumstances and context did this/these effort(s) occur?

- i. Did the investigation eventually come to an end? Did your institution propose formal sanctions before the Judiciary or other sanctioning agency? If so, have people been sanctioned? If not, why not?
- ii. If investigations are still ongoing, do you think they are moving on the correct path and speed? If they are too slow, where are the barriers?
- b. Do you think that the period in which the data on real estate property was disclosed, 2016, was favorable for your investigations, once in those years the anticorruption agenda was high in the public opinion?

E - Question on corruption persistence

12. Considering you and your team had information that you could potentially use to sanction corruption, why would you and your colleagues choose not to use it? What is/are the main barrier(s)? You can choose more than one option:

1. Corrupt actors who would be targeted in an investigation as such are very powerful
2. Corrupt actors' long-lasting and established network of relationships would likely mean impunity as the result, so as the benefits would not make up for the costs.
3. I assumed my institution would not provide me the level of support I would need

Annex 2 – Questionnaire to Civil Society Representatives

A - Question on attributions and ability.

- What position do you hold in the CSO you represent?
- Do you or your team use government data and information to inform your research and advocacy? If so, how often? If not, why not?
- Do you or your team usually request information to the government as one of the initial steps in research or advocacy work? Or do you look for open-source data before any information request?
- Do you sense that you and your team are adequately trained to read official data, including statistics and quantitative data?

B - Question on the knowledge about the dataset

- Are you and your team aware of critical datasets that could be used to inform and enrich research and advocacy activities?
- Are you and your team aware of the dataset on real estate and urban land taxation published by the City Government of Sao Paulo in 2016? If so, when did this come to your attention and knowledge? If not, why do you think this information has not reached you or your team?

C - Question on data interpretation

- Have you and your team been using open-source data for a long time? If so, how long? If not, why not?
- How would you describe your and your team's ability to interpret data and datasets? Regular? Good? Excellent?
- Have you or your team ever worked on interpretation and analysis of the dataset on real estate and urban land taxation published by the City Government of Sao Paulo?
 - If so, can you describe this or these specific effort(s)? When and under what circumstances and context did this/these effort(s) occurred?
 - If not, why not?
- To your knowledge, do accountability institutions based in Sao Paulo have the adequate training to interpret data and datasets?

D - Questions on data use (or barriers to usage) for accountability purposes

- Have you or your team been aware of accountability institutions effectively using the dataset on real estate and urban land taxation to move an investigation forward?
 - If so, can you describe this or these specific investigation(s)? When and under what circumstances and context did this/these effort(s) occur?
 - Did the investigation eventually come to an end? Did the institution propose formal sanctions before the Judiciary or other sanctioning agency? If so, have people been sanctioned? If not, why not?
 - If investigations are still ongoing, do you think they are moving on the correct path and speed? If they are too slow, where are the barriers?
 - Do you think that the period in which the data on real estate property was disclosed, 2016, was favorable for official investigations, once in those years the anticorruption agenda was high in the public opinion?

E - Questions on corruption persistence

○ Have you ever come across or heard reports of civil servants that even when gathering data, information, making analysis from public data, public documents, they chose not to go ahead with an investigation or a denunciation? Why in your opinion have they decided not to act? What is/are the main barrier(s)? You can choose more than one option:

- Corrupt actors who would be targeted in an investigation as such are very powerful
- Corrupt actors' long-lasting and established network of relationships would likely mean impunity as the result, so as the benefits would not make up for the costs.
- I assumed my institution would not provide me the level of support I would need